From proteomics to biotechnology. Using the resurrection plant *Eragrostis nindensis* to genetically engineer drought tolerant crops.

By: Llewelyn van der Pas

Dissertation presented for degree of Doctor of Philosophy

In the Department of Molecular and Cell Biology

Faculty of Science

University of Cape Town

Supervisors:

Prof. Jill Farrant[1]

Associate Professor Henk Hilhorst[1,2]

Associate Prof. Inga Hitzeroth[1,3]

Associate Prof. Suhail Rafudeen[1]

1: University of Cape Town

2: Wageningen University and Research

3: Biopharming Research Unit, University of Cape Town

Abstract

Global climate change is increasingly putting pressure on finding innovative solutions to ensure future food security in particular to developing African nations. Of great relevance are regionally adapted crops, known as orphan crops, which tend to have very little economic value but can provide a source of alternative food security. Vegetative desiccation tolerance is a remarkable feat of selective evolution and is only present in a small number of angiosperms. The ability of some plants, such as *Eragrostis nindensis* to survive complete cellular water deficit provides an attractive model for discovery-based omics to not only understand the mechanisms involved in driving desiccation tolerance but to explore the feasibility of potential target genes for orphan crop improvement. The work presented herein was aimed at complementing a transcriptomic study using the same leaf tissue from that study to evaluate the changes from RNA to protein and to determine whether there were proteomic signatures that could differentiate the desiccation-tolerant non-senescent (NST) leaves from the desiccation-sensitive senescent (ST) leaves. The data presented here illustrate that several important metabolic pathways are significantly reprogrammed, that only a small subset of proteomic-matching transcripts were translated, and that proteomic differences between the NST and ST were noted despite their being significant similarities between the two in general oxidative and osmotic stress. For instance, the prevention of ferroptosis and accumulation of raffinose synthase and starch synthase in the NST exclusively illustrated that small and subtle increases in protein abundance are likely responsible for enabling resurrection in the NST and not in the ST, which we hypothesise here is likely due to sacrificing of ST upon rehydration as a means to act as a source of nutrition for the NST during resurrection. The study also focussed on functional characterisation of a heat shock 70 protein from *E. nindensis* as a target for genetic engineering. The selected *EnHSP70* was shown to localise to the chloroplast and was able to undergo liquid-liquid phase separation *in vitro* in a protein concentration and polyethylene glycol dependent manner which could have broad impacts on its role in maintaining proteostasis. In *Arabidopsis thaliana*, overexpression of *EnHSP70* resulted in a stunted germination phenotype whereas expression in BL21 *Escherichia coli* did not enhance tolerance towards salt or mannitol stress. Furthermore, incubation of EnHSP70 with lactate dehydrogenase did not confer improved thermotolerance. Taken together, the selected HSP70 from *E. nindensis* did not appear to be involved in stress response and is likely involved with general proteostasis. Lastly, a method for generating embryonic calli from *Eragrostis tef* is presented with the goal of using this developed protocol for the genetic improvement of the Ethiopian orphan crop.

Table of Contents

1 Chapter 1: Scope of research……..1

1.1 Global climate change and food security....1

1.1.1 Global climate change....1

1.1.2 Future food security....3

1.2 Drought and resurrection plants....4

1.2.1 Drought stress....5

1.2.2 Drought response....8

1.3 Orphan crops for food security....9

1.3.1 *Eragrostis tef* (Zucc.) Trotter....10

1.3.2 *Eragrostis nindensis* as a model for desiccation and senescence....10

1.4 Research aims and objectives:....12

1.4.1 Aim 1: Investigating the proteome for *E. nindensis.* 12

1.4.2 Aim 2: Functional characterisation of HSP70 from *E. nindensis*. 12

1.4.3 Aim 3: Developing a method for *E. tef* transformation and regeneration. 12

2 Chapter 2: Proteomic investigation of *Eragrostis nindensis* leaf tissue*:* unravelling proteomic signatures of desiccation tolerance……..13

2.1 Introduction....13

2.1.1 *S*enescence....13

2.1.2 Physiology and ultrastructural differentiates NST and ST from one another.....14

2.1.3 The lipidome identified *E. nindensis* as a C18 plant.....15

2.1.4 The transcriptome indicated similar processes between NST and ST leaf tissue 17

2.1.5 Importance of proteomics....21

2.2 Methods....24

2.2.1 *E. nindensis* growth and sampling....24

2.2.2 Optimisation of protein isolation....25

2.2.3 Mass spectrometry and analysis....26

2.2.3.1 On-bead HILIC digest....26

2.2.3.2 Liquid Chromatography mass spectrometry 27

2.2.3.3 Data analysis....29

2.3 Results and Discussion....32

2.3.1 Protein quantification....32

2.3.2 Gene ontology enrichment 35
2.3.1 Gene ontology enrichment: General conclusions 46
2.3.2 KEGG analysis 48
2.3.2.1 Malate and pyruvate are key players in central carbohydrate metabolism. 52
2.3.2.2 *E. nindensis* is a raffinose oligosaccharide resurrection plant. 58
2.3.2.3 Hormonal regulation implies defence response and programmed senescence. 62
2.3.2.4 Amino acid metabolism is in favour of oxidative and osmotic stress mitigation 64
2.3.2.5 Lipid metabolism is directed towards degradation and protection. 69
2.3.2.6 Genetic information processes imply universal stress. 76
2.4 Conclusion 84
3 Chapter 3: A consolidated omics view of desiccation tolerance in *Eragrostis nindensis* 88
3.1 Introduction 88
3.2 Methods 88
3.3 Results and discussion 89
3.3.1 Differential senescence-associated proteins hints at suppression of senescence in the NST 89
3.3.2 Proteomic model for desiccation tolerance in *E. nindensis* 95
4 Chapter 4: Functional characterization of the heat shock protein 70.4 from *Eragrostis nindensis* 103
4.1 Introduction 103
4.2 Heat shock proteins: a review 103
4.3 Liquid-liquid phase separation 106
Methods 109
4.3.1 Bioinformatic analysis of EnHSP70 109
4.3.2 Molecular cloning 110
4.3.2.1 Amplification of EnHSP70 110
4.3.2.2 Overexpression construct 111
4.3.2.3 Liquid-liquid phase separation construct 112
4.3.2.4 Subcellular localisation construct 112
4.3.3 *A. tumefaciens* transformation 113
4.3.4 *Arabidopsis thaliana* floral dipping and selection of transformants 113
4.3.5 Liquid-Liquid phase separation assay 115
4.3.6 Subcellular localisation 116

4.3.7 Bacterial abiotic stress assays 116
4.3.8 Lactate dehydrogenase enzyme activity assay 117
4.3.9 EnHSP70 expression detection in T_2 lines 118
4.4 Results 119
4.4.1 Bioinformatic analysis of EnHSP70 119
4.4.2 Molecular cloning 124
4.4.3 *Arabidopsis thaliana* transgenic screening 125
4.4.4 mCherry::EnHSP70 expression and purification 128
4.4.5 EnHSP70 liquid-liquid phase separation 129
4.4.6 Bacterial abiotic stress assay 131
4.4.7 Lactate dehydrogenase assay 132
4.5 Discussion 133
4.6 Conclusion 137
5 Chapter 5: Developing a method for generating transgenic *Eragrostis tef* 138
5.1 Introduction 138
5.1.1 Transgenic plants 138
5.1.2 Selection of plant expression vector 139
5.1.2.1 Ti Plasmid 139
5.1.2.2 Binary vectors 141
5.1.2.3 Viral vectors 143
5.1.3 Selection of Agrobacterium strain 145
5.1.4 Selection of transformation method 146
5.1.4.1 Callus induction and somatic embryogenesis 148
5.2 Methods 150
5.2.1 *Eragrostis tef* germination and growth 150
5.2.2 Agrobacterium preparation and initial tef leaf infiltration trial 150
5.2.3 Callus induction 152
5.2.3.1 Immature embryos (IE) 152
5.2.3.1 Cotyledons and roots 152
5.2.3.1 Mature seeds 152
5.2.4 Co-cultivation of calli and Agrobacteria with viral vectors 154
5.2.5 Co-cultivation of calli and Agrobacteria with EnHSP70 155
5.2.6 Plant regeneration 155
5.3 Results 156
5.3.1 Initial tef infiltration 156

5.3.2 Callus induction and plant regeneration 157
5.3.3 Callus co-cultivation 160
5.4 Discussion 162
5.4.1 Transient expression in *E. tef* leaves is not a feasible method. 163
5.4.2 Somatic embryogenesis from *E. tef* required exhaustive optimisation to establish. 163
5.5 Conclusion 167
6 Chapter 6: Concluding thoughts and future work……169
7 References……173
8 Appendix……204
8.1 Sequences used for molecular cloning 209
8.2 EnHSP70 multiple sequence alignment 214
8.3 Global proteome for *Eragrostis nindensis* 219

List of Figures and Tables

Figure 1-1: Global temperature anomalies. 2
Figure 1-2: Photosynthetic electron transport. 6
Figure 1-3: ROI scavenging in plants.. 7
Figure 2-1: *Eragrostis nindensis* global transcriptome:. 18
Figure 2-2: E. nindensis AWC curve and tissue distinction: 25
Figure 2-3: Proteome sample concentrations: 33
Figure 2-4: Venn diagram for all DAPs: 35
Figure 2-5: Venn diagrams for all accumulating and diminishing DAPs 36
Figure 2-6: GO Terms associated with biological and molecular processes in *E. nindensis*. 38
Figure 2-7: Ferritin protein and transcripts:. 42
Figure 2-8: Over-represented cellular components in *E. nindensis*:. 43
Figure 2-9: Koala annotation and KEGG pathways for *E. nindensis* DAPs 49
Figure 2-10: Carbohydrate metabolism in *E. nindensis:* 53
Figure 2-11: Fold change for selected carbohydrate related terms:. 54
Figure 2-12: Summary of selected carbohydrate enzymes: 55
Figure 2-13: Fold change for selected starch and sucrose terms: 59
Figure 2-14: Fold change for selected defence-related terms:. 64
Figure 2-15: Amino acid metabolism in *E. nindensis*: 65
Figure 2-16: Fold change for selected amino acid metabolic terms 67
Figure 2-17: Lipid metabolism in *E. nindensis.* 70
Figure 2-18: Fold change for selected lipid metabolic terms 71
Figure 2-19: Fold change for selected Genetic Information Processing, Chaperone, and Protein Processing terms 79
Figure 3-1: Selected senescence related terms for 25% RWC 90
Figure 3-2: Drought to desiccation response in *E. nindensis* 97
Figure 3-3: Desiccated state in *E. nindensis* 99
Figure 3-4: Rehydrated state in *E. nindensis*. 100
Figure 4-1: HSP70 Chaperone model 105
Figure 4-2: Maps of EnHSP70 final constructs: 112
Figure 4-3: EnHSP70 localisation in castor protoplasts 121
Figure 4-4: EnHSP70 percent identity matrix 122
Figure 4-5: EnHSP70 disorder and phosphorylation prediction 123
Figure 4-6: Three-dimensional model for EnHSP70 124
Figure 4-7: Diagnostic digests of EnHSP70 clones 125
Figure 4-8: Arabidopsis thaliana transgenic screening. 126
Figure 4-9: T_2 PCR Screening. 127
Figure 4-10: EnHSP70 expression in T_2 plants 127
Figure 4-11: mCherry::EnHSP70 purificationl. 128
Figure 4-12: EnHSP70 liquid-liquid phase separation assay 130
Figure 4-13: Bacterial abiotic stress assay 131
Figure 4-14: Lactate dehydrogenase assay. 132
Figure 5-1: Octopine-type Ti Plasmid 140

Figure 5-2: Agrobacterium tumefaciens infection 141
Figure 5-3: A binary vector 142
Figure 5-4: Viral vector used in agroinfiltration. 144
Figure 5-5: Summary of some transformed monocotyledons. 148
Figure 5-6: eGFP fluorescence in vacuum infiltrated *E. tef* leaves 156
Figure 5-7: eGFP detection of vacuum infiltrated *E. tef*: 157
Figure 5-8: Calli from *E. tef* immature embryos and seeds. 159
Figure 5-9: Plant regeneration from somatic embryos 160
Figure 5-10: Relative eGFP fluorescence of infected calli 161
Figure 5-11: Dot blot for eGFP detection 161
Figure 8-1: *E. nindensis* TEM images: 205
Figure 8-2: Highest and lowest fold change DAPs per RWC 206
Figure 8-3: String network interaction for selected malate dehydrogenase proteins 207
Figure 8-4: EnHSP70 Amino acid composition 207
Figure 8-5: mCherry::EnHSP70 three dimensional model 208
Figure 8-6: Dot blot of EnHSP70 transient expression in tobacco 208

Table 2-1: Liquid chromatography gradient for the LFQ liquid-chromatography mass spectrometry/mass spectrometry 28
Table 2-2: Mass spectrometry data acquisition parameters 28
Table 2-3: Recorded protein search configuration 30
Table 2-4: *Eragrostis nindensis* proteome and transcriptome summary 34
Table 2-6: Summary of pathways and interaction (ligand) output from DAVID 45
Table 2-7: Highest and lowest fold changes per RWC as annotated by Koala Blast 52
Table 2-8: Summary of highest and lowest fold changes of DAPs assigned to Genetic Information Processing, Chaperone, and Protein Processing 78
Table 4-1: Primers used in study. 111
Table 4-2: EnHSP70 subcellular location predictions 120
Table 4-3: Predicted isoforms of the upregulated *HSP70* sequences from the global transcriptome 122
Table 5-1: Commonly used Agrobacterium strains in plant transformation 145
Table 5-2: Media components used during plant tissue culture 153
Table 5-3: Summary of callus induction 158
Table 8-1: *Eragrostis nindensis* isolated protein 204
Table 8-2: Summary of Spearman correlation and covariance and number of transcript and protein pairs between shared KEGG terms from the transcriptome and proteome per RWC and tissue. 206
Table 8-2: Differentially abundant proteins from KEGG Koala Blast 220

Acknowledgements

Undertaking a PhD is no mean feat and the success of it is highly influenced by the people who support you. I would firstly like to give my thanks to my main supervisor Jill, who allowed me to shape this PhD and explore aspect of our research group. She always encourages my curiosity and has always been a great mentor. As one of the leading figures in plant desiccation tolerance, it was a tremendous honour to work with her and gain insight into her brilliant mind. Henk was added mid-way of the project as a co-supervisor and his wisdom on seeds and the connection between them and vegetative desiccation tolerance is invaluable. He provided great constructive criticism of the work and greatly impacted the final product. Inga was invaluable in the troubleshooting and execution of the molecular cloning and plant transformation work presented herein. Inga also provided a great source of information on a variety of obscure transformation-related topics. In the same breath, I want to thank Ann Meyers who always availed herself to assist with Agrobacterial strains and was also a great source of plant transformation information. Thanks to Suhail Rafudeen for his insight in proteomics and assisting with the study design. To Naadirah Moola te Water Naude a huge thanks for providing a robust pipeline for the gene ontology enrichment analysis. To Bronwyn, thanks for always listening to my problems and for providing invaluable information on callus induction methods. A huge thanks to Shannon Sparks who always helped in troubleshooting various aspect of the project and for being a constant throughout this journey. To Abigail Russel and Jasmin Knopp, a huge thanks is assisting with the liquid-liquid phase separation assays. In particular, Abigail for assisting in capturing the fluorescent images. To the members of the Plant Stress Laboratory, past and present, thanks for your regular support and helping. In particular to Robyn Craythorne and Michael Wittenberg for always being a great source of support and experimental design.

To my family, thanks for your regular encouragement, in particular to my mother, thanks for always pretending to understand what I was doing. To my friends, in particular Emma Frickle, Natalie Nel, and Molopo Lipali, our little PhD group had an enormous impact on both my emotional and mental health during the last stretch of my journey. Last, but certainly not least, to my partner, Werner, the journey from undergraduate to here has been a challenging and long one but you have always been a constant source of encouragement and support.

To the Molecular and Cell Biology department thank you for providing the support and facilities to conduct the research and for challenging each one of us to expand our scientific curiosities. I also want to partially thank the National Research Foundation and UCT Postgraduate Funding Office for providing some personal funding.

Declaration

I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own. I have used the Harvard convention for citation and the Chicago convention for referencing. Each contribution to, and quotation in, this thesis from the work(s) of other people has been attributed and has been cited and referenced. This thesis is my own work and is in my own words (except where I have attributed it to others). I have not allowed and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

Llewelyn van der Pas

14 June 2023

Conventions used:

Species names where used are reflective of their name at time of the relevant publication cited. Enzyme names are based on KEGG classification at the time of writing.

“…whilst this planet has gone circling on according to the laws of gravity, from so simple a beginning endless forms most beautiful and wonderful have been and are being evolved.”

Charles Darwin, The Origin of Species

Chapter 1: Scope of research

One of the primary goals of the Plant Stress Laboratory in the Department of Molecular and Cell Biology at the University of Cape Town, is to investigate how vegetative desiccation tolerant plants also described as resurrection plants, maintain viability under complete water deficit conditions and subsequently resume metabolic activity upon rehydration. The work thus far done within our group has shed important light on the mechanisms, and ultimately the genetic factors involved in their response to drought and desiccation. As a group we are now at the precipice of applying what has been learnt to answering the difficult questions of ensuring food security for developing countries in Africa. Chapter 1 thus sets the scene for subsequent Chapters and presents a general literature review and lays the foundation for the research done herein. Chapter 2 focusses on the proteomic investigation of leaf tissue from the resurrection grass *Eragrostis nindensis* with particular emphasis on comparing the proteome with a previously done transcriptomic study on the same leaf tissue to elucidate the proteomic signatures that drive desiccation tolerance. Chapter 3 considers a small subset of differentially abundant proteins and how they may play a role in the suppression of senescence in the desiccation tolerant leaves. Chapter 3 also consolidates the main findings from Chapter 2 by considering the proteomic changes from early drought response to desiccation response and finally explores to what extent the proteome validates the transcriptome. Chapter 4 focusses on functional characterisation of a heat shock protein 70 from *E. nindensis* as a possible candidate for genetic improvement in orphan crops by evaluating its effect in the model organism, *Arabidopsis thaliana*. This Chapter also explores some characteristics of EnHSP70 such as its liquid-liquid phase separation propensity and thermotolerance. Chapter 5 focusses on the development of a method for generating transgenic *Eragrostis tef* plants via somatic embryogenesis of explant tissue which is then co-cultivated with Agrobacteria harbouring plant expression vectors. Chapter 6 considers everything presented herein and lays the foundation for future work.

Global climate change and food security

1.1.1 Global climate change

The post-industrial revolution saw the world change with a large shift towards technological innovation and societal restructuring. However, it also led to the rapid deterioration of the environment with ramifications being felt today. Early mathematical models predicted the adverse effects of increasing CO_2, yet for the post-industrial world, these were simply ignored until 1988 when it was realised that climate change was real. In 1997, governments from

across the globe met and adopted the Kyoto Protocol with the goal of reducing six greenhouse gas emissions by 5.2%. The goal was not achieved for several reasons and another global agreement was signed in Paris in 2015. To this day, and by generalisation, not much has been done to combat global warming. In fact, by looking at CO_2 data collected by the National Oceanic and Atmospheric Administration, global CO_2 has increased by 70 parts per million in 40 years. Other greenhouse gasses such as methane, sulphur hexafluoride, and nitrous oxide share similar trends (see https://www.esrl.noaa.gov/gmd/ccgg/trends/ for more detail). Anomalies associated with increases in greenhouse gasses has indicated circa the 1940s a steady increase in temperature anomalies erring on increased global temperatures is evident (Figure 1-1).

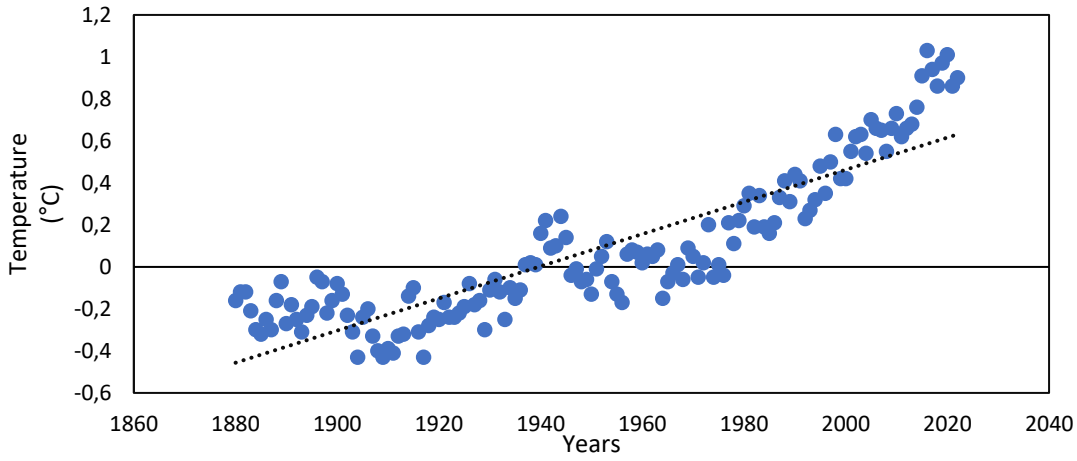


Figure 1-1: Global temperature anomalies. Global temperature anomalies from 1880 to 2023. Data from the National Centers for Environmental Information (NOAA, 2023)

It is important to distinguish global climate change (GCC) from global warming (GW). GCC can be defined as a change in climate patterns, such as precipitation or drought and can be brought about by anthropogenic activities or natural activities such as volcanism or plate tectonics, whereas GW can be defined as the increase in Earth's temperature mainly attributed by anthropogenic activities such as the burning of fossil fuels. According to data collected from the National Centers for Environmental Information (NOAA), there has been a 2°C increase in global temperature from the pre-industrial era to 2023 (Figure 1-1) and is further projected to increase even further.

By means of modelling, a 3 to 5°C increase in annual temperature over land areas is predicted to will have occurred by the end of the twenty-first century (Feng, et al., 2014). The Coupled Model Inter-comparison Project (CMIP5) model further shows that the northern hemisphere will become much warmer than the southern hemisphere (Feng, et al., 2014). This model further shows that there will be shifts in Köppen-Trewartha (K-T) climate classifications with regions such as temperate, tropical, and dry climate types predicted to likely increase in their

land cover whilst climate types such as polar, sub-polar, and subtropical are predicted to recede (Feng, et al., 2014). Under the Intergovernmental Panel on Climate Change Representative Concentration Pathway 8.5 (RCP8.5), also known as the 'business-as-usual' scenario, which takes higher greenhouse-gas emissions into account, these climate type shifts are projected to be much greater. For instance, under a low-emission scenario (RCP4.5), a change of 31.4% is projected for climate types, whereas under the RCP8.5 scenario, this change is projected to be 46.3%. Couple this with the projected decrease in precipitation, or as (Sherwood & Fu, 2014) refer to a decrease in the ratio of precipitation to potential evapotranspiration, tropical and mid-latitude regions which span across countries such as Mexico, Brazil, large areas of central Africa, India, and all the island nations between China and Australia, will see significant increases in dryland acreage.

1.1.2 Future food security

Many of the countries that span this mid-latitude geographical region are considered to be either 'developing' or 'under-developed' countries, often relying, to some degree, greatly on agriculture to drive their economy. According to the Famine Early Warning System Network (FEWS NET), at present, many countries within the tropical and mid-latitude regions are experiencing acute food insecurity. In Africa in particular, countries such as Nigeria, Zimbabwe, regions within Southern Africa, the Democratic Republic of Congo, and regions within the Horn of Africa are of concern. This illustrates that ensuring future food security is an important social and scientific endeavour. According to the 2019 State of Food Security and Nutrition in the World report from the Food and Agriculture Organization of the United Nations (FAO) there has been a reduction in the total number of people suffering from undernourishment from 947.2 million in 2005 to 811.7 million in 2017. However, when looking at specific regions such as Africa, the total number of people suffering from undernourishment was at 277 million people in 2018 (representing 21.5% of the total population of Africa). Whilst the reduction in undernourishment is tremendous, the proportion of people undergoing malnutrition is still significantly high. The FAO further provides statistics on global crop production stating that a 1.4% per annum growth in crop production from 1997/99 to 2030 is projected to occur, however this is down from the 2.1% growth of the past 30 years. Approximately 1.2 billion hectares (80%) of global cultivated area is rainfed and produces roughly 60% of global crops (Dempewolf, et al., 2017). Global population growth annually is estimated to be roughly 1.1%, or 83 million people. The population has grown from one billion in 1800 to over 7.7 billion in 2019 (Roser, et al., 2020) and is further projected to reach 8.6 billion people by 2032 (FAO). Whilst the absolute rate of annual increase has declined from

2.2% in 1962-1963 to 0.8% from 2022-2023 (FAO), the actual number of people has more than doubled since 1963 (Roser, et al., 2020).

This brings to light the triadic problem on socioeconomic development in these 'developing' or 'under-developed' countries under the banner of impending climate change: firstly, more people demand more food and by virtue of their existence, are (in)directly responsible for anthropogenically induced decay or destruction of natural resources. Secondly, more food requires the expansion of current agricultural endeavours often resulting in the devastation of already sensitive ecosystems. Lastly, less arable land is projected to be available in the imminent future for many of these already penurious countries. In essence, an endless loop is created. This naturally leads to the urgency in developing crops which are able to withstand the predicted climate type shifts (Varshney, et al., 2011).

Drought and resurrection plants

As the most abundant molecule on Earth, water is essential in sustaining life. As a polar molecule, water acts as a solvent for an array of ions and compounds and exhibits some unique physical properties that has made it a key component in maintaining life. All living cells require water for many biochemical processes, including the generation of cellular energy through oxidative 'burning' of organic molecules such as glucose. Plants, algae, fungi, and prokaryotes such as cyanobacteria are enveloped by a cell wall which can only provide structural support so long as there is a force acting upon it. Water provides such a force in the form of turgor pressure. While tracheophytes have vascular tissue that provide support to some degree, absence or irregular water acquisition can lead to xylem cavitation which thus hinders the structural component obsolete. Maintaining turgor pressure is a complex system but sufficed to say is vital in ensuring plant cell integrity and by extension, tissue, and organ viability. Not only is it required to ensure that cells remain turgid, but water is also necessary during the initial light-dependent photochemical reduction of CO_2. For most angiosperms that display desiccation sensitivity, early onset of senescence is noted with a water loss as low as 20% relative water content and can rapidly lead to cell, tissue, and organ death (Veljovic-Jovanovic, et al., 2006; Farrant, et al., 2015).

Evolution and natural selection have provided some 240 angiosperms (Marks, et al., 2021) with the remarkable ability to withstand water loss of approximately 95% (Gaff & Oliver, 2013) for periods that can span months and resume metabolic activity upon rehydration (Hoekstra, et al., 2001; Farrant, et al., 2007). These plants, coined 'resurrection' plants by Gaff (1971), represent a well-adapted subset within the plant kingdom capable of remaining viable at a water content of 0.1g $H_2O.g^{-1}$ dry mass. As far as these unique plants are concerned, two

distinct groups can be identified based on their response to drying down. Those that are characterised as homoiochlorophyllous, such as *Myrothamnus flabellifolia*, rely on curling their leaves and the accumulation of pigments along the exposed surface to protect cellular components such as the photosynthetic apparatus since this group retains thylakoid membranes and chlorophyll in their dry state (Sherwin & Farrant, 1998). The other group known as the poikilochlorophyllous resurrection plants include some well-studied grass representatives such as *Eragrostis nindensis*, the subject of this study, and other monocotyledons belonging to the *Xerophyta* genus. These resurrection plants do not maintain any chlorophyll upon dehydration, but rather dissemble chlorophyll along with the dismantling of the thylakoid membrane and rapidly resynthesise these upon rehydration. The various stresses associated with water loss in desiccation sensitive (DS) plants will be briefly compared and contrasted with desiccation tolerant (DT) resurrection plants.

1.2.1 Drought stress

During photosynthesis, H_2O is oxidised to molecular oxygen following photon excitation of the P680 reaction centre of photosystem (PS) II with concomitant electron transport occurring via plastoquinone to cytochrome complexes which are eventually transferred to $NADP^+$ via ferredoxin, schematically represented in Figure 1-2. Under normal circumstances, for every four photon P680 excitation events one molecule of molecular oxygen is generated. However, photosynthesis can become a lethal process under drought conditions. It is common for reactive oxygen species (ROS) to be generated through photosynthesis. In fact, it is such a common event that antioxidizing enzymes are present in the lumen to quench these ROS. However, when water-deficit stress occurs the production of ROS is significantly accelerated via the Mehler reaction (Sharma, et al., 2012). In general, when DS plants perceive a lack of water their first response is to mechanically close their stomata to prevent any additional transpiration. The mechanisms involved in stomatal conductance are well studied and understood with the closing, and opening, being primarily driven by plant hormones and the movement of potassium ion between stomata and guard cells (Bharath, et al., 2021). Owing to the absence of water and closure of stomata, it is common for there to be a general lack of turgor pressure within the mesophyll and vascular tissue and leaf drooping occurs.

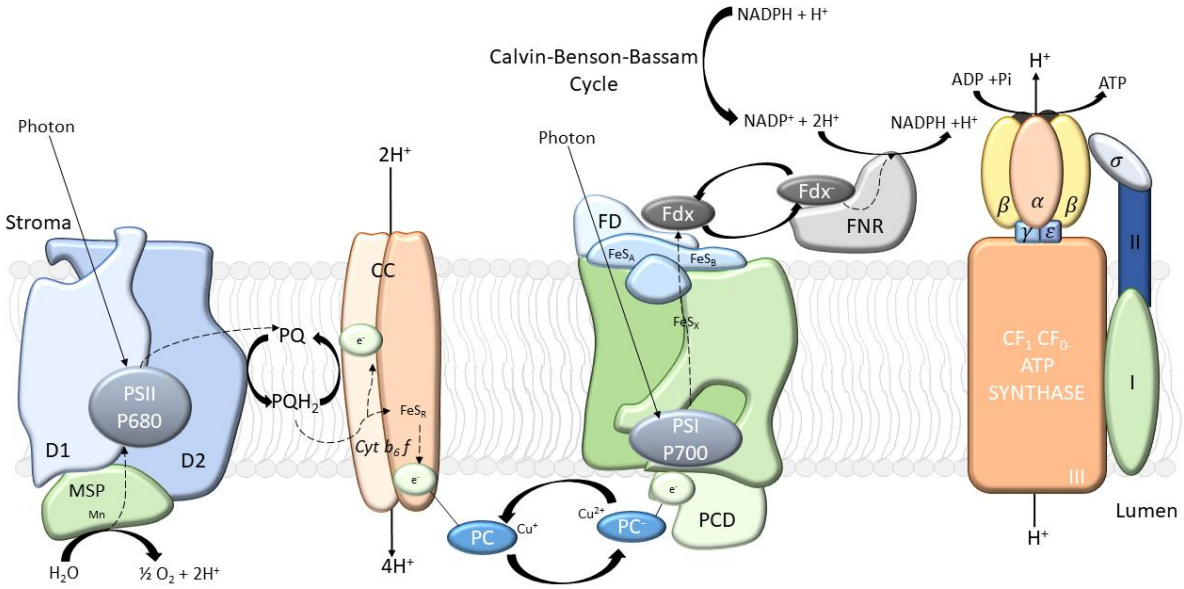


Figure 1-2: Photosynthetic electron transport. Schematic representation of the electron transport chain of photosynthesis. Dotted lines indicate the flow of electrons. MSP= manganese stabilising protein; D1 and D2= PSII reaction proteins; PQ= oxidised plastoquinone; PQH_2= reduced plastoquinone; CC= cytochrome complex; $FeS_{R,X,A,B}$= Fe-containing Rieske proteins; PC= plastocyanin; PCD= plastocyanin docking; FD= ferredoxin docking; Fdx= ferredoxin; FNR= ferredoxin-NADP+ oxidoreductase. Figure modified from (Campell & Farrell, 2012; Graham, et al., 2014)

For monocotyledonous plants, it is also common to observe leaf blades rolling in on themselves. This decreases the surface area exposed to irradiance. The closing of the stomata reduces the amount of CO_2 able to enter mesophyll tissue, though it should be noted that rapid closure of stomata does not rapidly lead to an absence of CO_2 within mesophyll tissue. As $NADP^+$ is used during the reduction of CO_2 during the Calvin-Benson-Bassam cycle (Figure 1-2), an absence of CO_2 leads to an accumulation of NADPH resulting in fewer $NADP^+$ being present to accept electrons via photosynthetic electron transport. As photosynthesis occurs in an aerobic environment, oxygen becomes a great acceptor of the excitation energy generated by P680*. Molecular Oxygen can be transformed to its singlet state ($^1O_2$) via energy absorption which initiates a chain reaction when it is given electrons to become a superoxide anion (O_2^-) which can undergo spontaneous dismutation to form hydrogen peroxide (H_2O_2) (Sharma, et al., 2012). These ROS can have severe effects on macromolecules such as lipid through peroxidation which can lead to membrane damage, protein inactivation through modification of peptides, and nucleic acid damage such as DNA strand breaking which can lead to genotoxicity and genome instability. It is worthwhile to note that although ROS has for the most part been considered a damaging agent in cells; more and more evidence is pointing to how valuable ROS is in signalling and communication (Tripathy & Oelmüller, 2012; El-

Maarouf-Bouteau & Bailly, 2008; Choudhury, et al., 2013; Foyer, 2018). A common method employed by many plants to diminish the harmful effects of ROS production is to either increase the expression of enzymes involved in antioxidative reactions such as superoxide dismutase, catalase, guaiacol peroxidase, and enzymes of the ascorbate-glutathione cycle; or to increase the production of non-enzymatic compounds such as ascorbate, glutathione, tocopherols, carotenoids, and phenolics such as flavonoids, tannins, hydroxycinnamate esters, and lignin (Sharma, et al., 2012) or both. Within the plant cell itself, various processes occur to scavenge ROI (reactive oxygen intermediates) which are typically formed through the excitation of O_2 to form compound such as the singlet oxygen, the superoxide radical, hydrogen peroxide and the hydroxyl radical (the latter three are formed through the addition of one, two, and three electrons each to O_2) (Mittler, 2002). These include pathways such as the use of catalase within the peroxisome (Figure 1-3 A), the ascorbate-glutathione cycle (Figure 1-3 C) in various organelles such as the mitochondria, chloroplast, the cytosol, the apoplast, and peroxisomes (Mittler, 2002), the water-water cycle (Figure 1-3 D) found in the chloroplast, and the glutathione reductase pathway (Figure 1-3 B) (Mittler, 2002).

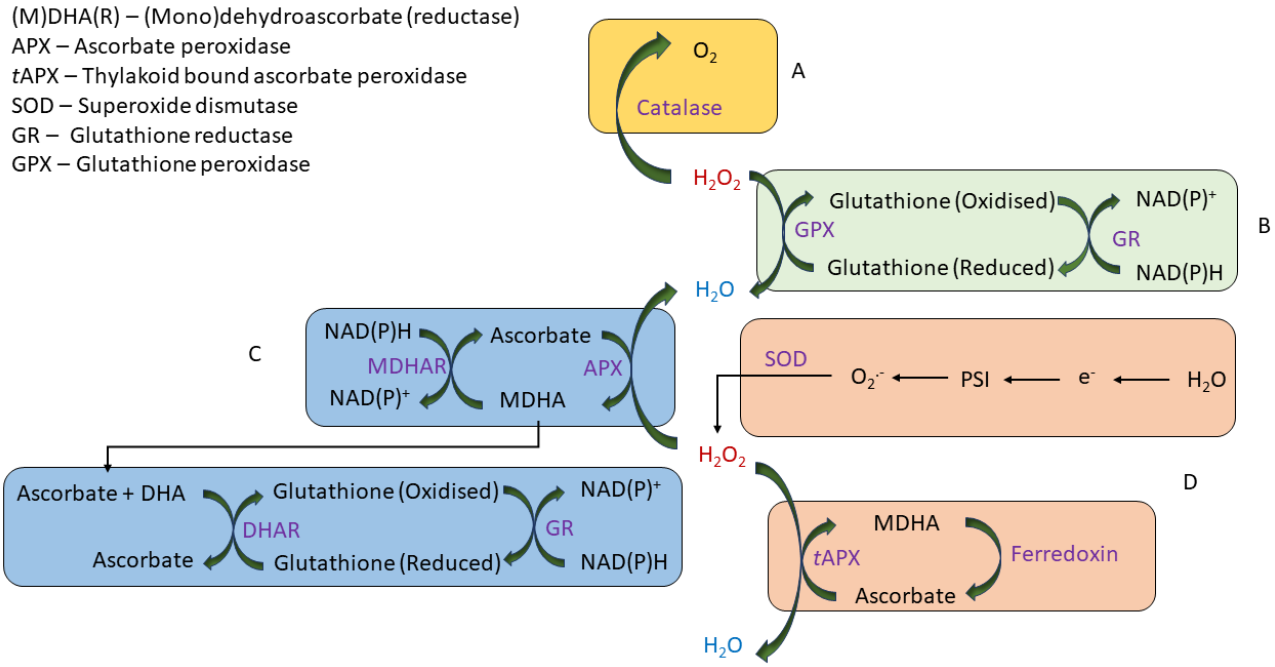


Figure 1-3: ROI scavenging in plants. Schematic representation of various pathways involved in scavenging of ROIs within plant cells. A: peroxisome catalase; B: glutathione reductase pathway; C: ascorbate-glutathione pathway; D: water-water cycle. Adapted from Mittler, (2002).

1.2.2 Drought response

Whether it be the seasonal perception of light intensity that determines dormancy in deciduous plants or the production of volatile compounds to attract pollinators and protectors, plants sense and respond to their environment. To this end and being restricted by movement, plants have to contend with fluctuations in water availability in the age old adage *adapt or die* with orthodox seeds being the ultimate survivors in dry conditions. DS plants have developed methods to tolerate a small loss in water. For instance, some plants such as the ephemeral bryophytes are often found in harsh environments such as *Vrolijkheidia circumscissa* (Pottiaceae) found in the succulent Karoo of South Africa (Hedderson & Zander, 2008) have short lifecycles so as not to experience periods of water limitation. Others would try to avoid water loss by controlling stomatal conductance and by virtue transpiration rates as well as maintain water uptake by having an expanded root system (Farooq, et al., 2009; Abobatta, 2018), while others would reduce the number of leaves to reduce the amount of water consumed. From a physiological perspective, DS plants can accumulate osmoprotectans as a means to modulate the osmotic state of cells under an array of stressed conditions (Suprasanna, et al., 2016). Hormones such as abscisic acid directly regulate stress response by not only closing stomata and thereby reduce transpiration (Wani, et al., 2016) but also in enhancing root proliferation (Giuliani, et al., 2005). Other hormones involved in drought response are the cytokinins whose activity is antagonistic to ABA (Pospíšilová, 2003) and is decreased during water-deficit stress. Jasmonates (JAs), though *sensu stricto* often more involved in pathogenic defence responses, activate plant responses to abiotic stresses such as drought (Seo, et al., 2011). Salicylic acid, yet another pathogenesis-associated gene regulator, has been shown to have a much greater concentration under drought stress (Bandurska & Stroinski, 2005). Naturally, these hormones seldom function in solitude, and it must therefore be stressed that although one hormone has one effect, the crosstalk and the various signal transductions can result in synergistic or antagonistic responses to drought (Wani, et al., 2016). The large problem, however, is that DS plants can only lose a fraction of their cellular water before irreversible damage and subsequent cellular death is encountered.

While these responses are to some extent mirrored in resurrection plants, for example the rolling of leaves seen with *E. nindensis*, resurrection plants employ an array of unique mechanisms to ensure that they are able to go beyond the drought response and enter the desiccation response. As indicated above, resurrection plants are broadly grouped as either retaining thylakoids or dismantling them. The dismantling of thylakoids, seen in *E. nindensis*, during dehydration ensures that one of the largest producers of ROS during dehydration is completely shut down (Van der Willigen, et al., 2001). The homoiochlorophyllous plants

accumulate large quantities of alternative pigments such as anthocyanins to quench excess light energy but maintain some degree of photosynthetic capacity (Farrant, 2000). Desiccation tolerant plants have also been characterised by the large accumulation of disordered proteins known as late embryogenesis abundant (LEA) proteins, commonly found in seeds during embryo development (Farrant, et al., 2012). These plants also accumulate large amounts of heat stable proteins such as HSP70. The presumed purpose for the accumulation of these thermotolerant proteins is to act as protectants. Other than accumulation of protective proteins, resurrection plants also increase the activity of a suite of antioxidant enzymes to quench ROS. A very striking feature of these plants is the accumulation of certain carbohydrates (recently reviewed in Dace et al., (2023). Of note are the plants that accumulate raffinose and related carbohydrates and those that accumulate trehalose. The proposed reason for the accumulation of these carbohydrates is to allow the cells to enter what has been termed as the 'glassy' state in an effort to ensure mechanical stability in the absence of water but also to act as a buffer to prevent rupture of cells upon rehydration. The accumulation of other carboxylic acids such as citric acid and amino acids such as proline have been proposed to assist in the establishment of the glassy state through the transient assembly of natural deep eutectic solvents (NaDES) (Du Toit, et al., 2021). What is more striking is that resurrection plants seem to have a similar response to dehydration as seen in orthodox seeds, presumably through repurposing of pathways involved in seed maturation. This similarity between vegetative desiccation tolerance and seed maturation has been the most accepted hypothesis of what drives desiccation tolerance (Oliver, et al., 2020). These processes are seemingly missing in vegetative DS plants, akin to what is seen in recalcitrant seeds.

Orphan crops for food security

As noted above, ensuring a food secure future in many developing and under-developed countries is vital in ensuring their socioeconomic prosperity. The Green Revolution has brought with it tremendous advancements in terms of global and local agricultural production, however the over-dependence on what I deem 'super-crops' (rice, soya, wheat, and maize) has resulted in locally adapted crops being replaced by apparent high-yielding varieties and shifting the research interest of plant breeders, agronomist, agriculturists, and plant geneticists towards these four. Whilst there is strong merit in acknowledging that genetic manipulation of economically important crops such as maize where varieties that are to some degree more tolerant towards water-deficit stress have been produced, orphan crops, so called for their regional importance, low international trade, and lack of research interest, provide a yet to be fully explored option to food security.

1.3.1 *Eragrostis tef* (Zucc.) Trotter

One such orphan crop is *Eragrostis tef*, more commonly known as teff. *E. tef* is an annual tufted panicle grass with high grazing value (Van Oudtshoorn, 2012). In Ethiopia, *E. tef* comprises 28% of the total acreage and accounts for 19% of the total annual grain production (Assefa, et al., 2011; Van Oudtshoorn, 2012). As *E. tef* is an annual C4 grass, it is intrinsically adapted to warmer climates being able to withstand prolonged warm periods and has a lower moisture requirement than C3 grasses. Nutritionally, *E. tef* contains many important elements such as K, S, P, Ca, Mg, Fe, Mn, Zn, and Cu, many of which are used as cofactors for an array of enzymes (da Silva Georsch, et al., 2019). Additionally, *E. tef*, is a source of almost all of the essential amino acids required by humans and is also rich in unsaturated fatty acids such as linoleic and oleic acids (da Silva Georsch, et al., 2019). Importantly, it is gluten free.

In terms of its tolerance to drought, it was observed that *E. tef* is able to withstand water loss to 50% relative water content (RWC) (Kamies, et al., 2017) before any adverse effects were noted. When compared to its resurrection relative *E. nindensis*, two varieties denoted as white and brown of *E. tef* were found to be mildly tolerant to drought with the latter being able to tolerate approximately 57% water loss compared to the white variety which lost viability at as little as 39% RWC (Ginbot & Farrant, 2011). Their results suggest that *E. tef* is not necessarily drought tolerant but is rather desiccation sensitive being able to withstand some degree of water loss.

1.3.2 *Eragrostis nindensis* as a model for desiccation and senescence

The Eragrostis genus owes its etymology from *eros* the Greek god of love and *Agrostis* which is Greek for grass. Members of the Eragrostis genus have spectacular inflorescence with noteworthy spikelets such as *E. superba* whose common name, the saw-tooth love grass, refers to its spikelet. *E. nindensis* occupies much of the mid- to northern areas of South Africa and is widespread from South Africa to as far as the Democratic Republic of Congo (Van Oudtshoorn, 2012). According to Van Oudtshoorn, (2012), it has average grazing quality commonly utilised by sheep and is classified as a subclimax increaser III grass. *E. nindensis* is a close relative of the Ethiopian orphan crop *E. tef* which has been described as a drought tolerant crop but lacks the capacity to become fully desiccated (Kamies, et al., 2017), although there is a fair degree of genomic overlap between the two having a 2:4 synteny of *E. nindensis* to *E. tef* (Pardo, et al., 2019). In the realm of vegetative desiccation tolerance, *E. nindensis* is unique in that it produces both desiccation-tolerant leaves (henceforth referred to as non-senescent) and desiccation-sensitive leaves (henceforth senescent). Upon dehydration and desiccation both tissue types resemble one another but it is only the non-senescent tissue

(NST) which resumes metabolic activity upon rehydration (Van der Willigen, et al., 2001) whereas the senescent tissue (ST) simply does not and commits to full senescence. Interestingly, the ST was once desiccation tolerant, but akin to desiccation and drought sensitive plants, loses this ability during development. In fact, we had noted that *E. nindensis* seedlings, approximately one month old, became fully desiccated and all of the leaves fully rehydrated and resumed growth which demonstrated that the 'switch' from tolerant to sensitive must be intertwined with plant and leaf development. Further investigation into this phenomenon is currently being undertaken. This unique phenotype of *E. nindensis*, makes it a useful model to study the drivers of senescence in a single species, particularly in the context of understanding the drivers that supress senescence in the NST.

Research aims and objectives:

The overarching aim of this thesis is to explore the proteomic signatures of desiccation tolerance in *E. nindensis* and investigate target genes for improved drought tolerance in *E. tef.*

1.4.1 Aim 1: Investigating the proteome for *E. nindensis.*

For this aim, senescent and non-senescent leaf tissue from the same dehydration and rehydration experiment done for a previously completed transcriptome (Madden, 2019) will be used to isolate protein for liquid chromatography mass spectrometry to determine the protein changes during dehydration in order to evaluate whether there are proteomic signatures that underpin desiccation tolerance and associated senescence. This will then be coupled with the transcriptome to investigate the fate of the transcripts and to generate a dual-omics model. The end goal for this work would be to potentially identify candidate genes for genetic engineering.

1.4.2 Aim 2: Functional characterisation of HSP70 from *E. nindensis*.

Heat shock proteins play an important role in proteostasis, and it has been hypothesised that protective proteins from resurrection plants may outperform their counterparts from drought-sensitive relatives. The transcriptome (Madden, 2019) indicated that one of the most up-regulated transcripts in *E. nindensis* was HSP70. A putative EnHSP70 was selected and the effect of overexpression of *EnHSP70* in the model organism *Arabidopsis thaliana* will be investigated. The propensity of EnHSP70 to undergo liquid-liquid phase separation will be investigated. Lastly, the thermo-tolerance and protection of EnHSP70 will also be investigated.

1.4.3 Aim 3: Developing a method for *E. tef* transformation and regeneration.

The final aim of this thesis is to develop a method for somatic embryogenesis using explants from *E. tef*. Successfully growing somatic embryos will be used as target for stable genetic transformation using agrobacterium-mediated transformation methods with the aim to regenerate these into transgenic plants. The work developed herein will be used to establish the foundation for future transformation using candidate genes identified from the transcriptome and proteome.

Chapter 2: Proteomic investigation of *Eragrostis nindensis* leaf tissue*:* unravelling proteomic signatures of desiccation tolerance.

Introduction

Omics refers to the study and characterisation of pools of biomolecules. It is all encompassing for that particular biomolecule (example, DNA for genomics, or RNA for transcriptomics) and allows scientists to predict the function, structure, and activity of the organism, tissue, or cell. In the context of studying water deficit stress, comparative omics of the transcriptome, proteome, and metabolome, provide vital information on key regulatory pathways, proteins, and genes involved in response to that stress. It is for this reason that having complementary omics profiles of *Eragrostis nindensis* could lay the foundation for possible genetic engineering. This Chapter will be aimed at analysing the global proteome of leaves of *E. nindensis* under water deficit stress and recovery therefrom, to gain insight into the protein changes in senescent and non-senescent tissue. The transcriptome for this time course and tissue type has already been completed (Madden, 2019), though at the time of writing this, it has not yet been published.

2.1.1 Senescence

Senescence is a universal developmental process associated with age or stress and is most commonly noticed upon seasonal shifts towards colder months when deciduous plants breakdown leaf contents as they prepare for dormancy. In a classical view, it is a well-orchestrated set of processes which eventually lead to cell death (Guiboileau, et al., 2010) and encompasses processes such as thylakoid degradation, vacuole collapse, chromatin condensation and eventual cell membrane degradation (Lim, et al., 2007). In this view, senescence is a way to redirect resources from older or senescing tissues to younger developing tissues in a typical source to sink manner (Lim, et al., 2007; Watanabe, et al., 2013). Plants have to contend with an array of environmental cues and signals and environmental stress can further promote the premature onset of senescence, which implies not only that it is a genetically-programmed age-dependent developmental fate but also governed by external stimuli (Lim, et al., 2007). The developmental fate of onset of senescence is a continuous process and is primarily controlled through cellular signals that initiate, maintain, and terminate it (Bresson, et al., 2018). Senescence is known to occur in a number of resurrection plants such as *Xerophyta schlechteri* (Radermacher, et al., 2019) *Craterostigma plantagineum* (Christ, et al., 2014) and *Sporobolus stapfianus* (Martinelli, et al.,

2007) but the underlining mechanisms that govern suppression of senescence in 'resurrected' tissue versus true senescence in non-resurrected tissues remains largely unknown. One might be tempted to question the relevance of further exploring senescence since it is both an age-dependent process as well as an environmentally-induced process. The answer is quite simple; if one can delay or supress senescence in economically important crops, they could remain greener and viable for longer under drought conditions especially since crop failure due to premature senescence is increasingly becoming more prominent (Rivero, et al., 2007). With this in mind a comprehensive physiological, ultrastructural, transcriptomic, and lipidomic study was undertaken to unravel signatures of senescence in *E. nindensis* (Madden, 2019) which will briefly be considered below.

2.1.2 Physiology and ultrastructural differentiates NST and ST from one another.

E. nindensis produces both senescent and non-senescent leaves as it matures and is classified as a poikilochlorophyllous resurrection plant owing to the complete breakdown of chlorophyll and thylakoids during dehydration. Morphologically, both tissues adopt a darkened brown colour at 40% RWC which progressively takes on a lighter syrup brown appearance. At 12hrs post-rehydration both tissues appear to be completely 'senescent' but at 72hrs post-rehydration the NST has completely unfolded and is green whereas the ST remains unchanged (Madden, 2019). During dehydration, the NST folds in on itself such that the abaxial surface of the leaf is exposed where reflective silica bodies were noted through scanning electron microscopy (Madden, 2019). Maximum photosystem II measurements indicated that PSII efficiency is maintained during early stages of dehydration circa 60% RWC in the NST but declined in the ST during this range. Both tissues had deactivated their PSII by 25% RWC. The NST reached pre-dehydration PSII efficiency approximately three days post-rehydration, whereas the ST failed to recover (Madden, 2019). This is further supported by the rapid decline in total chlorophyll which at approximately 60% RWC was approximately 5 and 2 $mg.mL^{-1}$ per mg dry weight compared to 7 and 15 $mg.mL^{-1}$ per mg dry weight in the NST and ST respectively at full turgor (Madden, 2019). Only the NST resumed accumulation of chlorophyll at approximately 30-40% RWC upon rehydration. Concomitant with the decrease in chlorophyll content, an increase in anthocyanin was noted in both tissues peaking at approximately 1.30 $mg.mL^{-1}$ per dry weight in the ST at the airdry state.

Being a C4 grass, *E. nindensis* displays Kranz anatomy with starch filled chloroplasts in the bundle sheaths at full turgor. At 60% RWC, a depletion of starch was noted with an increase in vacuolation in the NST whereas chloroplasts appeared to be packed with starch in the ST (Figure 8-1). Thylakoid membranes in the NST were however still intact indicating that commitment to breakdown of thylakoids is either at the early stages or has not started whereas

in the ST chloroplasts appeared swollen and thylakoids showed early signs of dismantling (Figure 8-1). Osmophilic bodies started accumulating at this stage predominantly around the bundle sheath and cell wall which increased in size at the 40% RWC in the NST (Figure 8-1). Similar bodies were seen in the ST, but their overall shape and size was irregular compared to those in the NST. Madden, (2019) proposed that the accumulation of the enlarged lipid bodies which appeared structurally dissimilar to the smaller osmophilic bodies could be storage bodies for lipids from thylakoid membranes. These lipid bodies were noticeably absent in the ST (Figure 8-1). At the 25% RWC point, the structure of the cell did not resemble that of the full turgor state and instead had chloroplast devoid of thylakoids with vesicles packed at the cell wall. Madden, (2019) noted the appearance of starch once more at this stage (Figure 8-1). Further desiccation to <5% RWC indicated that the ER was still intact with what appeared to be polyribosomes in proximity to the ER and that cytoplasm was preserved. Starch containing chloroplasts from the bundle sheath had accumulated in proximity to the vascular bundle presumably to act as a source of carbohydrate upon rehydration (Madden, 2019). Upon rehydration, the cell returned to a pre-dehydrated state with chloroplasts containing assembled thylakoids and the lipid droplets noted earlier seemingly missing (Figure 8-1). The ST was characterised as not laying down the foundation to tolerate desiccation such as vacuolation and lipid droplet accumulation along the cell wall. Madden, (2019) proposed that the disappearance of these osmophilic bodies in the desiccated state of the ST could be indicative of sink transportation from the ST and concluded that the ST displayed what is likened to classical senescence. The studies when taken together demonstrated that the two tissues morphologically and at the cellular architectural level were behaving fundamentally differently during dehydration stress. The apparent accumulation of lipid bodies prompted the further investigation into the lipid composition of *E. nindensis* which will be briefly considered next.

2.1.3 The lipidome identified *E. nindensis* as a C18 plant.

In plants, *de novo* synthesis of fatty acids occurs in plastids (chloroplasts, leucoplasts, and chromoplasts) from carbon assimilated during CO_2 fixation and are then exported to the endoplasmic reticulum for further chain elongation or modification to generate a suite of highly diverse lipids. Plants can synthesise glycerolipids, an important constituent of the cellular membrane (Li, et al., 2015) using one of two pathways. In the prokaryotic pathway, phosphatidic acid mainly comprising a C_{16} fatty acid moiety at the *sn*-2 position of glycerol are synthesised through sequential activity of acyl-ACP glycerol 3-phosphate acyltransferase and acyl ACP: lysophosphatidic acid acyltransferase (Ohlrogge & Browse, 1995; Li, et al., 2015; Mongrand, et al., 1998). The lipid produced here can be further used for the synthesis of

phosphatidylglycerols or further chemically altered through hydrolysis to produce diacylglycerols (DAG) which can then be further processed into monogalactosyldiacylglycerols (MGDG), digalactosyldiacylglycerol (DGDG) or sulfoguinovosyldiacylglycerol or through incorporation of cytidine triphosphate to produce cytidine diphosphate-diacylglycerol (CDP-DG) which can be further processed to produce phosphatidylglycerol (PG) (Ohlrogge & Browse, 1995; Mongrand, et al., 1998). In the eukaryotic pathway phosphatidic acid synthesis in the ER produces lipids containing C_{18} fatty acids at the *sn*-2 position of the glycerol where the DAG can be further processed to produce phosphatidylcholine or phosphatidylethanolamine and the CDP-DG can be further processed to produce PG or phosphatidylinositol (Ohlrogge & Browse, 1995; Mongrand, et al., 1998). Plants which predominantly rely on the prokaryotic pathway are denoted as C16:3 whereas those relying primarily on the eukaryotic pathway are denoted as C18:3, however there is plasticity between the two pathways and their combined contribution to the glycerolipid production in the leaves (Li, et al., 2015). Some plants exclusively use the prokaryotic pathway for PG synthesis, relying on the eukaryotic pathway for the synthesis of MGDG and DGDG and thus contain no C16:3 (Mongrand, et al., 1998).

Upon abiotic stress, lipids are readily mobilised as seen for example in the desiccation tolerant *C. plantagineum* (Frank, et al., 2000). The identification of a core set of desiccation tolerant genes further corroborates the role that lipids play in vegetative desiccation tolerance (Costa, et al., 2016). The lipidomic study identified *E. nindensis*' glycerolipid composition as C18:3 comprising lipids belonging to the sterol lipid, sphingolipid, glycerolipid, and glycerophospholipid categories (Madden, 2019). Differential abundances of lipid classes were also observed between the two tissue types with an increase in lipid content during dehydration in the NST tissue. Further analyses indicated that the composition of triacylglycerols (TAGs) and the manner in which they are packaged are more important for *E. nindensis*' desiccation tolerance rather than the accumulation of TAGs in other resurrection plants such as in *X. humilis* (Tshabuse, et al., 2018). Accumulation of lipid droplets (LDs) identified through transmission electron microscopy also varied between the two tissue types with the NST having smaller droplets along the plasmalemma whereas the droplets in ST were irregular and large (Madden, 2019). Their abundance also differed with the droplets disappearing upon rehydration of the NST tissue but still present in ST upon rehydration. The LD's distribution implies a much greater control over their distribution and size in NST than in ST. One of the possible conclusions drawn from the LD's distribution is the presence and accumulation of oleosin. The interplay between the TAG composition and oleosin is in the manner in which they are assembled. Upon synthesis in the ER, TAG is sequestered into the phospholipid bilayer before budding off (Huang, 2018). The LDs are composed of TAGs and

metabolites and are encased in a mono phospholipidic layer embedded with oleosin (Huang, 2018). Oleosin's structure allows it to interact with the TAG core while dispelling other oleosin-encased droplets through electrostatic repulsion. This ordered association of oleosin, and LDs give rise to a much tighter control over their assembly and distribution in NST than in ST. The large accumulation of irregular LDs in ST clearly demonstrated the fate of improper translational control in the assembly of the droplets. By accumulating along the plasmalemma, the LDs provide structural support during dehydration in NST tissue which is vital to ensure cell viability upon rehydration.

2.1.4 The transcriptome indicated similar processes between NST and ST leaf tissue

By considering the sum of differentially expressed genes in Figure 2-1 below, there is a lot to unpack. A noticeable down regulation of genes involved in photosynthesis (typical for poikilochlorophyllous plants such as *E. nindensis*), cell wall processes, nutrient uptake processes, and solute transport is suggestive of tissue moving towards a shut-down state. Though not clearly evident in Figure 2-1 below is the reshuffling of important metabolic processes involved in carbohydrates, amino acids, and coenzymes. Both tissue types have upregulation of genes involved in external stimuli perception which is indicative of plants being subjected to stress.

By looking at the number of differentially expressed genes between the NST and ST, the NST had 2690 transcripts compared to only 542 in the ST relative to a fully hydrated control, suggesting that at the transcript level desiccation tolerance is programmed in the NST further implying that the NST is under much greater regulatory control than that of the ST. Interestingly, transcripts were detected at RWC below 25% in the NST which is suggestive that either active transcription is occurring or transcripts are being saved for rehydration. The relative abundance however was considerably less hinting that the metabolic activity or rate was at or close to a quiescent state. This is also reflected in the accumulation of seed dormancy related transcripts.

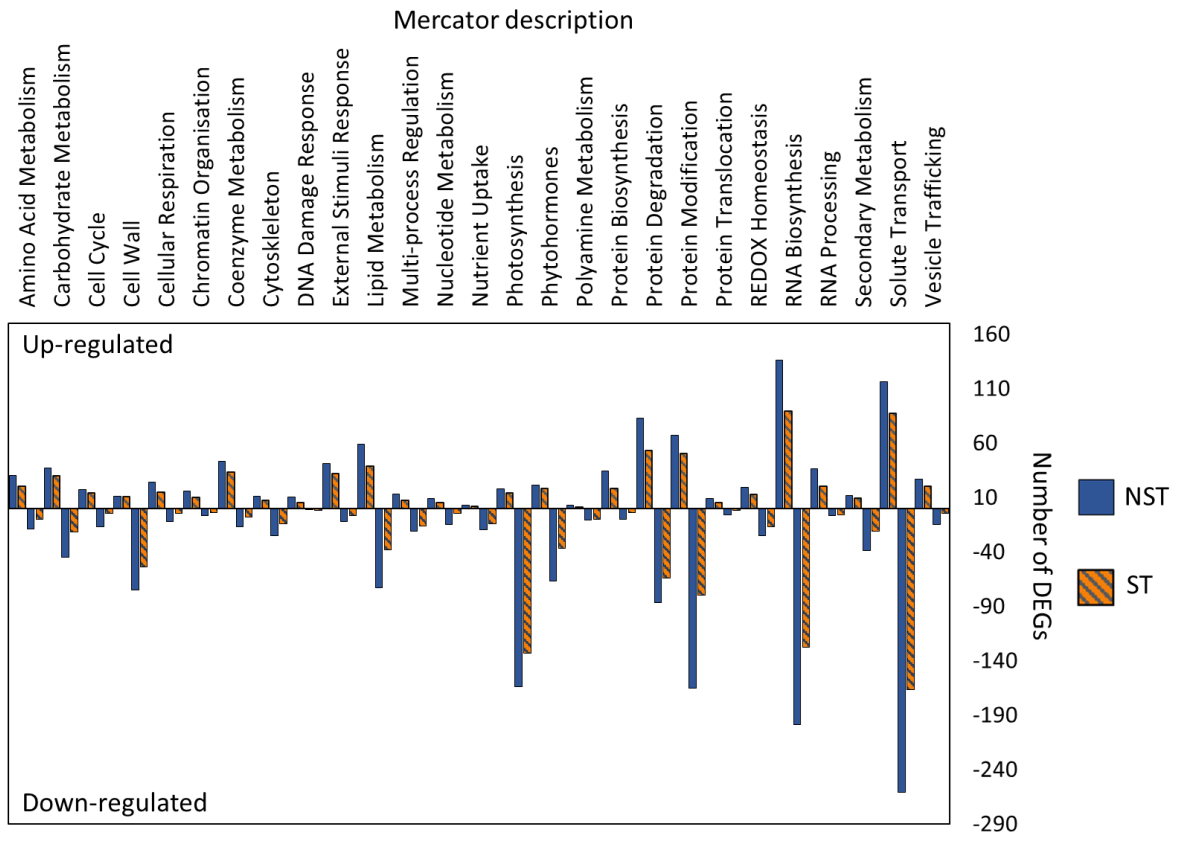


Figure 2-1: *Eragrostis nindensis* transcriptome summary: total number of up- and down-regulated differentially expressed genes in E. nindensis where NST refers to the non-senescent and ST the senescent tissue collected following a dehydration and rehydration treatment. Data from Madden (2019).

A consensus hypothesis for desiccation tolerance in the vegetative state of resurrection plants lies in the similarities observed in the desiccation response and seed development of orthodox seeds (Costa, et al., 2017a; VanBuren, et al., 2017). When considering seeds, two distinct classes can be drawn, those belonging to the orthodox class and those belonging to the recalcitrant class (Roberts, 1973). The former class are seeds which undergo a programmed dehydration during maturation and are in essence desiccation tolerant once mature. Upon imbibition, the seed 'resurrects' and resumes growth. The latter class are seeds which do not undergo programmed dehydration and require immediate planting upon maturation in order to resume growth. The same analogy can be drawn between vegetative desiccation tolerance as seen in resurrection plants and desiccation sensitive plants. This analogy extends beyond a simple anecdotal observation as it was reflected in the transcriptome with the accumulation of late embryogenesis abundant (LEA) protein transcripts. LEA proteins become much more abundant during the later stages of embryo development in seeds and are proposed to be involved in processes that offer protection. The abundance of these seed-specific proteins has been confirmed in a number of other resurrection plants such as *Xerophyta humilis* (Collett, et al., 2004), *X. schlechteri* (Costa, et al., 2017b; Gabier, et al., 2021), *Myrothamnus flabellifolia*

(Ma, et al., 2015) and more recently *Ramonda serbica* (Vidović, et al., 2022) to name a few. In *R. serbica*, a resurrection plant found in areas such as Bulgaria, Albania, and Serbia, 21% and 14% of the upregulated LEAs were from the LEA4.3 and LEA1 subfamilies respectively (Vidović, et al., 2022). The accumulation of LEA is not a desiccation-tolerant-only trait as its transcripts were found in the ST of *E. nindensis* and it has been shown that *E. tef* also accumulates LEAs upon water-deficit stress but apparently lacks the ability to express these accurately (Pardo, et al., 2019). Recently and by means of genome data comparison, the authors demonstrated that although LEAs are present, in general dehydration stress response among desiccation-sensitive and -tolerant grasses, LEA5 and LEA6 subfamilies were noticeably induced in *E. nindensis* and *Oropetium thomaeum* but not in *E. tef* (Pardo, et al., 2020).

In addition to LEA, heat shock proteins (HSP) were also significantly increased in both NST and ST (Madden, 2019). Both LEA and HSP are heat-stable proteins with a varying degree of disorder. This disordered nature makes these proteins particular useful as they are able to bind to an array of substrates but ultimately have the fundamental role of offering protection to these substrates. In *X. schlechteri*, a noticeable accumulation of HSP90 was noted during the late response to drying (Gabier, et al., 2021). Beyond their role as protectants under drought and heat stress, it is postulated that given their native intrinsically disordered nature, they along with other biomolecules such as mRNA, metabolites, fatty acids, and mRNA processing proteins, can form plant stress granules which under stress conditions form through liquid-liquid-phase separation and disappear upon stress alleviation (Maruri-López, et al., 2021).

While there was a considerable amount of mirroring that occurred between the NST and ST, transcripts exclusive to the NST were predominantly abundant in those associated with RNA regulation and amino acid metabolism, RNA processing, protein folding, response to heat, response to glucose, and regulation of seed germination, with many of them showing much higher abundance at lower RWC. The accumulation of RNA processing proteins was also observed in *X. schlechteri* during the late response to drying (Gabier, et al., 2021), which the authors propose is required for recovery. In addition, transcripts related to the ribosomal 40S and 60S proteins were significantly overrepresented in the lower RWCs of the NST. Focusing on those related to RNA biosynthesis it is clear that at the 25% to 10% RWCs there are considerable differences between transcripts accumulating and diminishing.

RNA biosynthesis transcripts related to growth such as auxin response, B3 DNA binding (Sasnauskas, et al., 2018), morphogenesis (HOX) (Bürglin, 2001), shoot development (no apical meristem -NAM) (Souer, et al., 1996), and flowering time (CCT motif) (Li & Xu, 2017)

are noticeably diminished at these RWCs (Madden, 2019) which is not all that surprising considering that at 25-10% RWC the plant is experiencing severe drought and would thus not partition energy reserves to promote growth. When looking at the transcripts that accumulate a much different state becomes evident. The AP2 domain, or similarly the ethylene responsive factor (ERF), though varying in relative transcript abundance, was predominantly accumulating in NST at the ~10% RWC. As part of the ethylene insensitive pathway (EIN2), the presence of these domains is highly suggestive that the NST is still responsive to the drought stress. The B-Box Zn-finger, a subgroup of the Zn-finger of proteins is thought to be involved in processes associated with protein-protein interactions and could possibly occur at complexes such as those involved in transcription (Khanna, et al., 2009) suggesting that the presence of these transcription factors could be indicative of active transcription. These transcripts and similar proteins as seen in *X. schlechteri* may be associated with natural deep eutectic solvents (NaDES) as proposed by Du Toit, et al., (2021) which could facilitate active transcription and translation. This possible active transcription is further supported by the presence of transcripts associated with the RNA polymerase subunits 1,2,6, M/15 Kd, and Rtr1/RPAP2. An even more interesting support towards active cellular activity is the presence of ribonuclease III and RNA recognition motif (RRM) associated with cleavage of dsRNA to produce mature rRNAs, tRNAs, and mRNAs to name a few (Zamore, 2001), and post-transcriptional gene expression respectively (Maris, et al., 2005). Other noteworthy mentions include the G-box binding protein MFMR, proposed to be involved in signal transduction pathways brought upon by environmental stress and involved in gene expression (de Vetten & Ferl, 1994; Sibéril, et al., 2001); the TCP family of transcription factors involved in a number of plant developmental processes such as leaf senescence, defence, and the regulation of pathways involved in bioactive compound synthesis to name a few (Li, 2015).

The response to stress is a complex pathway involving innumerable steps and intermediates to ensure that a plant responds appropriately and rapidly for failure to do so will most likely result in cell death. One can argue that during the drying down stage both desiccation sensitive and tolerant plants respond in a similar manner, for instance in the accumulation of seed related genes such a LEAs (Pardo, et al., 2019) or the activation of ROS scavenging and quenching proteins (Lei, et al., 2022; Li, et al., 2022b), however, those that are desiccation tolerant are able to tightly regulate their transcription and translation more effectively and are thus able to withstand the stress and enter this quiescent state. In the case of *E. nindensis*, the presence of active transcription at <25% RWC is strongly suggestive that it is still actively transcribing RNA involved in appropriate stress responses. When considering the transcripts involved in protein biosynthesis, some noticeable candidates emerge. Factors involved in translation initiation, elongation, and release are present at <25% RWC, even in ST.

Ribosomal protein transcripts L10e, 18, 18e, 36e, S6, 12, and 23, all accumulate at the 10% RWC suggesting that some degree of ribosomal assembling is occurring. Coupled with the required factors involved in protein biosynthesis, it is plausible to speculate that there is some protein turnover albeit minimal. In fact, one protein investigated during the transcriptome study, oleosin, showed expression at <10% RWC when probed on a western blot (Madden, 2019). Taken together, the transcriptome of *E. nindensis* clearly demonstrates that at RWCs where most desiccation-sensitive plants would perish, *E. nindensis* is still able to transcribe and translate (Madden, 2019). What remains to be further elucidated however, is the time course of these transcript and the fate of those transcript during the early stages of drought stress and those that are there at severe drought stress. There is also no conclusive evidence as to the degree of reprogramming upon recovery and the persistence of hypothetical protective proteins.

2.1.5 Importance of proteomics

There have been a large number of omics-based studies done on various resurrection plants reviewed in Lyall & Gechev, (2020) and Dinakar & Bartels, (2013) which have assisted in elucidating the mechanisms employed by resurrection plants during dehydration and desication stress. By considering some of the proteomic studies, early studies such as those on *X. viscosa* (Ingle, et al., 2007), *Boea hydrometrica* (Jiang, et al., 2006) and *Selaginella bryopteris* (Deeba, et al., 2009) and a more recent study on *Cupressus gigantea* (Lei, et al., 2022) all employed the use of 2-dimensional polyacrylamide gel electrophoresis (PAGE) to identify differentially expressed proteins commonly reported as differences in dot appearance and intensity on 2D gels. Though rudimentary by today's standards, these early studies helped shape our understanding of the importance of not just considering the transcriptional state of the stressed plant, but also its translational state. When these types of studies are compared to studies such as the one done on *X. schlechteri* (Gabier, et al., 2021) or *R. serbica* (Vidović, et al., 2022) where protein extracts were analysed, the latter approach yielded much more informative data than only those presented based on visual differences on a gel. Proteomic studies, either in tandem with other omics studies or alone, provide insightful information not only on the response networks but also hints to key proteins involved in priming the defence. For instance, by means of complementing the transcriptome with an iTRAQ-based proteomic study of a drought-resistant peanut cultivar (*Arachis hypogaea*), the authors demonstrated that 69 proteins had altered expression under drought conditions with many of them involved in stress response (Li, et al., 2022a). For example, the LEA2, ABA-responsive protein, peroxidase, and histone H1 all had upregulation at the transcription and translation level and similar patterns were observed for some of the downregulated proteins too. Though it should

be noted that it was reported that there were differences between the expression profile of the transcriptome and proteome for the same treatment type where transcripts might show upregulation but proteins for the same transcript had opposite expression , implying that post-transcriptional modification and regulation thereof are key regulators for protein turnover and expression (Vogel & Marcotte, 2012) and emphasising the caution that must be applied to transcriptomic studies since the up-regulation of a particular gene does not necessarily imply functional protein. This apparent delay in transcription was also noted for the resurrection plant *C. plantagineum* (Xu, et al., 2021) where the authors too used a multi-omics approach to investigate key drivers of desiccation tolerance. The appearance of the LEA2 in such abundance in *A. hypogaea* led the authors to expressing this protein in *A. thaliana* where they observed improved physiological parameters such as a significant reduction in the degree of lipid peroxidation under drought stress indicating that the LEA2 plays an important role in the drought-resistance phenotype of this particular cultivar (Li, et al., 2022a). Although there have been a large number of transcriptomic studies done on vegetative desiccation tolerant plants and has led to our core understanding of key mechanisms such as the accumulation of sucrose, raffinose, and trehalose, the overexpression of chaperones such as LEAs and HSPs, and the increase in antioxidative response by increased expression of antioxidant enzymes such as peroxidase, superoxide dismutase and catalase, a fundamental flaw in transcriptomic-based gene discovery and model implementation is that the presence of a transcript is not reflective of protein activity. Proteins are the functional units in the cell, and as such, it is surprising that not that many proteomic investigations have been done on resurrection plants. The author of the transcriptomic study was tasked with identifying whether there were RNA signatures which could differentiate the senescent and non-senescent leaves of *E. nindensis.* The overarching conclusion drawn from it was that both tissues appeared to behave in a similar manner with differential expression of related transcripts present in both. The author concluded that translational regulation is probably involved in ensuring that the non-senescent tissue reaches a quiescent state and is able to resume metabolic activity upon rehydration.

Research questions to be addressed:

The overall aim of this Chapter is to complete a proteomic study using the same leaf tissues as used for a previous transcriptomic study with the purpose of understanding the proteomic profile of *E. nindensis.* Key questions to be addressed here are:

1. What are the changes in protein abundance occurring during dehydration and upon rehydration?
2. Are there proteomic signatures which distinguish the NST from the ST?
3. To what extent can identified differentially abundant proteins (DAPs) complement the transcriptome?
    a. Are there temporal and spatial differences between gene expression and mRNA translation?
    b. Does the relative abundance of matching proteins correspond to the same relative expression of transcripts?

Methods

2.2.1 *E. nindensis* growth and sampling

All *E. nindensis* tissue used for the global proteome were previously sampled except for a few additional samples which had to be supplemented. The growth parameters and sampling that was done for the transcriptome will be briefly summarised below. Sterilized *E. nindensis* seeds were germinated on 4.4 $g.L^{-1}$ MS media solidified with 0.8% agar in a growth chamber at 24/15 °C day/night with a 16-hour photoperiod until seedlings were well developed. Seedlings were hardened off first by removing the tissue culture lids before being transferred to soil and transferred to a plant growth room where they were watered 2-3 times a week and fertilised once a week. Plants were then allowed to grow for approximately six months. In order to achieve uniform age and to initiate priming, these plants were first subjected to a dehydration time course upon which aerial tissue was removed and they were allowed to resume growth until such time biomass was sufficient for sampling. Tissue denoted as senescent (also referred to as desiccation-sensitive) and non-senescent (also referred to as desiccation-tolerant) (Figure 2-2 C) were sampled from randomised pots at different relative water contents (RWC). Multiple dry-down curves were generated in Madden (2019) and was used to determine the relationship between RWC and absolute water content (AWC). The AWC and RWC were then correlated with one another and RWC bins was assigned when there was congruence. *E. nindensis* rapidly lost AWC percentage within a short period of time and only the NST resumed reabsorption of water upon rehydration (Figure 2-2 A). For the remainder of this thesis, RWC will used instead of AWC During dehydration, total pigment is lost in both leaf tissue by 40% RWC. Distinction between the two tissues was primarily done by noting the position, age, and general appearance of the NST and ST during sampling. Given the nature of relative water estimation, not all RWCs were included in the final transcriptome for the two tissue types. Similarly, not all RWCs were included for the proteome. For rehydration time points, the transcriptome used 12hrs post-rehydration whereas the proteome opted to focus on 48hrs and 1 week post-rehydration. The proteome did not include airdry tissue from ST in leu of including an additional rehydration time point for the NST. Collected tissue was snap frozen in liquid nitrogen and stored thereafter at -80°C until further use.

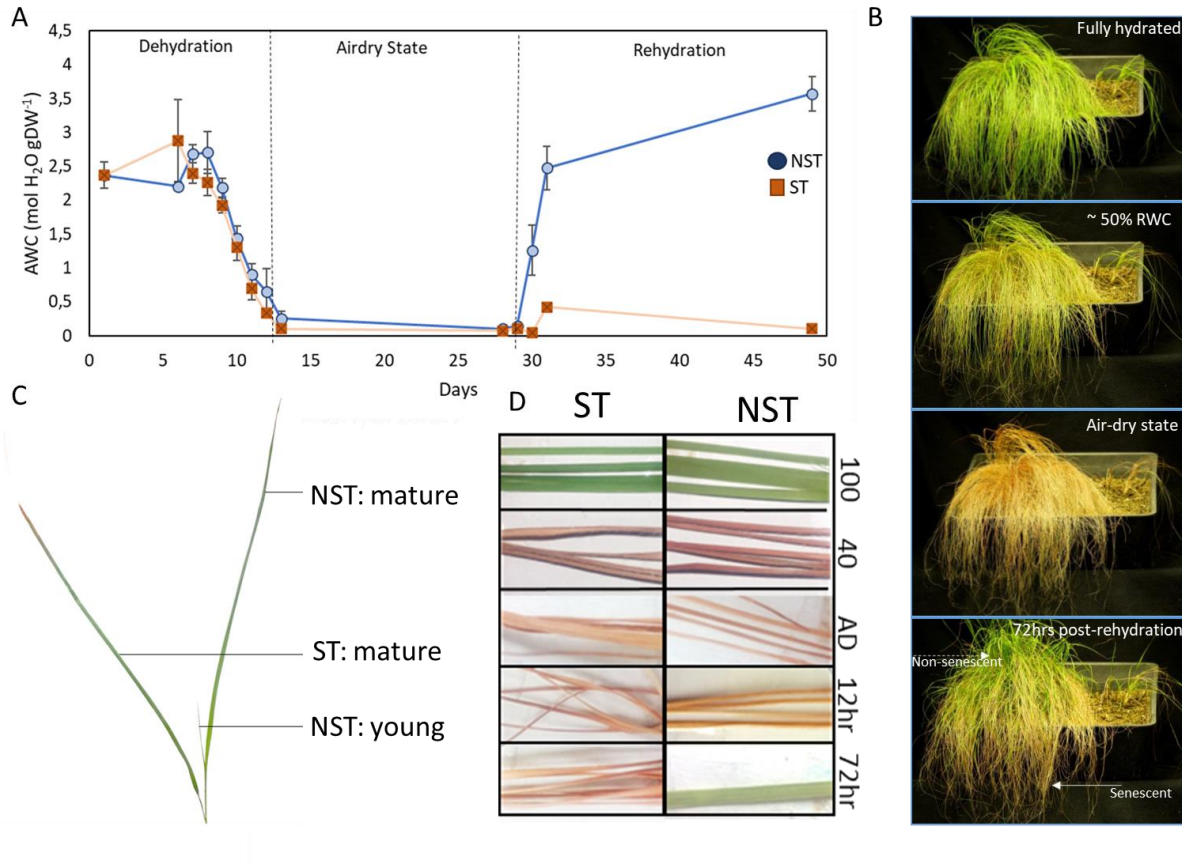


Figure 2-2: *E. nindensis* AWC curve and tissue distinction: (A) Absolute water content dry-down graph generated by Madden (2019) showing rapid decline in AWC content with only the NST resuming water re-absorption upon rehydration. Dry-down was used to bin tissues into RWC categories for sampling of transcriptome and proteome leaf tissue. (B) Shows time lapse images of aerial tissue during dehydration and rehydration. Time lapse footage can be viewed here: https://www.youtube.com/watch?v=gV6_CJXM0sw. (C) Close up image of the individual tillers of *E. nindensis* indicating the differentiation between the NST and ST used during sampling. (D) Close up images of individual leaf sections of NST and ST at different RWCs. The NST resumed re-greening by 72hrs post-rehydration. NST= non-senescent; ST= Senescent. Images kindly provided by C. Madden and also used in Madden (2019).

2.2.2 Optimisation of protein isolation

Various methods for protein isolation were evaluated on 50-100 mg of ground *Eragrostis curvula* fresh leaves before being optimised for *E. nindensis*. *E. curvula* was selected on the basis of it being much more fibrous than *E. nindensis* and provided a good proxy for method development for the drier *E. nindensis* tissue. Of the methods tested, a modified TRIZOL method was used for all further protein isolation as described by Gabier, et al., (2021). Briefly, ground tissue of approximately 30-60 mg fresh weight tissue was added to a tube to which 1 mL TRI-reagent (SIGMA) was added, vortexed, and then incubated at room temperature for 5 mins. To this, 200 µL of chloroform was added and again briefly vortexed and incubated

once more at room temperature for 5 mins. The samples were then centrifuged for 15 mins at 12 000 *x* g at 4°C. Upon completion, the upper phase which contained RNA was removed and processed separately. To the lower phenol phase (containing DNA and protein) 300 µL of ice-cold 100% absolute ethanol was added, briefly vortexed and centrifuged at 2500 *x* g for 15 mins at 4°C. The supernatant was removed and added to 1.5 mL ice-cold isopropanol, inverted, and then incubated for 10 mins at room temperature before being centrifuged at 10 250 *x* g for 15mins at 4°C. The resulting protein pellet was then washed with ice-cold 0.1 M ammonium acetate in 100% methanol and then again with 100% ice-cold acetone. The pellets were airdried in a fume hood for 30 minutes. The pellets were then resuspended in 100 µL resuspension buffer (50 mM triethylammonium bicarbonate and 2% SDS) and incubated at 95°C for 5 minutes to assist in the dissolving of the proteins, vortexing twice during the incubation. Upon completion, an additional 50 µL of the resuspension buffer was added and the contents pipetted. Owing to the presence of the SDS in the buffer, the samples were returned to the heating block for 5 mins before removing 100 µL of the supernatant to be used for analysis. Total protein was quantified using the BCA quantification method (Pierce) where 20 µL of 1:10 dilution of the sample was incubated with 200 µL of the BCA reagent in triplicate at 37°C for 30 minutes before measuring the absorbance at 562 nm. Total protein was estimated based on a standard curve using bovine serum albumin in half-strength resuspension buffer and corrected to account for dilution.

2.2.3 Mass spectrometry and analysis

2.2.3.1 On-bead HILIC digest

In preparation for the HILIC magnetic bead workflow, the HILIC beads (ReSyn Biosciences, HLC010) were aliquoted into a new tube and the shipping solution removed. The beads were then washed with 250 µL wash buffer (15 % (*v/v*) acetonitrile, 100 mM ammonium acetate (Sigma 14267) pH4.5) for one minute and repeated twice. The beads were then resuspended in loading buffer (30% acetonitrile, 200 mM ammonium acetate, pH4.5) to a concentration of 2.5 $mg.mL^{-1}$. A total of 20 µg from each protein sample was transferred to a protein LoBind plate (Merck, 0030504.100). Protein was reduced and alkylated by addition of 20 mM dithiothreitol (DTT, Sigma D9779) and 30 mM iodoacetamide (IAA, Sigma 16125) and incubated at 95°C for 10 minutes. HILIC magnetic beads were then added at an equal volume to that of the sample and a ratio of 5:1 total protein. The plate was then incubated at room temperature on a shaker at 900 RPM for 30 minutes to allow binding of the protein to the beads. Upon completion of binding, the beads were washed four times with 500 µL 95%

acetonitrile for one minute. For digestion, trypsin (Promega PRV5111) made up in 50 mM TEAB was added at a ratio of 1:20 total protein and LysC (Pierce 90307) was added at a ratio of 1:250 total protein. The plate was then incubated at 45°C on the shaker for two hours. After digestion, the supernatant containing peptides was removed and dried down. Samples were then resuspended in liquid chromatography loading buffer comprising 0.1% (*v/v*) formic acid and 2.5% (*v/v*) acetonitrile.

2.2.3.2 Liquid Chromatography mass spectrometry

Liquid-chromatography mass-spectrometry was conducted with a Q-Exactive quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, USA) coupled with a Dionex Ultimate 3000 nano-ultra performance liquid chromatography system. Data was acquired using Xcalibur v4.1.31.9, Chromeleon v6.8 (SR13), Orbitrap MS v2.9 (build 2926), and Thermo Foundations 3.1 (SP4). Peptides were dissolved in 0.1% (*v/v*) formic acid (Sigma 56302) and 2% (*v/v*) acetonitrile and loaded on a C18 trap column (PrepMap100, 902790500, 300 µm x 5 mm x 5 µm). Approximately 400 ng of peptide was injected per sample. Samples were trapped onto the column and washed for 3 minutes before the valve was switched and peptides eluted onto the analytical column. Chromatographic separation was performed with a C18 column (0.75 µm x 15 cm x 1.7 µm). The solvent system comprised of solvent A (liquid chromatography grade water (Burdick and Jackson BJLC365) and 0.1% formic acid) and solvent B comprising acetonitrile and 0.1% formic acid. A multi-step gradient (summarised in Table 2-1) for peptide separation was generated at 300 $nL.min^{-1}$ as follows: time change: 1.5 mins, gradient change: 2-5% solvent B, time change: 50 min, gradient change: 5-18% solvent B, time change: 1 min, gradient change: 18-30% solvent B, time change: 4 mins, gradient change 30-80% solvent B. The gradient was then held at 80% solvent B for 5 mins before returning it to 2% solvent B for 5 mins.

Table 2-1: Liquid chromatography gradient for the LFQ liquid-chromatography mass spectrometry/mass spectrometry

Time (min)	**Flow rate ($\mu L.min^{-1}$)**	**Solvent A (%)**	**Solvent B (%)**
00.00	0.450	98	2
00.00	0.450	98	2
03.00	0.450	98	2
04.40	0.450	95	5
04.50	0.300	95	5
55.00	0.300	82	18
56.00	0.300	70	30
56.00	0.300	70	30
60.00	0.300	20	80
60.10	0.450	20	80
64.90	0.450	20	80
65.00	0.450	98	2
70.00	0.450	98	2

All data acquisition was obtained using Proxeon stainless steel emitters (Thermo Fisher TFES523). The mass spectrometer was operated in a positive ion mode with a capillary temperature of 320°C. The applied electrospray voltage was 1.95 kV. Details of data acquisition are summarised in Table 2-2

Table 2-2: Mass spectrometry data acquisition parameters

Full Scan	
Resolution	70 000 (@ m/z 200)
AGC target value	3e6
Scan range	350-2000 m/z
Maximal injection time (ms)	100
Data-dependent MS/MS	
Inclusion	Off
Resolution	17 500 (@ m/z 200)
AGC target value	1e5
Maximal injection time (ms)	50
Loop count	10
Isolation window width (m/z)	2
NCE (%)	27
Data-dependent settings	
Underfill ration (%)	1
Charge exclusion	Unassigned, 1,7,8, >8
Peptide match	Preferred
Exclusion isotopes	On
Dynamic exclusion (s)	60

2.2.3.3 Data analysis

Relative quantification was conducted using Progenesis QI for proteomic v2.0.5556.29015 (Nonlinear Dynamics, UK). Data processing included peak picking, run alignment and normalisation (singly charged spectra were removed from the processing pipeline). Protein quantification was run using the "relative quantification using non-conflicting peptides" method. Database interrogation was performed with Byonic Software v3.8.13 (Protein Metrics, USA) using *E. curvula* reference proteome sourced from UniProt dated 03.07.2022. A pilot study using fully hydrated leaf tissue was done to determine which method of protein isolation yielded most optimal results as well as which database returned the highest annotation for peptide spectra. This pilot study used *Zea mays*, *Arabidopsis thaliana*, and the *E. curvula* UniProt reference proteomes with the latter returning a greater number of annotated peptide spectra. Details of search parameters are summarised in Table 2-3.

Accession numbers associated with each comparative pair and RWC were uploaded onto DAVID (https://david.ncifcrf.gov/home.jsp) using the full differentially expressed protein accession numbers from all RWCs as reference for gene ontology enrichment analysis. GO terms selected were biological process, molecular function, and cellular compartment. GO terms with an FDR of <0.05 were selected to filter the range of terms. The FDR selected GO terms were then uploaded onto REVIGO (http://revigo.irb.hr/) without the accompanying p-value generated from DAVID to summarise the GO terms by removing redundancy. All FDR of <0.05 selected terms were then assembled into a master file displaying the up and down regulated GO terms per RWC per GO category and processed in R 4.2.2 using modified scripts from Bonnot, et al., (2019) to generate a GO term plot for each GO category. DAVID was also used to generate a list of InterPro terms for the accessions per RWCs separated by whether they were up- or down-regulated. The combined list was then reduced by only selecting terms that fell within a <0.05 FDR. Finally, DAVID was used to determine which pathways were involved during each RWC for each regulation. The online bioinformatics website http://www.bioinformatics.com.cn/static/others/jvenn_en/index.html was used to generate Venn diagrams for the GO terms. KEGG Koala Blast was used to annotate the full proteome owing to instances of unidentifiable proteins from original annotation.

Table 2-3: Recorded protein search configuration

Rule	Value
Protein database	E.curvula_RefProt_UP0003248975365prot_030722.fasta
Spectrum-level FDR	Auto cut
Cleavage residues	RK
Digest cutter	C-terminal cutter
Peptide termini	Fully specific
Maximum number of missed cleavages	2
Precursor tolerance	10.0 ppm
Fragment tolerance	Frag:qtof_hcd 20.0 ppm
Fragment tolerance version	2
Charge applied to charge unassigned spectra	1,2,3
Precursor mass max	10000.0
N-glycan search	None
O-glycan search	None
Off by x isotopes	-2,-1,0,+1,+2
Contaminants added	True
Decoys added	True
% Additional parameters	
Disulfide enable	False
Trisulfide enable	False
DSS Crosslink enable	False
Custom crosslink enable	False
Wildcard enable	0
Combyne cut off score	Auto
Protein FDR cut off	1%
Focused DB created	False
Export mzIdentML	True
Score version	2
Precursor_assignment_flags	2
Po_NumberMonosReturn	1
Lock mass list	None
% Modification searches:	
Common_ modifications_ max	1
Rare_ modifications_ max	1
% Fixed and variable modifications:	
Carbamidomethyl / +57.021464 @ C	Fixed
Oxidation / +15.994915 @ M	Common2
Deamidated / +0.984016 @ N, Q	Common2
% Custom modification text below	
% Glycan modifications:	
Show all N-glycopeptides	0
% Addition parameters:	
Product version	PMI-Byonic-Com: v3.8.13

The taxonomic ID for *E. curvula* (38414) was used as reference and the database termed *family_ eukaryotes. pep, genus_ prokaryotes. pep* was used. The '*reconstruct pathway'* option was used to generate metabolic pathways for the accessions wherein annotated accessions

were assigned to different broader categories either as a KEGG pathway or within a Brite table. For complementation with the transcriptome, the assigned Arabidopsis homologues was treated in the same way as the UniProt accessions *id est*, DAVID and Koala annotation and the associated fold changes as determined in Madden (2019) was assigned to the KEGG descriptions. For complementation to the proteome, KEGG descriptions which matched were combined into one master file which showed the expression across RWCs and tissues for the transcriptome and proteome. To determine whether the expression profile between the transcriptome and proteome were the same covariance was determined for 60, 25%, and AD NST, and 60 and 25% ST as these represented the RWC and tissue types that matched between the two data sets. ANOVA was also used to analyse variance between RWCs and tissues. Lastly, Spearman rank correlation was done per RWC and tissue between the two data sets to determine degree of correlation. For simplicity and illustrative purposes, the average fold-change for transcripts with multiple gene IDs for the same term with the same expression pattern was used. Likewise, for proteins with the same expression for multiple accessions, the average fold-change was used. The full summary of all Koala annotated DAPs can be found in Table 8-3.

Results and Discussion

Note on terminology used: terms such as up-regulated, overrepresented, and increased abundance all refer to proteins or terms that were significantly more abundant relative to the fully hydrated control. Likewise, terms such as down-regulated, underrepresented, and diminished refer to proteins or terms that were significantly lower in abundance relative to the fully hydrated control. For all values shown they are relative to the fully hydrated control and do not represent individual changes between RWCs. DAPs refer to differentially abundant proteins. Differentially expressed transcripts whose description matched that of the proteome are included where applicable and are indicated by [RNA] in the heat maps.

2.3.1 Protein quantification

For mass spectrometry analysis, the Centre for Proteomic and Genomic Research (CPGR) required protein to be at a minimal concentration of 1.25 $mg.mL^{-1}$ and requested that a total of 50 µg protein be supplied. In order to accomplish this from a limited amount of starting tissue, various methods were used to determine the most optimal protein isolation method. Of all the methods evaluated, the TRIZOL (TRI-reagent) from Sigma proved most reproducible yielding the highest protein concentration. A total of 30-60 mg fresh weight *E. nindensis* tissue was used for all isolations and quantified using the BCA protein quantification kit (Figure 2-3) as per the manufacturer's recommendations and estimated based on a bovine serum albumin standard curve. Protein concentration was corrected to factor in dilution and are presented in Table 8-1 in the appendix. Protein concentration was not corrected for starting material used as CPGR requested a fixed amount of protein for analysis. On average, 2.25 ± 0.65 $µg.µL^{-1}$ soluble protein was isolated using the modified TRIZOL method. Highest mean protein concentration was obtained for the 25 NST at 3.002 $µg.µL^{-1}$ with the lowest concentration within the acceptable range obtained for the 100 NST at 1.346 $µg.µL^{-1}$. This is to some degree representative of the added mass effect that water has on more hydrated tissues compared to more dehydrated tissues. An initial pilot study done at CPGR comparing their isolation method and the optimised method used herein indicated that the optimised method used herein yielded greater protein concentration suitable for mass spectrometry.

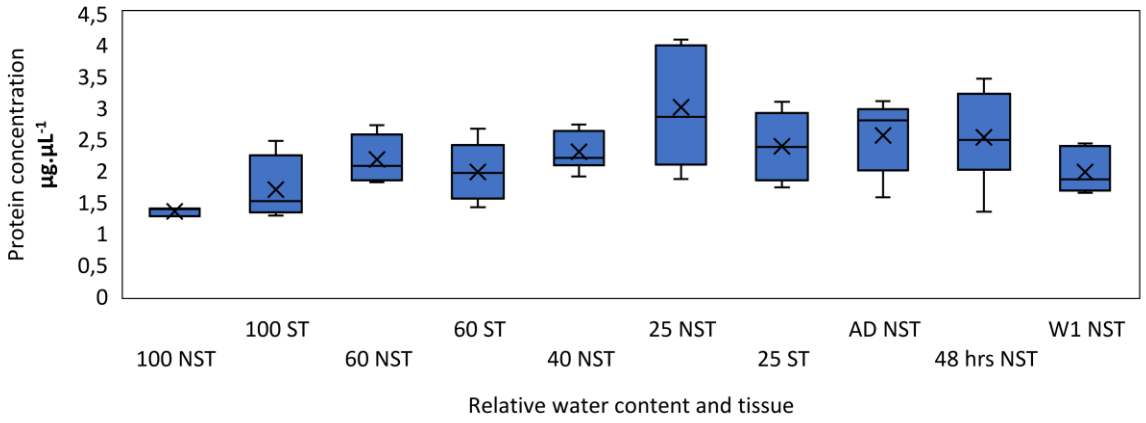


Figure 2-3: Proteome sample concentrations: Protein concentration per RWC as estimated from a bovine serum albumin standard curve using the BCA protein quantification kit following incubation at 37°C for 30 minutes. Standards and samples were done in triplicate of which 20 µL per standard and *E. nindensis* protein extract was added to 200 µL of quantification reagent. This standard curve equation was y=2.086x + 0.0982 R^2=0.997. n= at least 5 per RWC. N = 55. NST=non-senescent; ST= senescent; AD= airdry; W1= one week post-rehydration. All other numbers refer to the RWC.

CPGR requested at least three biological replicates per relative water content (RWC) but recommended five. Where possible, five biological replicates were included with some RWCs having 6 to decrease the potential variability for those particular tissue types at those RWCs. In addition, three different reference data bases, namely *Arabidopsis thaliana*, *Zea mays*, and *E. curvula* were used to determine which of these could annotate the *E. nindensis* peptide spectra best with *E. curvula* being able to map the most. The *E. curvula* protein data base was subsequently selected for annotating the proteome.

Data returned from CPGR contained only the significantly abundant proteins with a fold-change of 2 or more at an FDR <0.05 per RWC and was reported as normalised abundance values. Accession numbers associated with the comparisons (RWC vs fully-hydrated control) that indicated up- or down-regulated proteins were then uploaded onto DAVID for gene ontology enrichment selecting biological process, molecular function, and cellular compartment using the full proteomic accessions from all RWCs as background. Upon summarising using REVIGO and using a dispensability cut off of <0.05 the number of terms were dramatically reduced, and it was decided to not use the summarised terms but rather just the FDR selected terms from DAVID for functional enrichment analysis. Table 2-4 below summarises the outputs from CPGR as well as the number of GO Terms (not separated based on category) that were filtered for further analysis. A total of 3 965 differentially abundant protein accessions were returned across all RWCs. The 48hrs post-rehydration and 25% for NST returned the least amount of differentially abundant protein accessions whereas the 60%,40%, and air-dry (AD) for NST and 25% for ST returned differentially abundant protein

accessions greater than 500. Upon submitting the individual accession numbers for up- and down-regulated proteins onto DAVID, the number of differentially abundant GO terms was significantly reduced, implying a high degree of duplication and/or redundancy in the annotated differentially abundant protein file. On average the DAVID analysis reduced up-regulated accessions to GO terms by 92.34% and reduced down-regulated accessions by 91.61%. In accordance with accepted bioinformatic practices, by applying a <0.05 FDR cut off on DAVID GO terms the number of GO terms was further reduced to 71 up-regulated and 101 down-regulated across all RWCs. Using this stringent cut off, no up- or down-regulated GO terms were isolated for the 60% and 25% ST, and 25%NST. Further evident in Table 2-4 is a decline in the number of DAPs relative to differentially expressed genes (DEGs) as RWC declines, for example at 25% RWC the transcriptome reported 750 DEGs as up-regulated whereas the proteome only had 134 DAPs.

Table 2-4: *Eragrostis nindensis* proteome and transcriptome summary indicating total number of differentially expressed genes (transcripts) and abundant proteins (accessions) and the filtering numbers based on gene ontology enrichment for proteome only using DAVID. Numbers do not take duplicate terms into account. *RWCs only present in transcriptome.

		Transcriptome		**Proteome**					
		Transcripts		**UniProt accessions**		**DAVID GO enrichment**		**<0.05 FDR**	
Tissue	**RWC**	**Up**	**Down**	**Up**	**Down**	**Up**	**Down**	**Up**	**Down**
NST	60	368	342	377	325	55	52	6	13
	40	627	700	312	334	46	55	13	31
	25	750	941	134	55	23	3	11	0
	AD	869	967	357	350	45	46	11	24
	12hrs*	886	1059	-	-	-	-	-	-
	48hrs	-	-	100	54	24	13	11	9
	W1	-	-	238	218	42	62	17	8
ST	60	374	299	206	231	42	61	2	11
	25	667	807	319	365	27	41	0	5
	18*	747	783	-	-	-	-	-	-

2.3.2 Gene ontology enrichment

All GO terms irrespective of FDR values were taken together to determine the degree to which the terms overlapped with one another after removing duplicate GO terms per RWC. By only considering the dehydration RWCs, a total of 79 up- and 85-down DAPs were returned with 19 terms overlapping between the two. By including the rehydration samples (Figure 2-4), three terms across the RWCs were shared by all, those being terms associated with *peroxidase activity*, *heme binding*, and *response to oxidative stress* which comprised predominantly of antioxidant enzymes such as catalase, glutathione peroxidase, L-ascorbate peroxidase, and peroxidase as well as nitrate and sulphite reductases.

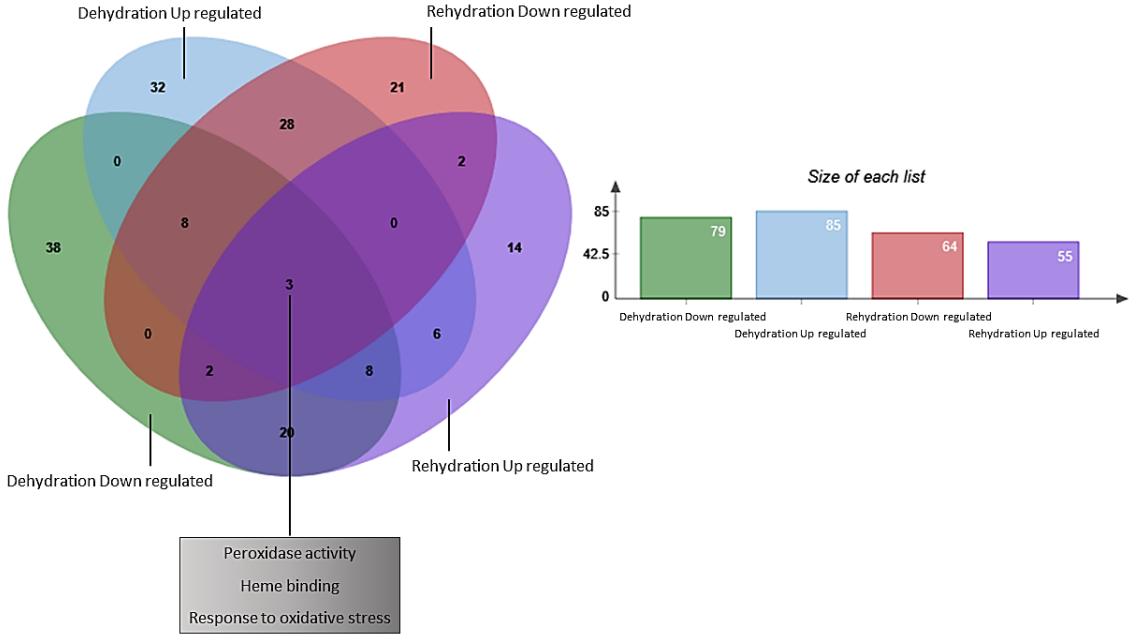


Figure 2-4: Venn diagram for all DAPs: Venn diagram generated using all up and down DAPs from across RWCs for *E. nindensis* proteome. Graph includes the 3 terms that were significantly abundant across all RWCs. All DAPs are significant with a fold-change of >2 and FDR <0.05. (Bardou, et al., 2014)

To further elucidate which terms were associated with up and down DAPs, the same list of terms was used however, this time considering the up and down DAPs for the dehydration RWCs separately, shown in Figure 2-5. A total of 5 terms were shared across all RWCs that were in abundance relative to the fully hydrated control. These terms were associated with *ion transport* which included ferritin and ferroxidase, *Clathrin binding and assembly*, and *1-phoshatidylinositol binding* both of which comprising proteins with an ANTH-domain which are membrane-bound proteins involved in PIP2 binding in Clathrin-coated pits. There were no overlapping terms for the 60% RWCs in either tissue. A total of 13 and 3 unique terms for 25% ST and AD NST respectively were isolated. Those terms associated with the 25% ST suggest

that there was a high degree of activity occurring at the mitochondrion with an equally high degree of activity with DNA, RNA, and a suggestion of RNA translation.

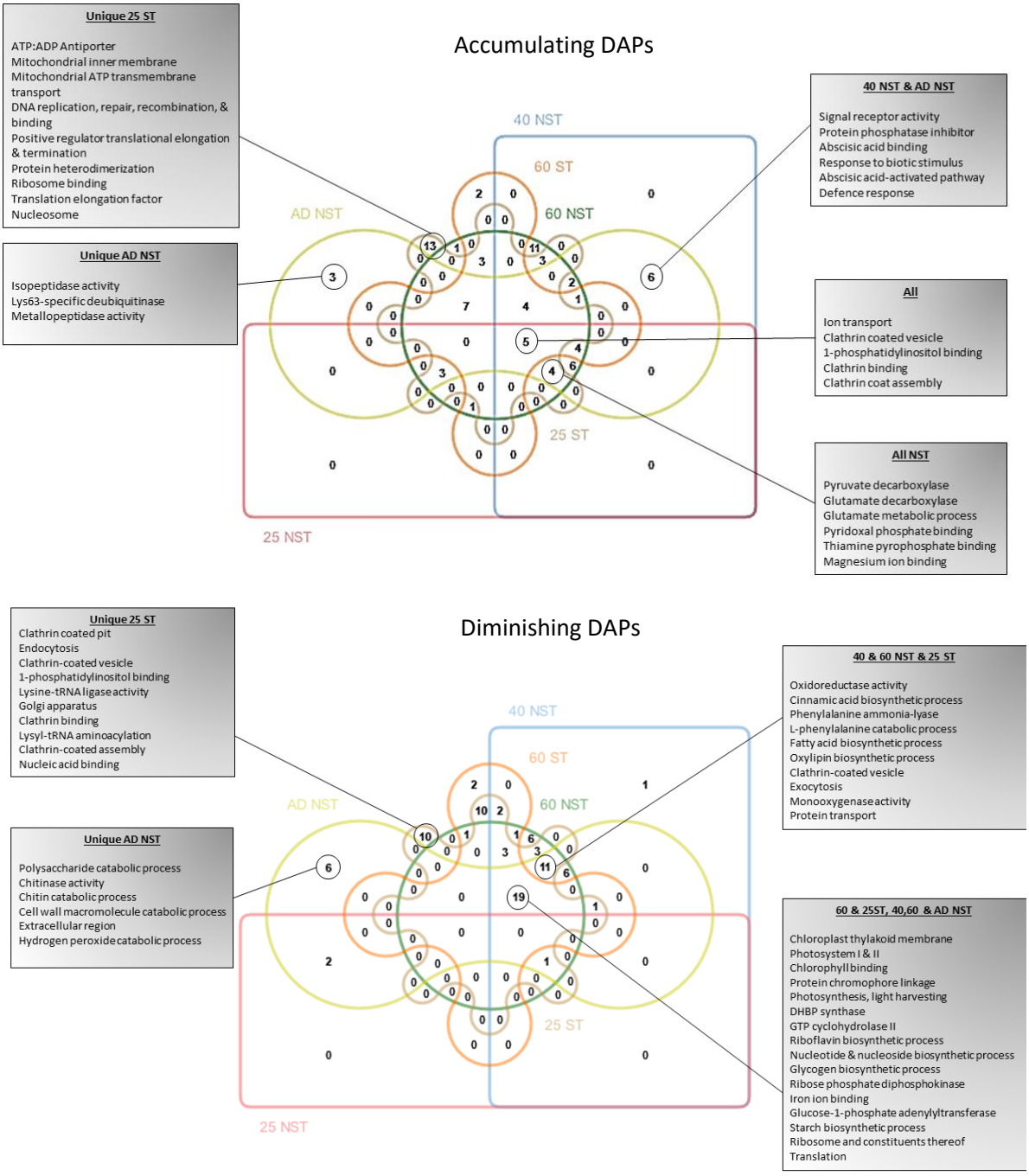


Figure 2-5: Venn diagrams for all accumulating and diminishing DAPs for dehydrated RWCs: Venn diagrams in Edward's style showing unique and overlapping DAPs for those accumulating (top diagram) and those that diminished (bottom diagram) for all dehydration RWCs. Some areas of overlap and uniqueness have the associated GO terms displayed. All DAPs are significant with a fold-change of >2 and FDR < 0.05. (Bardou, et al., 2014). NST=non-senescent; ST= senescent; AD= airdry All other numbers refer to the RWC

An interesting observation was that there were 6 shared terms between the 40% and AD NST with terms associated with hormone and signal reception. The non-senescent tissue RWCs had a total of 4 shared terms associated with pyruvate and glutamate processes which seems to imply that these processes were switched on in the NST and were maintained throughout dehydration. With regards to the diminished DAPs there were no overlapping terms shared by all the RWCs. The greatest overlap occurred at the 60% and 25% ST, and 40%, 60%, and AD NST with a suite of terms associated with a number of processes, most notably terms associated with photosynthetic processes which are related to the poikilochlorophyllous nature of *E. nindensis*. Other noteworthy diminishing DAPs include a number of terms associated with amino acid metabolism, notably phenylalanine in the 40% and 60% NST, and 25% ST. Terms associated with Clathrin appeared to be both in abundance and diminished in the 25% ST, but this could be due to accessions that are related but when assigned to a GO term were then placed in the same category.

By applying a more stringent <0.05 FDR filter on all the GO terms returned from DAVID, the number of terms per RWC per regulation was greatly reduced. By considering the biological processes as shown in Figure 2-6 a key feature of the poikilochlorophyllous nature of *E. nindensis,* which is the disassembly of the thylakoid and associated photosynthetic components, was apparent with the underrepresentation of *light-harvesting complexes* of photosynthesis in most of the dehydrating tissue with a markedly greater reduction at 40% NST relative to the other RWCs and only shows up-regulation 1-week post-rehydration. This suggests that although thylakoid disassembly starts at 60% RWC, 40% RWC can be seen as a tipping point for committed disassembly. This was also reflected in the chlorophyll abundance and general appearance of *E. nindensis* where below 40% RWC little to no chlorophyll was detected and the appearance of leaves was akin to early senescence (Madden, 2019) . Other noteworthy changes include the apparent underrepresentation of amino acid metabolism, in particular valine, isoleucine, and phenylalanine all of which comprising a hydrophilic functional group as well as underrepresentation of the polar charged amino acids glutamine and asparagine at 48hrs post-rehydration.

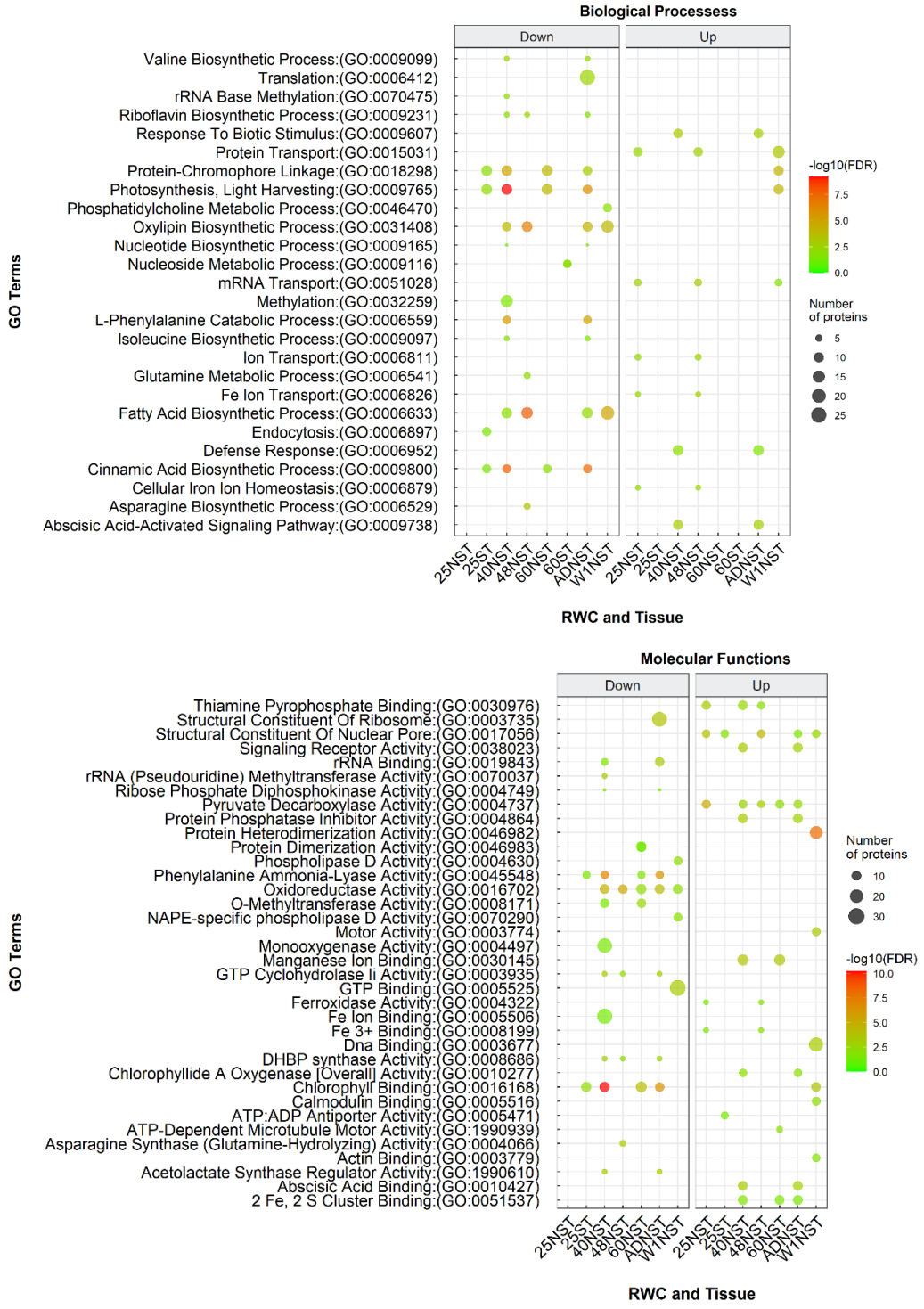


Figure 2-6: GO Terms associated with biological and molecular processes in *E. nindensis*: GO Term plot showing the selected GO Terms with an FDR <0.05 and the log(FDR) as a scale with number of proteins associated with that term for all RWCs down- and up-regulated relative to the fully hydrated control. Plot generated in R 4.2.2 (Bonnot, et al., 2019). All DAPs are significant with a fold-change of >2 and FDR < 0.05. NST=non-senescent; ST= senescent; AD= airdry; 48=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

Interesting was the overrepresentation of protein transport in 25% NST, 48hrs and 1 week post-rehydration with concomitant up-regulation of mRNA transport, suggesting active transport of presumably important proteins and mRNA in preparation for rehydration and then

subsequent redistribution upon partial and near full metabolic activity. Which proteins and mRNA, and the direction of transport these might be taking, is unclear from the enrichment analysis but could be indicative of active transcription and translation occurring at 25% RWC. There is precedent in the literature for such activities at such low water contents (Farrant & Hilhorst, 2021). The absence of these GO terms in the ST suggests that at 25% RWC the two tissues are behaving differently. Biological processes involved in lipid metabolism also showed underrepresentation with the first evidence of decreasing fatty acid metabolism at 40% NST with the greatest decrease occurring at 48hrs post-rehydration. This seems to suggest that the synthesis of new fatty acids was halted, presumably as a means to redirect carbon towards other processes. The small yet present underrepresentation of nucleotide biosynthesis in 40% and air-dry NST suggests that the production of nucleotides, the monomeric units for nucleic acids, is not favoured in the NST at these two RWCs and could be indicative of a cessation of nucleic acid biosynthesis which in turn can be interpreted as a halt in RNA synthesis. The down regulation of nucleoside metabolic processes at 60% NST is likely what is contributing to the low abundance of nucleotide biosynthesis in the 40% NST. Considering the 40% and air-dry and associated molecular functions, the term *ribose-phosphate dehydrogenase* is underrepresented. Ribose-5-phosphate, which is primarily produced via the pentose phosphate pathway (PPP), is a precursor for a myriad of biomolecules including nucleic acids.

The apparent underrepresentation of ribose-phosphate dehydrogenase, together with associated down-regulation of nucleotide biosynthesis, provides the first evidence of a remodulation to carbohydrate related processes, further corroborated with the marked accumulation of *pyruvate decarboxylase*, an enzyme that decarboxylates pyruvate to acetaldehyde and CO_2 in all the dehydrating NSTs. Over and above the underrepresentation of nucleotide biosynthesis at these two RWCs, there were a number of other GO terms that only emerge at these two RWCs, those being overrepresented terms associated with *response to biotic stimulus, defence response, abscisic acid (ABA)- activated signalling and binding, 2-Fe-2S- cluster binding, chlorophyllide a oxygenase activity, protein phosphatase inhibitor activity,* and underrepresented terms associated with *RNA binding* and *acetolactate synthase regulator activity* in the molecular functions category and *ribosome* in cellular component (Figure 2-8).

The occurrence of ABA-related GO terms at 40% and AD NST RWCs is highly suggestive of these two serving as tipping point RWCs and that once *E. nindensis* reaches the 40% state it has crossed the point of drought response and needs to ensure that the foundation of surviving desiccation is in place as it enters the desiccation response. The resurrection plant, *Xerophyta schlechteri*, has been proposed as the model plant for possible enhancement of economical monocotyledonous crops such as maize (Farrant & Hilhorst, 2022). Through numerous

studies focusing on its anatomical and metabolic (Radermacher, et al., 2019) to proteomic (Gabier, et al., 2021) changes a clear distinction between early to mid to late stages of desiccation are evident. In the model proposed for *X. schlechteri*, Farrant and Hilhorst, (2022) demonstrated that during the early stages of drying (up to approximately 60% RWC) the response between sensitive crops and *X. schlechteri* are in a broad sense the same but are crucial for the continuation into the mid-stage of desiccation (approximately 40% RWC). At this stage commitment to the desiccation response is initiated in *X. schlechteri* such as the dismantling of thylakoids to the extent that the leaves resemble dry seeds (Radermacher, et al., 2019). For *E. nindensis*, it appears that at this 40% RWC state, commitment to the desiccation response was also initiated, demonstrating that perhaps with respect to *X. schlechteri* and *E. nindensis* the 40% RWC (or mid-stage) is an important milestone and beyond this stage is when true vegetative desiccation tolerance emerges. At the air-dry state, which is at the desiccated state, *E. nindensis* might be preparing itself to enter a quiescent state and may even lay down proteins and mRNAs required for rapid recovery upon rehydration where for instance in *X. schlechteri* it has been proposed that during the dismantling of thylakoids, transcripts necessary for the re-synthesis of photosynthesis are made and stored (Costa, et al., 2017b). ABA of course has a direct effect on plant response to the environment where in orthodox seeds it promotes seed dormancy (Kim, et al., 2013). Protein phosphatases are primarily responsible for dephosphorylation of proteins and act oppositely to protein kinases. The phosphorylated state of a number of proteins and enzymes can to some degree be likened to their overall activity. Over and above the dephosphorylation role, these proteins, like the protein kinases are integral to signal transduction pathway which ultimately have an impact on a range of cellular processes from transcription to apoptosis. The overrepresentation of inhibitor activity on protein phosphatases suggests that at 40% and in air-dry tissues *E. nindensis* modulates the phosphorylated state of its protein pool. In the context of seed dormancy and its relatedness to vegetative desiccation tolerance, SNF1-related protein kinase 2 (a subgroup of protein phosphatase 2 henceforth referred to as SnRK2) positively regulate ABA through binding to bZIP transcription factors (Fujii, et al., 2007). Knock-out mutants in Arabidopsis displayed a lack of seed dormancy (Fujii & Zhu, 2009). The positive regulation of SnRK2 is greatly inhibited by the presence of another subgroup of protein phosphatases, PP2C, which in turn have an effect on the regulation of ABA. Their presence acts in antagonism to SnRK2 by promoting seed germination (Bhaskara, et al., 2012). Thus, in this context, it is tempting to speculate that something similar might be occurring where the *phosphatase* inhibition might be in favour of inhibiting PP2C to promote SnRK2 regulation of ABA such that a seed-dormancy-like state is achieved.

The increase in *chlorophyllide activity* is primarily attributed to the recycling of chlorophyll and is presumably in such high abundance owing to the disassembly of the thylakoids. Chlorophyllides have a conserved Rieske domain and thus also explains the increase in the *2-Fe-2S cluster binding*. The underrepresentation of *RNA binding* is a unique conundrum which at first glance seems to contradict the speculation of active transcription and translation occurring at lower RWCs but could simply refer to compartments or processes involved in RNA metabolism and does not necessarily reflect cessation of RNA and downstream translation. The underrepresentation of *acetolactate synthase* regulation clearly demonstrates that a halt in branched chain amino acid biosynthesis is evident since it is the first enzyme in this pathway. While it too has a Rieske domain and two thiamine pyrophosphate (TPP) domains, the Rieske clusters are clearly directed towards chlorophyllide and the TPP are redirected elsewhere, likely towards other enzyme complexes such as pyruvate dehydrogenase or succinate dehydrogenase.

Iron appears to be an important ion with processes such as *iron transport* and *cellular iron homeostasis* being overrepresented in 25% NST and then again at 48hrs post-rehydration. When looking at other iron related terms, *ferroxidase* and *Fe^{3+} binding* was functionally enriched and overrepresented at the same RWCs. The chloroplast can be considered as one of the highest iron dense organelles (Schmidt, et al., 2020) and disassembly of the thylakoids can inherently increase iron pools which, as indicated above, primarily occur at 40% RWC. Given that *E. nindensis* does not accumulate as much photosynthetically-derived ROS when compared to homoiochlorophyllous resurrection plants, it does however increase the iron pool, and as such a regulatory mechanism is required to alleviate iron toxicity as well as iron-dependent cell death. Programmed cell death (PCD) is an orchestrated set of processes which allows cells to drive specific sporo- and gametophytic developmental stages but also plays a vital role in stress response (Distéfano, et al., 2021). Ferroptosis, which is an iron-dependent form of PCD, is a relatively new concept in plant cell biology and is characterized by the accumulation of lipid ROS which is mechanistically dissimilar to vacuolar cell death for example. The process is generally very similar to what was first observed in mammalian systems (Dixon, et al., 2012), those being Fe-dependent ROS production and accumulation, subsequent lipid peroxidation (of which polyunsaturated fatty acids containing phospholipids are the main targets), and overall cellular depletion of glutathione. Heat has been shown to play a role in the activation of ferroptosis in Arabidopsis (Distéfano, et al., 2017). In order to prevent ferroptosis, it is proposed that *E. nindensis* up-regulates ferroxidase which is responsible for the conversion of Fe^{2+} to Fe^{3+} while also increasing the Fe^{3+} binding capacity (Figure 2-7). The associated protein identified through InterPro annotation is ferritin which is

responsible for the sequestration of Fe^{3+}. Ferritin is noticeably absent in the 1 week post-rehydration tissue and in the 25% ST (Figure 2.7).

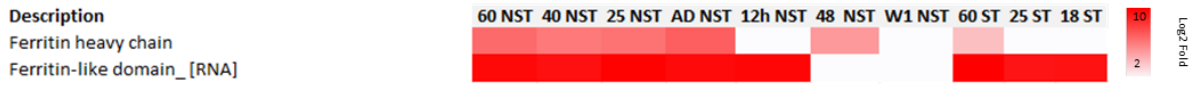


Figure 2-7: Ferritin protein and transcripts: Heatmap showing ferritin as log2 fold values. Note that 12h NST and 18 ST are exclusive to the transcriptome whereas 48h NST and W1 NST are exclusive to the proteome. Where there were no values for missing protein and transcripts for the exclusive RWCs, 0.0001 was inserted. Fold changes displayed here represent the greatest/lowest for the respective description. Transcript indicated by [RNA]. NST=non-senescent; ST= senescent; AD= airdry; 12h= 12hrs post-rehydration; 48=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

The presence of these two GO terms at 48hrs post-rehydration implies that the simple re-greening of tissue is not a benchmark for concluding that the plant has returned to pre-dehydration status and further demonstrates the need of *E. nindensis* to ensure adequate and frequent water supply before returning to normal metabolic activity. This seemingly simple yet elegant process removes iron which could otherwise compromise the integrity of the plasma membrane which needs to be maintained in order to ensure that the cell does not collapse during dehydration. The absence of ferritin at 1-week post-rehydration implies that at least in terms of sampling times, is when *E. nindensis* is back to normal activity. The absence of it at 25% ST but presence at 60% illustrates two unique features. First is that at 60% RWC the ST is likely to be capable of returning to regular metabolic activity should it receive water since it has already been suggested that 40% RWC is when commitment to desiccation response is made. Secondly, its absence at 25% illustrates for the first time a positive association to PCD in the ST but suppression of PCD in the NST.

The air-dry NST displayed a marked reduction in *rRNA binding* with a great number of proteins involved with structural components of the ribosome also with marked reduction. This seems to suggest that either translation has ceased to occur or *E. nindensis* has entered a quiescent state. However, this is seemingly in contradiction to the number of proteins assigned to *nuclear pore* in the cellular compartment GO analysis (Figure 2-8).

Cinnamic acid, an allelochemical involved in a myriad of plant functions, is strikingly diminished throughout dehydration. The protein associated with it is *phenylalanine-ammonia lyase* whose GO terms was also denoted as functionally diminished in the biological terms. The link between the two terms is quite simple. Phenylalanine-ammonia lyase catalyses the conversion of phenylalanine to trans-cinnamic acid during phenylpropanoid biosynthesis. Cinnamic acid when applied exogenously to *Glycine max* (soy) resulted in increased indole-3-acetic acid oxidase activity as well as increased lignification causing cell wall stiffening (Salvador, et al., 2013). Lignification and stiffening of cell walls were not observed through ultrastructural studies (Madden, 2019) and indeed Moore, et al., (2013) show an increased

“plasticisation” of cell walls during dehydration of *E. nindensis,* suggesting that a decline in cinnamic acid might facilitate the latter, required for mechanical stabilisation in the desiccated state.

An underrepresentation of *GTP cyclohydrolase II activity* at 40%, AD NST, and 48hrs post-rehydration implies a decrease in folate synthesis which is ultimately involved in overall riboflavin metabolism and further exemplified with the decreased abundance of *3,4-dihydroxy-2-butanine-4-phosphate (DHBP) synthase*. Riboflavin, a B-vitamin, is an important molecule in the cell and primarily functions as a cofactor in a number of biochemical reactions. However, the nature of it makes it very prone to a host of abiotic-stress-induced alterations such as oxidation and hydrolysis (Hanson, et al., 2016). In their review, the authors further argue that while application of riboflavin can improve overall stress response, also shown by Guhr, et al., (2017) in *Agaricus bisporus* for example, the associated ‘high-cost’ and ‘high-risk’ in making riboflavin under stressed conditions might not be favourable. In the case of *E. nindensis*, the down-regulation of biosynthesis of riboflavin might actually favour the recycling of carbon towards other processes that provide greater protection than producing compounds which might be targets for oxidative damage. There was a marked decrease in abundance with oxidoreductase activity, specifically where a redox reaction transfers hydrogen or electrons from one donor and two oxygen atoms are incorporated into a donor, throughout drying and rehydration. Oxidoreductase enzymes encompass one of the largest enzyme groups and are involved in innumerable processes ultimately involved in REDOX homeostasis. The protein primarily responsible for the associated GO terms is lipoxygenase which is discussed in more detail in section 2.3.2.5 below.

Over-represented terms for cellular components (Figure 2-8) further supports the cessation of photosynthetic activity during dehydration with a reduction in components related to the PSII, PSI, the thylakoid membrane, and oxygen evolving complex which only shows accumulation at 1-week post-rehydration with the 40% NST having the greatest enrichment in most of these terms and the air-dry tissue coming in second. At 40% and 60% NST exclusively, there was increased abundance of terms associated with the cell wall and apoplast. The associated protein with these GO terms was annotated as a germin-like protein. It is a type of glycoprotein which is associated with the cell wall (Patnaik & Khurana, 2001) and has been shown through recombinant expression of *Camellia sinensis* cDNA in *E. coli* to have oxalate oxidase or superoxide dismutase activity (Fu, et al., 2018). In rice, the expression profile of a suite of germin-like proteins were noted under salt and drought stress (Anum, et al., 2022) whereas in *Gossypium hirsutum* upregulation of a germin-like protein was noted when cotton was challenged by fungal pathogenicity and application of jasmonic acid (Pei, et al., 2019), implicating germin-like proteins as having a positive role in both biotic and abiotic stress. The

association given to the cell wall and apoplast could be supportive of increased antioxidant activity at these sites as a means to reduce membrane lipid peroxidation. Its functional enrichments in the NST exclusively during early drought stress and commitment to desiccation response illustrates tighter regulation of membrane lipid oxidation states in the NST compared to the ST.

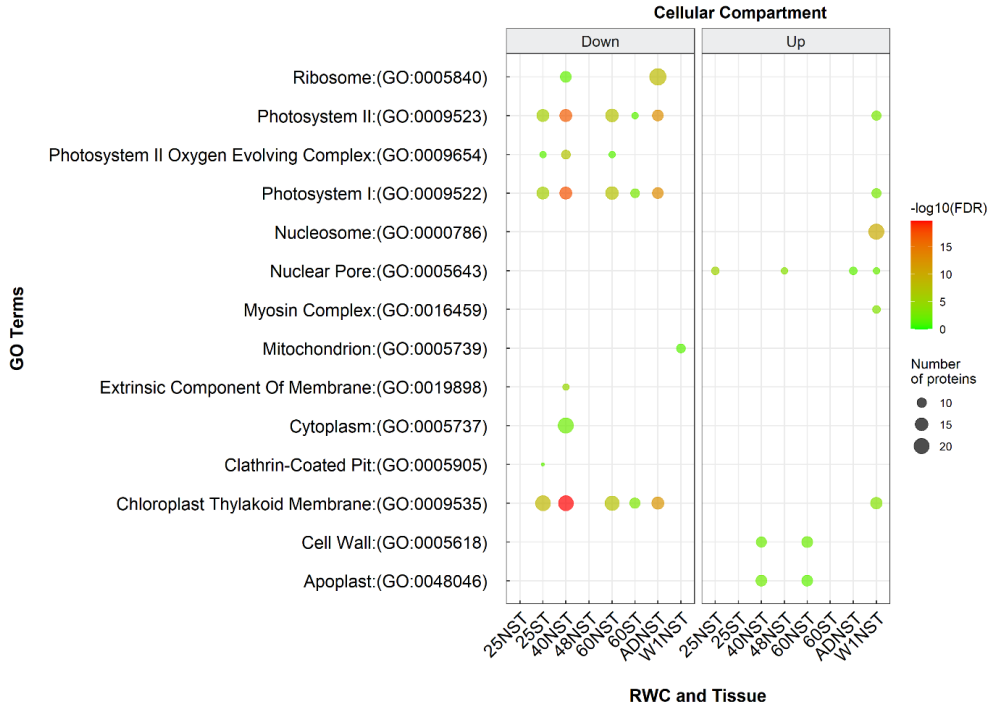


Figure 2-8: Over-represented cellular components in *E. nindensis*: GO Term plot showing the selected GO Terms with an FDR <0.05 and the log(FDR) as a scale with number of proteins associated with that term for ST and NST along the dehydration and rehydration grouped as down- and up-regulated. Plot generated in R 4.2.2 (Bonnot, et al., 2019). All DAPs are significant with a fold-change of >2 and FDR < 0.05. NST=non-senescent; ST= senescent; AD= airdry; 48=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC

Disruption to the cellular membrane integrity inevitably leads to cellular leakage which in turn can escalate the progression of PCD. The low number of proteins associated with the down-regulation of clathrin-coated pits in the 25% ST is a small yet interesting observation. Clathrin-related GO terms which did not make the FDR cut off are present in the majority of tissues during dehydration with concomitant increase in terms associated with Golgi and ER vesicle transport, suggesting that there is a high degree of vesicular transport occurring.

Pathway analysis using DAVID was done using the accessions from each RWC separated into those that were over- and underrepresented and compared against the reference proteome background of all accessions to elucidate whether key pathways were involved in dehydration. Table 2-5 lists the summary of pathways and interactions deemed significant with an FDR <0.05. Results from the pathway analysis demonstrated the importance of photosynthetic shutdown with ligands associated with photosynthesis showing decreased activity across dehydration RWCs and increased activity at 1-week post-rehydration.

Table 2-5: Summary of pathways and interaction (ligand) output from DAVID showing only terms with FDR <0.05 grouped by RWC and regulation

RWC and tissue	Regulation	Enzyme/Ligand	Count
25 NST	**Up**	Ferroxidase: (1.16.3.1)	4
		Pyruvate Decarboxylase: (4.1.1.1)	8
25 ST	**Down**	Chlorophyll: (KW-0148)	11
		Chromophore: (KW-0157)	11
		S-adenosyl-L-methionine: (KW-0949)	10
	Up	Pyruvate Decarboxylase: (4.1.1.1)	8
40 NST	**Down**	Chlorophyll: (KW-0148)	11
		Chromophore: (KW-0157)	11
		S-adenosyl-L-methionine: (KW-0949)	8
		Schiff base: (KW-0704)	8
	Up	Hydroxymethylglutaryl-CoA Lyase: (4.1.3.4)	9
		Iron: (KW-0408)	48
		Iron-sulphur: (KW-0411)	14
		Manganese: (KW-0464)	13
		Pyruvate Decarboxylase: (4.1.1.1)	8
48hrs NST	**Down**	Asparagine Synthase (Glutamine-Hydrolysing): (6.3.5.4)	5
		GTP Cyclohydrolase II: (3.5.4.25)	4
		Oxidoreductase : (1.13.11.-)	6
	Up	ATP-binding: (KW-0067)	25
		Ferroxidase: (1.16.3.1)	4
		Pyruvate Decarboxylase: (4.1.1.1)	6
60 NST	**Down**	Chlorophyll: (KW-0148)	12
		Chromophore: (KW-0157)	12
		Magnesium: (KW-0460)	27
		Schiff base: (KW-0704)	8
	Up	Iron-sulphur: (KW-0411)	14
		Manganese: (KW-0464)	15
		Pyruvate Decarboxylase: (4.1.1.1)	8
60 ST	**Down**	S-adenosyl-L-methionine: (KW-0949)	10
		Schiff base: (KW-0704)	9
	Up	Ferroxidase: (1.16.3.1)	4
AD NST	**Down**	Chlorophyll: (KW-0148)	9
		Chromophore: (KW-0157)	9
		Magnesium: (KW-0460)	31
	Up	Iron-sulphur: (KW-0411)	13
		Pyruvate Decarboxylase: (4.1.1.1)	8
W1 NST	**Down**	GTP-binding: (KW-0342)	28
		Phospholipase D: (3.1.4.4)	8
	Up	ATP-binding: (KW-0067)	40
		Chlorophyll: (KW-0148)	10
		Chromophore: (KW-0157)	10
		Nucleotide-binding: (KW-0547)	39
		Peroxidase: (1.11.1.7)	8

Pathways involved with pyruvate decarboxylase were once again overrepresented in the same RWCs as above. The overrepresentation of Hydroxymethylglutaryl-CoA (HMG-CoA) Lyase, an enzyme involved in the production of acetyl-CoA and acetoacetate, implicates the two products as important precursor metabolites for a myriad of pathways including type II polyketide biosynthesis for example. HMG-CoA can be derived through the degradation of many larger compounds, including amino acids such as leucine, isoleucine, and valine. Ferroxidase and iron-associated pathways/ligands were once more present in the overrepresented RWCs further implicating tight regulation over iron homeostasis.

2.3.1 Gene ontology enrichment: General conclusions

The stringent FDR filtering of GO terms provided a summarised understanding of the proteomic profile of *E. nindensis* during dehydration and subsequent rehydration. Taken together the terms clearly illustrated the importance of shutdown of photosynthesis, which not only validated the bioinformatic pipeline used herein but also demonstrates a fundamental flaw in enrichment analysis by simply ignoring lesser 'enriched' terms. The overall decline in amino acid and protein metabolism suggests to some degree control over which proteins are made and perhaps redirection of amino acids towards other pathways to produce metabolites which are required during dehydration. The overall activity surrounding pyruvate illustrates not only its central role in global metabolism but also alterations to central metabolism during dehydration. It is further evident that there is some degree of regulation occurring with RNA and subsequent translation and an overall decline in lipid metabolism. The functional enrichment further demonstrated that at moderate and severe desiccation stress, *E. nindensis* undergoes tipping point processes where upon reaching that state a suite of processes and proteins are (in)activated to ensure firstly that the plant prepares itself for severe desiccation stress and secondly that it prepares itself to enter a quiescent state, similarly to what one might argue is occurring in orthodox seeds during seed maturation. The enrichment of terms associated with iron demonstrated that perhaps *E. nindensis* while simultaneously disassembling its thylakoids, actively prevents ferroptosis from occurring in the NST. One might even argue that the prevention of ferroptosis is an indication that reaching air-dry state is tightly controlled such that programmed cell death is perturbed, further exemplifying the reaching of a quiescent state in *E. nindensis*. The shutdown of riboflavin biosynthesis illustrates that *E. nindensis* has preference when it comes to oxidative stress response by withholding the production of compounds which in a classical sense could alleviate oxidative stress in leu of sacrificing the carbon which could otherwise be shuttled elsewhere.

The gene ontology analysis done herein is not without its limitations. Since the algorithms associated with assigning terms and the statistics associated with deciding significance can

be heavily influenced by the total number of accessions related to those terms and by applying a concomitant <0.05 FDR stringency, a large number of GO terms are completely ignored in favour of more enriched terms. For example, in the non-filtered selection of terms chaperone activity is increased in the 40% and 60% NST but is completely ignored in favour of terms associated with iron homeostasis since there were 'functionally' more accessions. While the gene ontology enrichment analysis demonstrated the importance of photosynthetic shut down in *E. nindensis*, characteristic hallmark signatures of stress and desiccation-tolerance response such as the accumulation of sucrose, heat shock proteins, LEAs, and antioxidants to name a few were noticeably absent. Furthermore, the annotation using the *E. curvula* database returned a large number of proteins which were annotated as Uncharacterised protein or as domain of unknown function (DUF). While it is expected with mass spectrometry based sequencing to have peptide spectra that do not correspond to any known peptide, the larger than expected category of uncharacterised proteins and DUFs pointed towards the need for increased experimental functional analysis, especially in the realm of plant sciences. While it would be erroneous to put sole blame on bioinformatic limitations there are a few instances where the tools used for annotation presented problems. For instance, on multiple occasions an annotated description from CPGR would have multiple accessions assigned to it and upon closer inspection, one or two might be uncharacterised as is from the *E. curvula* database but a much greater number of them could actually be manually annotated. The InterPro annotation assisted in assigning annotations to the majority of proteins from the proteome but had one fundamental flaw in that it would return results that would state for instance *AAA-domain containing protein* which does not provide any context of where it might be involved or what its function might be. It did however allow for annotation of proteins involved in chaperone and seed development processes such as heat shock proteins and late embryogenesis abundant proteins.

Another inherent limitation is the fact that the *E. curvula* UniProt database was used to identify peptides based on the rationale that curvula is a closer relative to *E. nindensis* than the model organisms *Arabidopsis thaliana* or *Zea mays*. In fact, in a pilot study the *E. curvula* database mapped significantly more peptides than did the other two. Using a non-model organism's database to annotate another non-model organism's peptides resulted in severe limitations with regards to compatible analysis platforms. This inherently meant that complementation with the transcriptome which was annotated with Arabidopsis homologues made direct comparison between the two difficult to do. In the interest of further expanding on the gene ontology enrichment and to gain a better understanding of the fundamental processes which were being influenced by dehydration and rehydration between the senescent and non-senescent tissue, the entire proteome was re-annotated using KEGG's Koala Blast annotation

tool. The results for this will be presented next accompanied with a discussion of the processes and their implications in desiccation tolerance in *E. nindensis.*

2.3.2 KEGG analysis

Using the online Koala Blast, accessions were uploaded and blasted using the *E. curvula* taxonomic ID against the aforementioned database to assign general classifications and to reconstruct metabolic pathways. Of the 2857 accessions submitted, only 59.9% was annotated by Koala Blast. This could likely be due to the high number of related terms across all RWCs as well as not accounting for non-enzymatic proteins. A full list of annotated accessions separated by KEGG category per RWC can be viewed in Table 8-3 in the Appendix. Of the 1712 DAPs which were annotated (Figure 2-9 A), 17% were annotated as belonging to processes involved in carbohydrate metabolism, suggesting that during the dehydration and rehydration cycle remodulation of carbohydrate metabolism is important. Surprisingly, a significantly large number of accessions were annotated as belonging to processes involved with genetic information processing. Amino acid metabolism, originally thought to be functionally enriched by gene ontology, only accounted for 5% of the global proteome. Energy metabolism accounted for 6% of the annotated proteome suggesting that perhaps in concert with shifts in carbohydrate metabolism, redirection of energy generating processes are also modulated. Upon reconstructing the metabolic pathways involved and taking individual RWCs into account it became evident that between RWCs and tissues differences in KEGG and Brite categories are noted. Illustrated in Figure 2-9 B below are the mean log fold changes for each RWC taking no evidence as zero expression into account. General trends across RWCs were for instance fluctuating changes in amino acid metabolism with an overall trend of increased activity associated with *folding, sorting,* and *degradation* except in the lower RWCs of NST. An overall trend of increased fold-change related to *chaperone* and *folding* was noted with the lower RWCs in NST accounting for greatest fold changes. Carbohydrate metabolism appeared to also fluctuate between RWCs though was predominantly in favour of down-regulation. Significant variation was evident in genetic information processing which includes processes involved with DNA repair, transcription, and translation, where for instance at 60% NST the greatest fold change was noted, fluctuating changes beyond 60% RWC, and a seemingly underrepresentation at 1 week post-rehydration. In agreement with the gene ontology enrichment a great fold change in metabolism related to cofactors and vitamins was noted, though by looking at it from this perspective it is not evident whether it is in favour of catabolism or anabolism. Lipid metabolisms showed a marked decrease in mean fold change at the 1 week post-rehydration time point with a general trend of increasing fold change throughout dehydration, suggesting that lipid metabolisms in general

plays an integral part during dehydration, which the gene ontology enrichment attributed towards a decline in fatty acid biosynthesis.

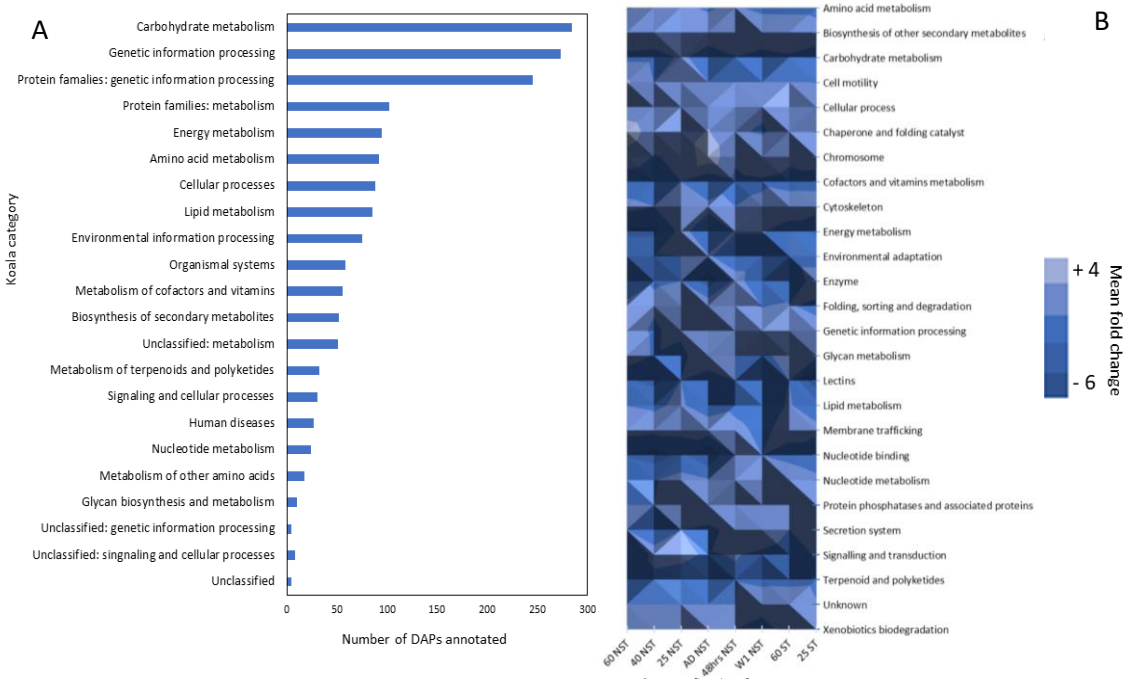


Figure 2-9: Koala annotation and KEGG pathways for *E. nindensis* DAPs: Bar graph showing the number of DAPs annotated by Koala Blast (A) and mean fold change of each RWC upon assigning to KEGG or Brite categories (B). All DAPs are significant with a fold-change of >2 and FDR < 0.05. ST= senescent; AD= airdry; 48=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

The biosynthesis of secondary metabolites was noticeably in favour of negative fold-change suggesting that although the accumulation of secondary metabolites in response to stress might be a common strategy, for instance in the synthesis of anti-pathogen alkaloids, *E. nindensis* would much rather redirect the carbohydrates elsewhere. There was a fair degree of fluctuation occurring with processes involved with membrane trafficking, suggesting to some level that during and after dehydration inter- and intra-cellular movement is tightly regulated. Protein phosphatase inhibitor activity, which was functionally enriched, showed a significant decrease in fold-change at the 40% and air-dry state in the NST and 25% in the ST, further corroborating the observation that *E. nindensis* actively modulates the phosphorylated status of proteins, presumably ensuring that essential proteins required for desiccation response remain phosphorylated. Processes assigned to *chromosome* showed an overall decrease in fold change across RWCs suggesting that at this scale without taking individual proteins into account, there is marked reduction with chromosomes implying to some degree that modifications done to chromosomes, such as histone associations, are favourably down-regulated. Overall energy metabolism also appeared to fluctuate throughout the RWCs implying that shifts to cellular energy production was evident. Since the

photosynthetic apparatus and associated energy production is shut down due to the breakdown of the thylakoids, energy needs to be supplied presumably either through increased mitochondrial oxidative phosphorylation or via β-oxidation of fatty acids in the peroxisomes.

By considering the individual highest and lowest fold changes associated with each RWC, summarised in Table 2-6, no clearly defined process appears to be favoured or shut down. A summary heatmap can be viewed in Figure 8-2 in the appendix. The 60% NST had the highest fold change associated with *Glutaredoxin 3* which was categorised as belonging to amino acid processes. Two associated annotations had decreased abundance at 25% ST and at 1 week post-rehydration whereas one accession was present in all. The highest abundance accession was present in all RWCs except 25% NST and 48hrs post-rehydration and was the same accession that had decreased fold change in the 1 week post-rehydration time point. The lowest fold change for *aquaporin PIP* in 60% NST was also diminished in 40% and airdry NST and then again at 1 week post-rehydration. The presence of the declined *PS II oxygen-evolving enhance protein 1* in the 40% NST further strengthens the argument that at 40% NST is when commitment to thylakoid disassembly and complete photosynthetic shut down occurs. Two accessions were assigned to this particular protein with 2 assigned to *protein 2*, and 3 to *protein 3*. Only one of the *protein 3* accessions had increased fold change at 1 week post-rehydration while the remainder all had negative fold changes for most dehydration RWCs except the 25% NST. This apparent differential abundance of proteins was largely also reflected in the transcriptome where related transcripts had differential expression at the same RWC (Madden, 2019). This could likely be due to the polyploidy of *E. nindensis* as seen for instance in the octoploid resurrection plant *C. plantagineum*, where divergent homologue expression was noted during desiccation stress (VanBuren, et al., 2023).

The *late embryogenesis abundant protein* (LEA) accounted for the greatest fold change in 40% and 25% NST. LEAs were present in all dehydration tissues with it accounting for greater than 30-fold increase in the NST having the highest fold change in airdry at >50 fold. One accession showed positive fold change in the rehydration time points and two accessions had negative fold changes across various RWCs. The *3-deamino-3-oxonicotianamine reductase* was categorised as belonging to processes involved in terpenoid and polyketide metabolism. Two accessions were assigned to this annotation and had marked reduction in all NST dehydration RWCs. One accession had marked increased activity in both NST and ST dehydration RWCs with one other accession having declined activity in the 40% NST. Its presence was completely absent in both rehydration time points. The *linoleate 9S-lipoxygenase* had two accessions assigned to it and was categorised as belonging to lipid metabolism. Its fold change was negative for all NST tissue except the 25% NST with one

additional accession annotated at 60% NST. A total of 13 accessions were assigned to *cell division protease FtsH* with the highest fold change present in the AD NST but was also positive in the 60% NST, 25% ST, and at 48hrs post-rehydration. The remaining accessions all had negative fold changes present in all dehydration RWCs except the 25% NST and it was categorised as belonging to processes involving protein folding, sorting, and degradation. The *cellular nucleic acid-binding protein* was also present in the 60% and AD NST and was categorised as being involved in genetic information processing with only 1 accession assigned to it. A total of 8 accessions were annotated as *pathogen-inducible salicylic acid glycosyltransferase*, categorised as carbohydrate metabolism, with the lowest fold change accession only present in the rehydration time point. One accession had positive fold change in the 60% and 40% NST while a negative fold change in the 1 week-rehydration for the same accession was noted. Five accessions had negative fold change at 48hrs post-rehydration.

The *malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)* had 8 accessions assigned to it with the one corresponding to the highest fold change present in all dehydration RWCs except the 40% NST. One accession had decreased fold change in all dehydration RWCs except the 25% NST, whereas 5 accessions had decreased fold change in the 1 week post-rehydration time point with one accession of the 5 having increased fold change in the 40% NST. The remaining accession had positive fold changes in all rehydration RWCs. The *protein FAM50* had two accessions assigned to it and was categorised as being involved in cellular processes, more specifically, the spliceosome and was only present in the two ST RWCs. This protein through blasting was further annotated as *Protein XAP5 CIRCADIAN TIMEKEEPER*. Lastly, the *long-chain acyl-CoA synthase* which is involved in lipid metabolism in the peroxisome had 3 accessions assigned to it and was diminished in the 60% ST and 1 week post-rehydration. Two accessions had both positive and negative fold changes in the 40% NST and 1 week post-rehydration respectively. One accession had positive fold changes in all dehydration RWCs except 25% NST whereas the remaining accession had decreased fold change in the 25% NST. The various differentially abundant proteins mentioned thus far can be seen in Table 8-3 in the Appendix, section 8.3.

Table 2-6: Highest and lowest fold changes per RWC as annotated by Koala Blast

	Highest		Lowest	
RWC	**Annotation**	**Fold change**	**Annotation**	**Fold change**
60 NST	Glutaredoxin 3	48.77	Aquaporin PIP	-38.23
40 NST	Late embryogenesis abundant protein	33.55	PS II Oxygen-evolving enhancer protein 1	-22.65
25 NST	Late embryogenesis abundant protein	46.05	3-deamino-3-oxonicotianamine reductase	-6.70
AD NST	Cell division protease FtsH	88.47	Linoleate 9S-lipoxygenase	-54.51
48hrs NST	Aspartyl-tRNA synthase	24.54	Linoleate 9S-lipoxygenase	-13.60
W1 NST	Cellular nucleic acid-binding protein	31.35	Pathogen-inducible salicylic acid glycosyltransferase	-32.36
60 ST	Malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	45.28	Protein XAP5 CIRCADIAN TIMEKEEPER	-10.55
25 ST	Malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	43.92	Long-chain acyl-CoA synthase	-50.97

2.3.2.1 Malate and pyruvate are key players in central carbohydrate metabolism.

Carbohydrate metabolism encompasses all processes involved not only in primary metabolism and energy generation such as glycolysis and the tricarboxylic acid cycle (TCA) or processes involved in the synthesis of amino- and nucleotide sugars, but also processes involved in the recycling and reshuffling of carbohydrates. As can be seen in Figure 2-10, the median fold change associated with Koala annotated DAPs for carbohydrates, a general trend towards lower fold changes was evident, with 66.26% of DAPs underrepresented. Three data points were removed from this plot in order to visualise the spread better. Those being a 45.2 fold change for 60% ST and a 43.9 fold change for 25% ST, annotated as a *malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)* enzyme, and a -32.4 fold change for the 1 week post-rehydration annotated as a *pathogen-inducible salicylic acid*

glycosyltransferase enzyme. The *malate dehydrogenase* enzyme was also present as the highest fold change for 60%, 25%, and air-dry NST at 14.44, 15.65, and 13.99 respectively (Figure 8-2 and Table 8-3). Highest fold change at 10.51 for the 40% NST was an enzyme annotated as *pyruvate decarboxylase. Fructose-bisphosphate aldolase class I* had the lowest fold change of -13.73 and -9.37 for the 60% and 40% NST whereas a *pyruvate orthophosphate* enzyme had the lowest fold change in the 25% and air-dry NST at -4.61 and -11.45 respectively. The senescent tissue did not have the same negative fold change proteins as did the non-senescent tissue. In the 60% ST, with a fold change of -9.33, the enzyme *fructose-1-6-bisphatase I* was the lowest whereas in the 25% ST the starch hydrolysing enzyme *β-amylase* was the lowest at the fold change of -15.97. Upon rehydration a clear shift is noted where for example in the 48hrs post-rehydration and 1 week post-rehydration the highest fold change of 17.89 and 11.05 was assigned to *protochlorophyllide reductase* respectively, with the lowest fold change at 48hrs post-rehydration assigned to *myo-inositol-1-phosphate synthase* at -5.49.

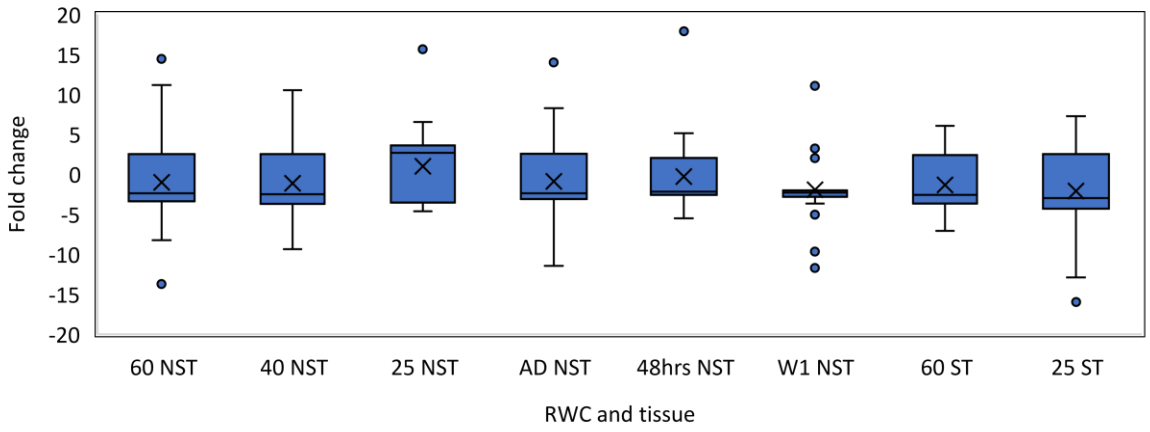


Figure 2-10: Carbohydrate metabolism in *E. nindensis*: Box and whisker plot showing the median fold change for DAPs annotated as belonging to processes involved in carbohydrate metabolism for all RWCs. Three outliers were removed. All DAPs are significant with a fold-change of >2 and FDR < 0.05. ST= senescent; AD= airdry; 48=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

When looking at these proteins across all RWCs, it is clear that although a protein might be the highest or lowest at a particular RWC it was present in other tissues and has the same pattern of expression. Clearly present in Figure 2-11 is a difference in protein and transcript expression between dehydration RWCs and rehydration time points, demonstrating that over and above highest or lowest protein abundance, there are clear differences between these two states.

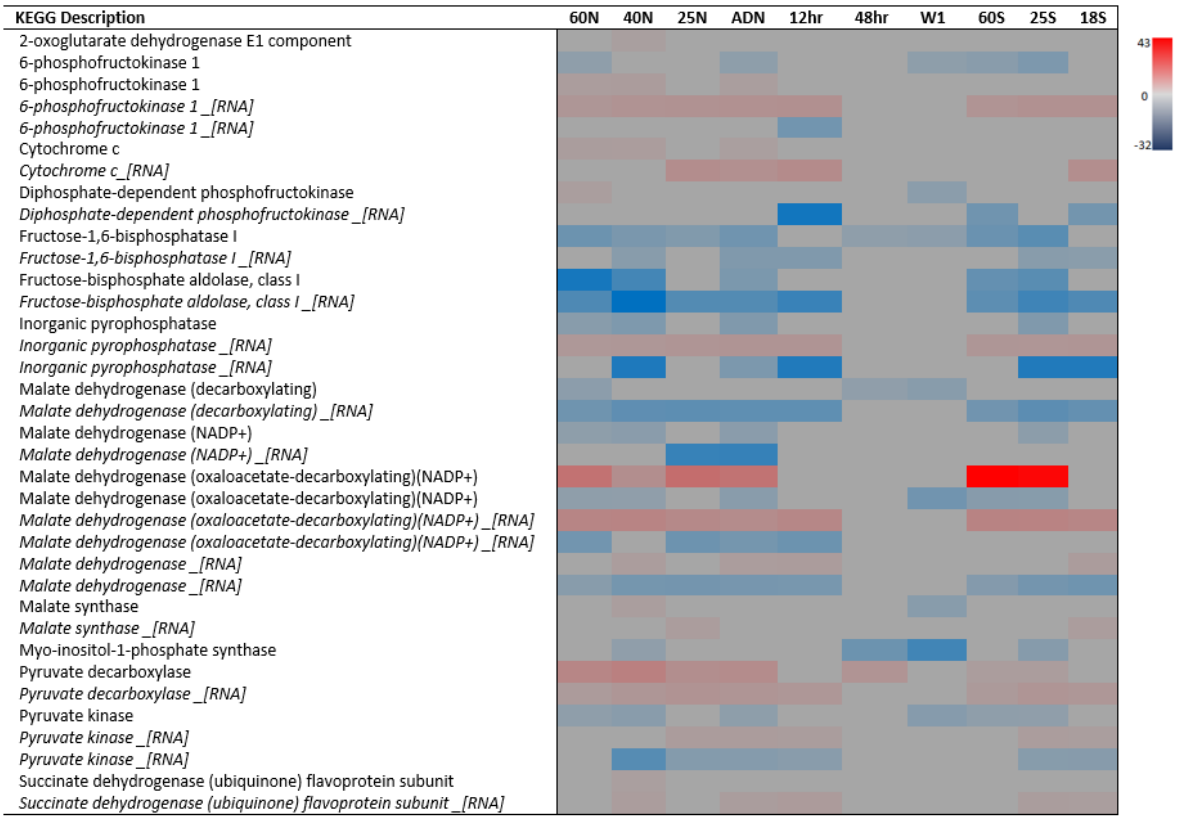


Figure 2-11: Fold change for selected carbohydrate related terms: Heatmap showing a selection of carbohydrate related terms from the proteome with matching transcripts from the transcription as log2 fold values. Transcripts indicated in italics and _[RNA]. All DEGs and DAPs represent log2 fold change >2, <0.05 FDR relative to fully hydrated leaves. N=non-senescent; S= senescent; AD= airdry; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

For instance, there was a total lack of *fructose-bisphosphate aldolase* in the 48hrs and 1 week post-rehydration and a switch in expression for *malate dehydrogenase* at 1 week post-rehydration. In order to appreciate the seemingly greatest and lowest abundances of these proteins one has to consider the protein within a greater scheme, especially since proteins, in particular enzymes, do not perform their functions in isolation. The direction of fructose-6-phosphate to fructose-1,6-phosphate is primarily driven by the *phosphofructokinase diphosphate dependent* enzyme at 60% NST with transcripts associated with it present throughout dehydration in both tissues. Its relative abundance was diminished at 1 week post-rehydration which implies that during the early stages of dehydration the formation of fructose-1,6-phosphate was produced at no ATP expense. The reverse reaction was highly underrepresented and was likely due to the reduced Calvin cycle occurring but could also be connected to a suspected active repression via a calcium binding protein simply annotated as MO25 which interacts with AMPK. The cascade of events eventually leads to the repression of fructose-1,6-phosphate illustrating that tight regulation of the phosphorylated state of fructose is important. *Triosephosphate isomerase* was underrepresented at 60% and 40%

NST and in both ST (Table 8-3) which begs the question of how glyceraldehyde-3-phosphate could be formed since fructose-2,6-phosphate formation is favoured.

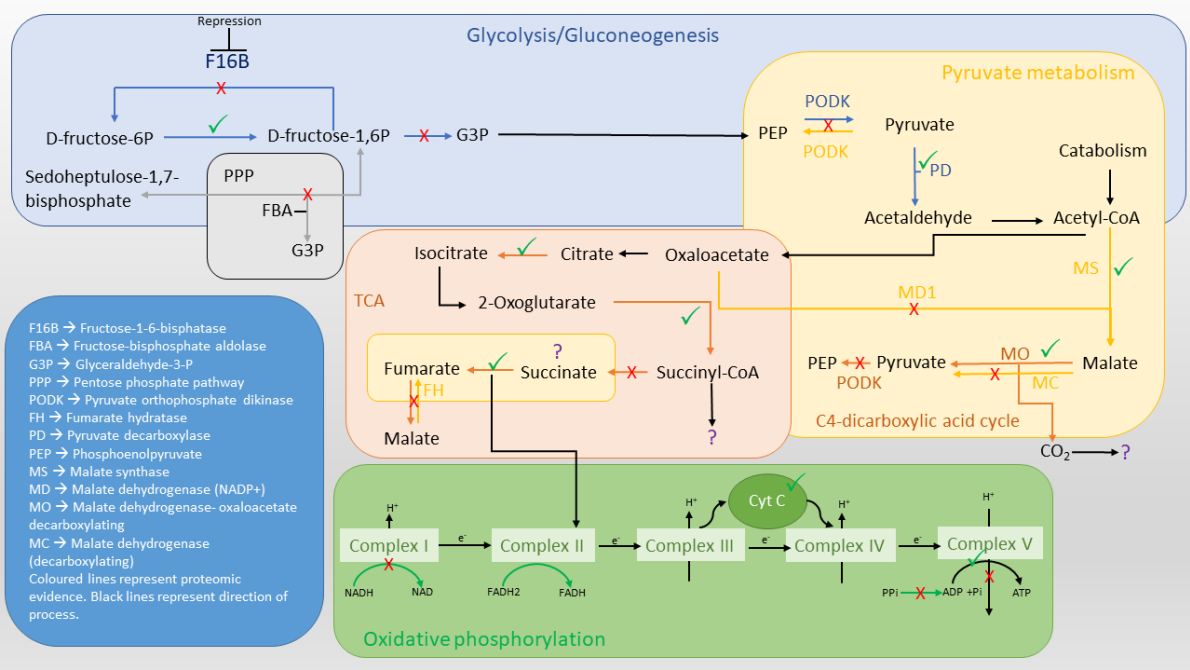


Figure 2-12: Summary of selected carbohydrate enzymes: Schematic showing proteins that have a connection to the over- and underrepresented proteins from Figure 2-10. Red crosses represent an overall decline in abundance whereas green ticks represent overall increase in abundance. Coloured lines matching to coloured areas represent proteins from the proteome and Figure 2-10. Black lines indicate general direction and are not present in the proteome.

The functional enrichment of *pyruvate decarboxylase* in all RWCs seems to imply that pyruvate is being funnelled from elsewhere and presumably in the direction of acetyl-CoA formation (further corroborated through increased activity of β-oxidation of fatty acids which is discussed below and network analysis of malate-related proteins (Figure 8-3)) for energy production. Pyruvate was not being formed from phosphoenolpyruvate (PEP) since *pyruvate kinase* protein abundance was underrepresented in all tissues despite transcripts being present at lower RWCs in both tissues and 12hrs post-rehydration as well as the great underrepresentation of *pyruvate orthophosphate dikinase* in all RWCs. Pyruvate was also not being produced via *malate dehydrogenase (decarboxylating)* (MD1) which was only underrepresented at 60% NST and then again at both rehydration time points. MD1 is similar to the Arabidopsis homologue At4G00570 which is annotated as an NAD-ME2 (NAD-dependent malic enzyme). KEGG pathway reconstruction places this enzyme in both carbon fixation and pyruvate metabolism. Given its underrepresentation at 60% NST, it is most likely involved in carbon fixation and not necessarily in pyruvate metabolism. *Malate dehydrogenase NADP+* (MD) is similar to the Arabidopsis homologue At5G58330 annotated as an NADP-MDH (NADP-dependent malate dehydrogenase) where it too through KEGG reconstruction is assigned to carbon fixation and pyruvate metabolism where its

underrepresentation is probably also related to photosynthetic shut down. The great overrepresentation of *malate dehydrogenase (oxaloacetate decarboxylating)* (MO) in all RWCs seems to suggest that it is playing an important role. Given that *E. nindensis* is a C4 plant it is not surprising that enzymes involved in C4 photosynthetic processes should appear, however Rubisco was greatly underrepresented (Table 8-3) which implies that its activity was likely not in favour of transferring sequestered CO_2 to Rubisco. MO functions in both photosynthetic and non-photosynthetic oxidative decarboxylation of malate (Liu, et al., 2007) where it uses NADP. Its role in non-photosynthetic processes have been reported to be involved in plant defence, fatty acid biosynthesis, and cytosolic pH maintenance (Liu, et al., 2007) and is similar to the Arabidopsis homologue At5G25880 and At2G199000 (NADP-ME3 and NADP-ME1 respectively). In Arabidopsis its expression is greatest during seed maturation and has been shown to be greatly induced under ABA, NaCl, and mannitol treatments where knockout mutants were less sensitive to ABA repression (Arias, et al., 2018), whereas in rice, an NADP-ME2 was induced under abiotic stress (Liu, et al., 2007). In *Ricinus communis* MO plays an important role in oil deposition during seed development (Wheeler, et al., 2016) and lipid bodies were noted in ultrastructural studies of *E. nindensis* (Madden, 2019). Ferredoxin, during photosynthetic electron transport, serves as one of the primary means to regenerate NADPH. NADPH is a vital cofactor for a wide range of enzymes and processes but most noteworthy is its role with the ascorbate-glutathione and thioredoxin antioxidant systems (Corpas & Barroso, 2014). Its great overrepresentation could thus not only be in the generation of pyruvate for TCA, which according to network interactions on String indicated close interaction with a number of TCA intermediates (Figure 8-3), but also in replenishing NADPH required by antioxidant enzymes which are discussed below.

Whether it is involved in fatty acid biosynthesis is contentious owing to evidence that would suggest fatty acid oxidation is favoured over biosynthesis (discussed below). Malate supply was replenished at 40% RWC through *malate synthase* using acetyl-CoA through catabolism of amino acids and fatty acids. Overexpression of malate synthase from *R. communis* enhanced the tolerance of Arabidopsis seeds under elevated temperatures and NaCl treatment (Brito, et al., 2020). The authors, through coupled metabolomics, postulated that it modulates the glyoxylate cycle and gluconeogenesis owing to the enriched metabolites involved in these processes (Brito, et al., 2020). Whether *malate synthase* is doing something similar remains unclear, but it is reasonable to speculate that when coupled with lower fold-change of MO at 40% RWC, it serves as a means to replenish the malate pool to ensure that during desiccation response NADPH is regenerated.

Diverting attention towards the TCA, citrate was presumably converted to isocitrate, and the cycle proceeded in the forward direction to produce succinyl-CoA from 2-oxoglutarate

exclusively in the 40% NST. The conversion of succinyl-CoA to succinate was surprisingly underrepresented in all dehydration RWCs despite there being an overrepresentation of succinate dehydrogenase at 40% NST with an increase in expression of associated transcript at 12hrs post-rehydration. This suggests that although there is underrepresentation of succinyl-CoA conversion, succinate is being converted to fumarate but the conversion of it to malate with incorporation of H_2O is underrepresented in the 1 week post-rehydration and in the 60% ST. This suggests that at 40% NST there is a large shift in the production of fumarate (likely as a result of complex II activity with the recycling of FADH) to malate which is then only switched off upon rehydration and surprisingly in the 60% ST. This further implicates the favourable production of malate in the NST during dehydration. Energy production is naturally hampered under abiotic stress where for instance energy is not being produced via photosynthesis owing to its shutdown, yet there were a number of proteins and transcripts associated with oxidative phosphorylation which seems to suggest active electron transport and oxidative phosphorylation even at the airdry state. There was significant up-regulation of transcripts associated with *V-type ATpase* with concomitant overrepresentation of associated proteins in all RWCs. Proton transporting ATpase was also overrepresented in all RWC with electron transfer proteins showing increased protein abundance at 60% NST, 40% NST, and at the airdry state with corresponding transcripts showing increased expression at 40% NST and 12hrs post-rehydration with associated overrepresentation of *cytochrome c* in 60% NST, 40% NST, and air-dry state. While there were some processes that showed decreased activity, the overwhelming favourable direction of electron transfer seems to imply mitochondrial activity however, the data would imply that it was not necessarily in the direction of ATP production since *inorganic phosphatase* protein abundance was underrepresented in all tissues except at 25% in both tissues and at rehydration time points despite transcripts being up-regulated in all tissues. What might be occurring here warrants further investigation as to the state of mitochondrial activity and ATP turnover during and after rehydration to fully understand the regulation of mitochondrial-derived energy production. It is however safe to assume that given the functional enrichment of *protein phosphatase inhibitor activity* the need to generate more ATP for the sake of protein phosphorylation is likely not that high. In addition, many of the favourable processes listed thus far are in favour of using ADP or no ATP at all. *E. nindensis* might be diverting ATP pools towards other processes and could likely make small amounts of ATP should it be absolutely necessary.

The changes to central carbohydrate metabolism illustrate a few unique features of *E. nindensis* during dehydration. First, glycolytic processes that favour ATP consumption were underrepresented in favour of ADP-dependent processes. Malate enzymes showed differential abundance with underrepresentation of enzymes associated with the shutdown of

photosynthesis (MD and MD1 in Figure 2-12) and overrepresentation of enzymes (MO in Figure 2-12) for the recycling of NADPH and pyruvate production to sustain TCA (Figure 8-3). Acetyl-CoA may play a larger role than anticipated and is likely being derived from amino and fatty acid catabolism. Lastly, the first evidence of selective translation is presented where *E. nindensis* would have transcripts but fails to translate them, presumably to divert energetically costly translation towards the synthesis of more important proteins such as chaperones.

2.3.2.2 *E. nindensis* is a raffinose oligosaccharide resurrection plant.

The complexity that governs desiccation response cannot be overstated, as evident from central carbohydrate metabolism. A common strategy employed by many resurrection plants is the accumulation of sugars (Zhang, et al., 2016). Between hydrated and dry leaves, *E. nindensis* accumulated an array of sugars including sucrose, glucose, fructose, raffinose, stachyose, and a small amount of trehalose relative to the other sugars (Ghasempour, et al., 1998). Processes involving starch synthesis and breakdown appear to be regulated at both stages. *Starch synthase* was overrepresented in the proteome in the air-dry state, at 60% and 25% ST, and then 48hrs after rehydration. The 60% ST and rehydration time point can easily be rationalised as residual photosynthesis and starch deposition and resumption of photosynthesis and starch deposition respectively but the presence of active starch synthesis in the airdry-state and 25% ST is puzzling. Transcriptomic evidence also indicated that there were transcripts at 40% NST, air-dry and 12hrs post-rehydration which would indicate that starch must be actively made (Madden, 2019). Starch acts as a carbohydrate reservoir for plants which is typically sent to sink organs such as roots via phloem transport as smaller oligosaccharides. The occurrence of starch synthesis in such water-limiting conditions could only imply that residual sugars which were not redirected towards osmoprotectant functions could rapidly be re-polymerised and most likely stored. Further investigation into the actual starch content of air-dry tissue is needed to determine whether *E. nindensis* actively polymerises sugars or prepares the tissue for rehydration by having starch synthesis proteins. Another case could be made in favour of the insoluble nature of starch and how it could be utilised once osmoprotectant carbohydrates have been partitioned by perhaps providing some structural integrity to the cell under near and complete water-deficit conditions. Ultrastructural studies of *E. nindensis* during desiccation indicated a presence of starch at 25% RWC in both tissues, corroborating to some extent that starch was actively being made (Madden, 2019). Indeed, the role of starch during desiccation is not well-explored and its degradation and synthesis appears to not follow a universal pattern (Thalmann & Santelia, 2017).

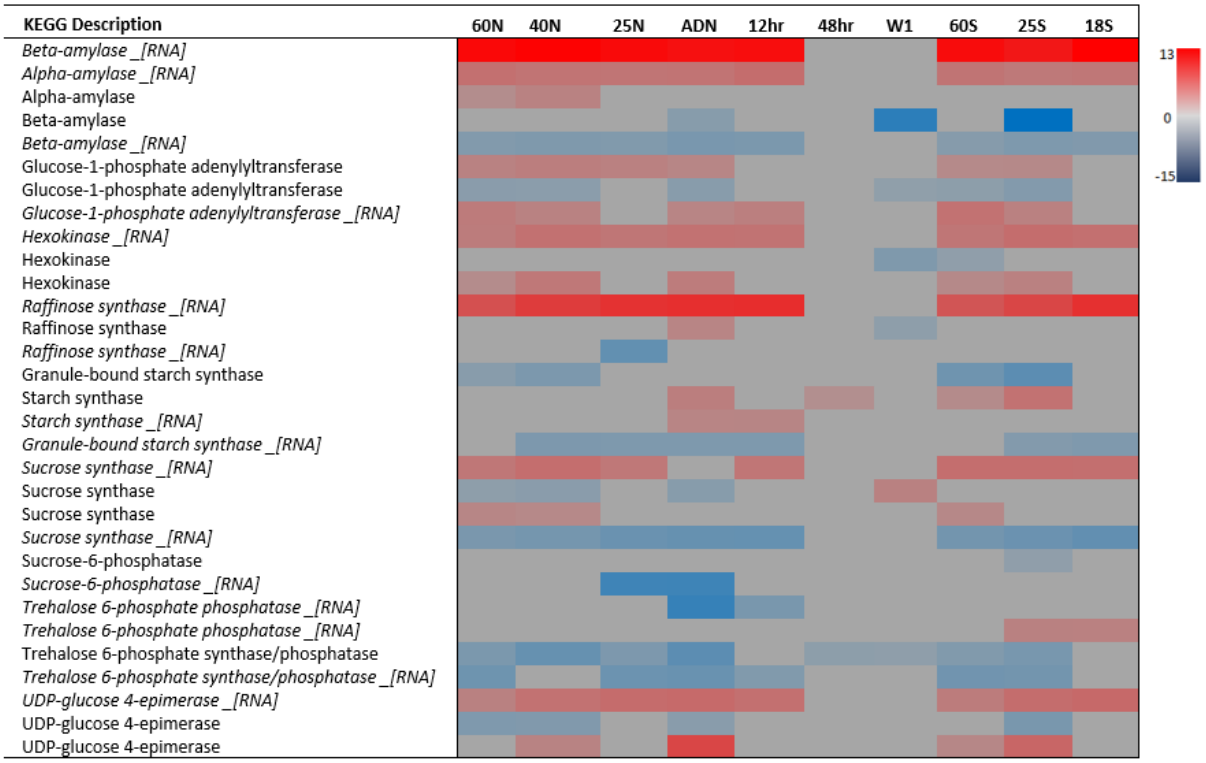


Figure 2-13: Fold change for selected starch and sucrose terms: Heatmap showing a selection of starch and sucrose metabolism terms from the proteome with matching transcripts from the transcriptome as log2 fold values. Transcripts indicated in italics and _[RNA]. All DEGs and DAPs represent log2 fold change >2, <0.05 FDR relative to fully hydrated leaves. N=non-senescent; S= senescent; AD= airdry; 12hr=12hrs post-rehydration; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

Interesting is the underrepresentation of *granule bound starch synthase* which was likely also the same transcript that showed down-regulation. Both enzymes are involved in final starch production but the substrates they use differ. Starch synthase can use UDP-glucose or ADP-glucose as substrates, whereas granule-bound starch synthase can only use ADP-glucose. The enzyme responsible for ADP-glucose production *glucose-1-phosphate adenyltransferase* was present as seemingly two different isoforms with differential abundance, with one predominantly overrepresented during dehydration and absent upon rehydration. The overarching situation is that *starch synthase* activity is favoured over *granule bound starch synthase* and that the ADP-glucose derived through enzymatic activity is being diverted to the former and not the latter. It was not just the synthesis of starch which displayed preferential enzyme choice. The breakdown of starch was favoured in the 60% and 40% NST by *alpha-amylase* which hydrolyses starch to form dextrin which can readily be converted into D-glucose. Its transcriptomic profile indicated that at least one isoform of it was upregulated in both tissues during dehydration but was only actively translated in the 60% and 40% NST. *Beta-amylase*, which primarily hydrolyses starch to maltose, was noticeably underrepresented in the air-dry, 25% ST, and 1 week post-rehydration. This would suggest that during early and moderate desiccation stress the need to generate glucose via dextrin is favoured over the

maltose route. Synthesis and catabolism of starch thus appears to be temporal and spatial in nature, except for transcriptomic evidence, which illustrated that production and breakdown of starch metabolism was tightly regulated. Hexoses, carbohydrates whose composition is that of a 6-member ring, are one of the primary sources of sugars during dehydration as seen in other desiccation tolerant plants such as *Sporobolus stapfianus* and *Xerophyta viscosa* (Whittaker, et al., 2001) with a corresponding increase in *hexokinase* activity. *E. nindensis* had overrepresentation of *hexokinase* activity at both omics' levels indicating that there was active supply of hexoses during dehydration.

Sucrose accumulation during vegetative desiccation tolerance is well-accepted nowadays where it was postulated nearly twenty years ago to be a replacement for water and in concert with proteins forms a 'glassy' state as seen in many orthodox seeds during maturation (Illing, et al., 2005). *Sucrose synthase*, which from the proteome annotation was simply annotated as such, had differential regulation between what can be concluded as two different isoforms. This was also reflected in the transcriptome. The choice of which isoform presents a unique situation and illustrates a fundamental flaw in annotation algorithms as well as the nomenclature employed by researchers who have annotated the genomes and proteomes of the reference models. This in turn is further complicated by the polyploidy nature of *E. nindensis*. Without knowing which isoform was active and which was not, it is difficult to make any conclusions. The transcriptome identified the upregulated *sucrose synthase* as SUS3 and the downregulated one as SUS4 and the regulation of the two was attributed to favouring pathways that are assistive to desiccation tolerance such as seed maturation. SUS3 cleaves sucrose to form its constitutive monomers UDP-glucose and fructose and has been shown to be an important enzyme during stressed conditions owing to it not requiring ATP (Barrat, et al., 2009). In fact, it has been suggested that owing to its energy efficiency it can be considered as a biomarker of sink strength (Xu, et al., 2012). The overabundance of *sucrose synthase* in the 60% ST could be indicative of source to sink transport between the senescent and non-senescent tissue, which according to Blomstedt, et al., (2010) plays an important role in the desiccation response of *S. stapfianus.* The transcriptome identified a collection of sugar transporters as being up-regulated but the proteome, through both InterPro and Koala annotation, could not find any overrepresentation of terms related to sugar transport. There were accessions annotated as transporters, but they all appear to be related to ion transport. It has been speculated and reported that different SUS isoforms play different roles during stressed conditions (Bieniawska, et al., 2007) and as such their relative accumulation would differ and also provides substantiation for the overrepresentation of the *sucrose synthase* at 1 week post-rehydration, which is underrepresented during dehydration. In light of the mirroring transcriptomic evidence of SUS3 and SUS4 between the two tissue types it might be a general

grass behaviour irrespective of senescent classification as suggested by Pardo, et al., (2020), however, it stands to reason that the over- and underrepresentation of the two isoforms in the NST is not a generalisation of grass behaviour but more programmed desiccation tolerance.

Our understanding of the importance of sucrose accumulation has broadened to the understanding of a preference of some species to accumulate raffinose (and related oligosaccharides such as stachyose) or trehalose (Dace, et al., 2023). Members of the Linderniaceae family appear to accumulate neither of these but prefer to accumulate octulose (Zhang, et al., 2016). Raffinose is synthesised via galactose metabolism where galactinol moieties, which are derived through the linking of myo-inositol to galactose, are added onto sucrose. Trehalose on the other hand is synthesised by combining UDP-glucose and glucose-6-phosphate by *trehalose-6-phosphate synthase*. In their study on hydrated versus dry leaves of several grasses, Ghasempour et al., (1998) demonstrated that accumulation of raffinose increased from 0.00 ±0.0 in the hydrated state to 3.00 ±0.04 µmole.g^{-1} dry weight in the dry state, whereas trehalose only increased by 0.20 µmole.g^{-1} dry weight. Coupled with proteomic evidence that showed a significant reduction in *trehalose-6-phosphate synthase* activity with significant up-regulation at the transcript level of *raffinose synthase* illustrates definitively that *E. nindensis* is a raffinose accumulating resurrection plant instead of a trehalose dependant one. Though interestingly *raffinose synthase* only appeared as differentially abundant in the air-dry state despite transcripts being present as early as 40% in the NST. What is even more conclusive is that it was significantly underrepresented in the 1 week-post-rehydration tissue, demonstrating that firstly there was a delay in translation of significantly up-regulated transcripts to significantly abundant protein but that upon resumption of ‘normal’ metabolic activity it was highly repressed.

The tight regulation of starch synthesis and breakdown coupled with the increase in hexokinase activity and sucrose synthase demonstrated that *E. nindensis* tightly controls the stages when monomeric carbohydrates were made available and what their need was. The underrepresentation of *trehalose synthase* activity coupled with high *raffinose synthase* activity demonstrated that during dehydration the need for sucrose accumulation far exceeds the needs for trehalose and that upon reaching the air-dry state the accumulation of raffinose exclusively in the NST safeguards the NST and could likely be one of the major factors that differentiates the NST from ST.

‘

2.3.2.3 Hormonal regulation implies defence response and programmed senescence.

Irrespective of the form of stress encountered by plants, there is generally significant overlap between processes which are switched on and off between biotic and abiotic stress. For most of the researchers exploring the effects and adaptations of plants to abiotic stress, biotic stress responses are often overlooked. For example, in the nickel hyperaccumulator, *Senecio coronatus*, a possible explanation for its ability to hyperaccumulate a heavy metal which would otherwise pose significant growth and development retardation to sensitive plants such as Arabidopsis is to prevent herbivory and pathogenicity (Boyd & Martens, 1998; van der Pas & Ingle, 2019). As evident thus far, the desiccation response requires a well-orchestrated set of processes which for the most part can be costly to maintain. For instance, the redirection of carbohydrates to create this 'glassy' state naturally enriches the tissue with carbohydrate sources which herbivores, insects, and pathogens would enjoy harvesting. In order to protect the valuable resources accumulated, as well as prevent the untimely removal of non-senescent leaves, *E. nindensis* arms itself accordingly. As noted in Table 2-6 and Figure 2-14, there was a significant reduction in a *pathogen-inducible salicylic acid glycosyltransferase* enzyme at 1 week post-rehydration with increased activity of it at the 60% and 40% NST. Salicylic acid (SA) is a key plant hormone involved not only in development, where it plays an important role in leaf senescence for example (Vlot, et al., 2009), but also in plant defence (Rivas-San & Plasencia, 2011). SA in cells is predominantly in either a methylated or glycosylated state, where the conversion of the methylated state (MeSA) to SA is responsible for the induction of systemic acquired resistance. The glycosylation of SA, mediated by *salicylic acid glycosyltransferase* (SAGT), is apparently not required for systemic acquired resistance (Kobayashi, et al., 2020). It has been reported that application of SA or pathogenic attack can greatly increase the expression of SAGT in Arabidopsis (Song, 2006) and that its abundance is closely related to the regulation of SA and an SA-derivative SA 2-O-β-D-glucoside (SAG). Song, (2006) suggests that the activiation of SAGT is an early pathogenic response likley required for adaptation. Given that the plants used for the transcriptomic and proteomic study were grown in controlled conditions it is highly unlikely that a pathogen or accidental application of SA occurred which would warrant the accumulation of SAGT. An alternative possibility is that it might result in an early activation of pathogen-response priming while the NST undergoes dehydration.

The overrepresentation of *jasmonoyl-L-amino acid hydrolase* (JIH), produced through alpha-linoleic acid degradation in the peroxisome (Reumann & Bartel, 2016), was present in the NST exclusively during dehydration. Plant hormones play a huge role in response to environmental stresses, for example in Arabidopsis, the accumulation of jasmonic acid can reduce water loss

by regulating the stomatal aperture (Wang, et al., 2020). ABA is a known inducer of stomatal closure upon water deficit stress. JIH has been shown in particular to regulate the levels of jasmonoyl-L-isoleucine levels in *Nicotiana attenuata* where the various carboxy- and hydroxylated states have an impact on overall defence against herbivory (Woldemariam, et al., 2012). The transcriptome indicated an upregulation of an *ethylene insensitive 3* transcript. Ethylene is a volatile plant hormone most famous for its effects on fruit ripening and senescence among other things (Dolgikh, et al., 2019). The presence of *gibberellin regulated protein* in the transcriptome points to some hormonal activation through gibberellin which like the aforementioned has a role in a myriad of plant developmental processes (He, et al., 2020). However, as has been made clear thus far, the presence of a transcript, irrespective of a great fold change does not necessarily correlate to active protein. It is likely that during dehydration these hormone-related transcripts are made but it would appear that JA and SAGT are of more relevance in the NST than ethylene or gibberellic acid. Two equally valid arguments can be made for the SAGT and JIH abundance at the mild and moderate drought stress time points. One is in priming a defence response due to environmental signal (water deficit) which could in turn prime the plant against potential herbivory or pathogenicity. An equally valid but most likely more applicable response is in the triggering of senescence since both are directly involved in regulating this process. The most convincing argument for defence response is the overrepresentation of proteins annotated as *chitinase* and *endochitinase* predominantly during dehydration in both tissues.

Chitin is an amino-polysaccharide polymer commonly found in the exoskeletons of insects and within fungal cell walls. Chitinase binds chitin and randomly cleaves glycosidic linkages and has in recent years garnered great interest for its role in plant defence and biopesticide biotechnology (Sharma, et al., 2011). The transcriptome also reflected the upregulation of a number of chitinase genes. The abundance of the protein, mirrored by upregulation in transcripts, implies that *E. nindensis* could prime itself for potential fungal pathogenicity and/or insect herbivore. This could possibly be as a means to deter biotic stress during dehydration and rehydration since both processes are metabolically expensive and directing resources towards abscission (mediated by abscisic acid) could not only affect hormonal balance during dehydration but could also deplete limited stores of resources which are required elsewhere. Though it should also be noted that during the first round of de novo assembly, fungal genes were noted and had to be removed (Madden, 2019). Thus the appearance of chitinase could also be due to latent fungi present on the leaf tissue.

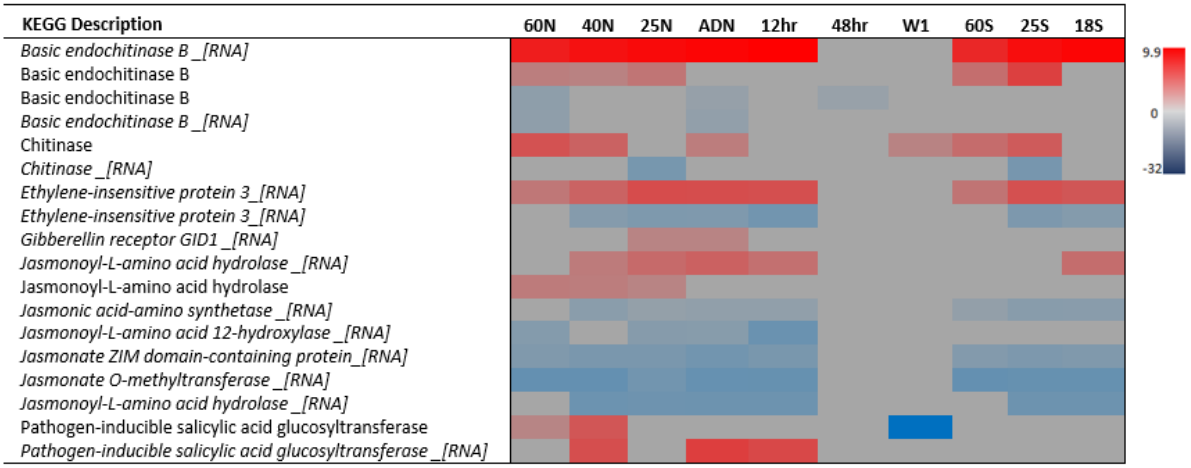


Figure 2-14: Fold change for selected defence-related terms: Heatmap showing a selection of defence-related terms from the proteome with matching transcripts from the transcriptome as log2 fold values. Transcripts indicated in italics and _[RNA]. All DEGs and DAPs represent log2 fold change >2, <0.05 FDR relative to fully hydrated leaves. N=non-senescent; S= senescent; AD= airdry; 12hr=12hrs post-rehydration; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

While not a standard approach when considering abiotic stress, the interconnectivity between biotic and abiotic stress cannot be ignored. The overabundance of chitinase suggested that should *E. nindensis* be attacked by insects or fungi it has the necessary enzyme to cleave the glycosidic linkages and could even potentially use this as a means of carbohydrate acquisition. The presence of these hormones could be in favour of either priming *E. nindensis* for potential pathogenicity or could more likely be involved in programming leaf senescence. The apparent absence of it in the ST demonstrates that although at first glance and as indicated by the transcriptome there are not any obvious differences between the two tissue type, this illustrates that there was overrepresentation of hormonal regulation occurring in the NST not present in the ST.

2.3.2.4 Amino acid metabolism is in favour of oxidative and osmotic stress mitigation.

Across RWCs there was a general trend towards underrepresentation of proteins involved with overall amino acid metabolism (Figure 2-15) with 66.09% of DAPs being underrepresented. Highest fold change was seen at 48hrs post-rehydration with a 24.54 fold change in *aspartyl-tRNA synthase* with the lowest fold change at 25% ST at -17.86 for *prolyl-tRNA synthase. Aspartyl-tRNA synthase* also had the highest fold change at 1 week post-rehydration. *Peroxiredoxin 6* with an average fold change of 12.13 had the highest fold change for both tissues at 60% and 25% whereas *glutamyl-tRNA synthase* had the lowest fold change for the 60% and 40% NST, and 60% ST with an average fold change of -10.11.

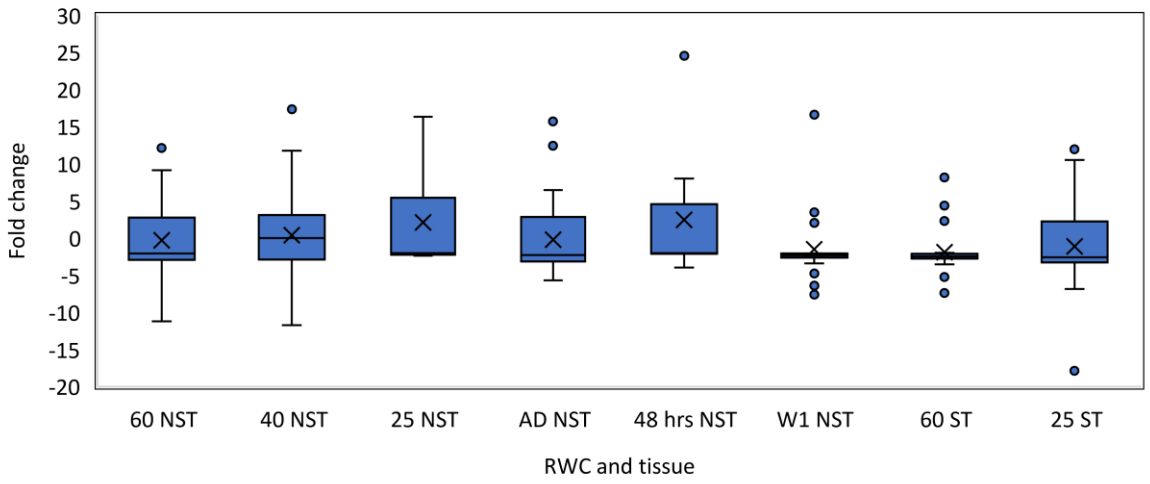


Figure 2-15: Amino acid metabolism in *E. nindensis*: Box and whisker plot showing the median fold change for DAPs annotated as belonging to processes involved in amino acid metabolism for all RWCs. NST=non-senescent; ST= senescent; AD= airdry; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

Thioredoxin reductase NADPH had differential expression in the 40% NST and 25%NST with fold changes of 17.35 and -2.40 respectively. *Arginine decarboxylase* with fold changes of -5.70 and -7.61 was underrepresented in the arid-dry and 1 week post-rehydration respectively. The air-dry tissue had the highest fold change for *L-ascorbate peroxidase* at 15.69 whereas at 48hrs post-rehydration with a fold change of -3.96 *polyphenol oxidase* was the lowest.

Across most of the dehydration RWCs, a large number of aminoacyl-tRNA synthase enzymes were significantly underrepresented (Figure 2-16). Most notably was a great underrepresentation in the airdry state. Of these, only two aminoacyl-tRNA synthase were overrepresented, the aforementioned *aspartyl-tRNA synthase*, and a *phenylalanyl-tRNA synthase* (annotated here as an alpha subunit) was overrepresented in the 60% NST. The occurrence of a significant underrepresentation of these aminoacyl-tRNA synthase enzymes implies that very little amino acid to tRNA ligation is occurring, which has far-reaching impacts on overall translation since they are required by the ribosomal units during peptide elongation. At first glance this appears to suggest that synthesis of new proteins was severely halted yet there was proteomic evidence of enrichment of certain proteins during dehydration.

All of the aminoacyl-tRNA synthase enzymes require ATP for the linking of an amino acid to a corresponding tRNA and would suggest that *E. nindensis* favours not making new amino-acyl-tRNAs as a means to redirect ATP towards sugar and protein phosphorylation. It would be erroneous to state that there were no pools of aminoacyl-tRNAs within the cell to sustain protein synthesis because it was evident that there was new protein turnover, even at the airdry state. The enrichment of *phenylalanyl-tRNA synthase* at the 60% NST implies that when considered with the decreased abundance of *phenylalanine ammonia lyase* from gene ontology enrichment, free phenylalanine is ligated to tRNA. Why this might be occurring is

puzzling. *Phenylalanine ammonia lyase* converts phenylalanine to trans-cinnamate which can then either be directed towards acetyl-CoA or fumarate production both of which leading into the TCA. Data would suggest some part of it was ligated to tRNA while another more reasonable path could be in the direction of succinyl-CoA and acetyl-CoA production via alternative pathways. When taken together by considering overall phenylalanine, tyrosine, and tryptophan metabolism as a whole, it appears that the main shutdown is centred around the shikimate pathway and apparent arrest of anthranilate production.

Glutamate decarboxylate which catalyses the conversion of L-glutamate to γ-aminobutanoate (GABA) was significantly overrepresented in most of the dehydration RWCs and at 48hrs post-rehydration. The presence of GABA is well-known for its role in stress tolerance (Fait, et al., 2008). Results from a comparison of a drought-tolerant wheat cultivar *Triticum aestivum* L. cv. Zagros and drought-sensitive cultivar cv. Marvdasht indicated that the drought-tolerant cultivar accumulated more GABA and proline (Saeedipour & Moradi, 2012). In *C. plantagineum*, GABA showed increased abundance upon dehydration (Xu, et al., 2021). In *S. stapfianus,* which displays a similar phenotype to *E. nindensis* an increase in GABA was noted in the older desiccation-sensitive leaves and the authors concluded that the accumulation of GABA was not 'essential' for desiccation tolerance (Martinelli, et al., 2007). Similarly in their transcriptomic and metabolomics study the authors too suggested that it did not play that great of a role in desiccation tolerance (Yobi, et al., 2017). This may well be the case for desiccation tolerance as there are a number of other processes such as the accumulation of sucrose and the overexpression of heat shock proteins that are more likely responsible for the desiccation phenotype. However, the overarching role *glutamate decarboxylase* plays in oxidative stress response cannot be ignored.

The overrepresentation of *proline iminopeptidase* involved in terminal proline peptide cleavage suggests that at 60% NST *E. nindensis* increases its proline pools. Proline has been associated with alleviating drought stress as it too acts as an osmoprotectant (Yoshiba, et al., 1997), although it was speculated several decades ago that proline is not necessarily important for vegetative desiccation tolerance (Tymms & Gaff, 1979). Akin to glutamate decarboxylase, the accumulation of proline may simply be a general drought response commonly also encountered by sensitive plants. However, the lack of *proline iminopeptidase* in the ST would suggest that although it might be universal it does not necessarily translate into a fitness gain in all tissue types.

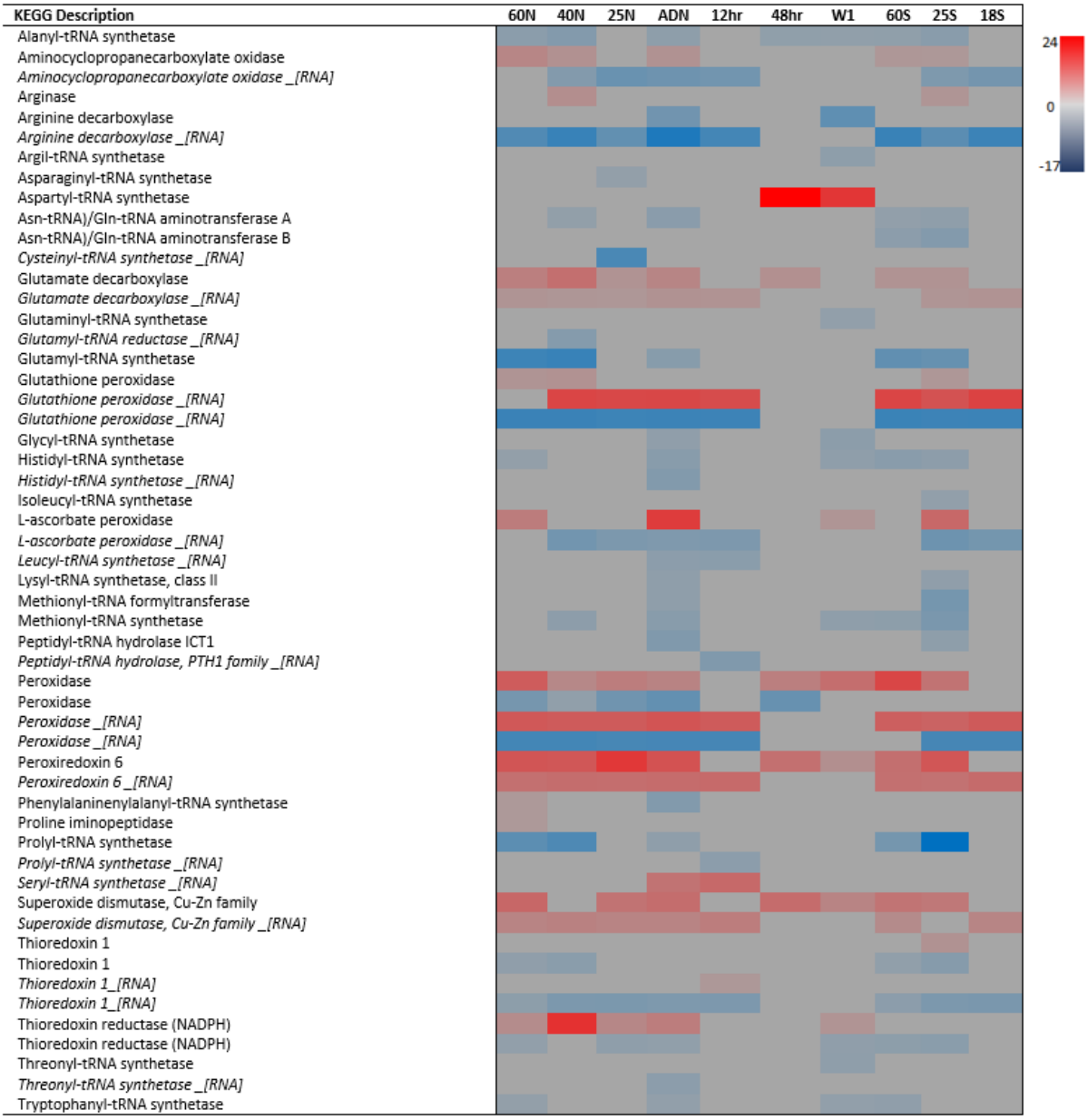


Figure 2-16: Fold change for selected amino acid metabolic terms: Heatmap showing a selection of amino acid metabolic terms from the proteome with matching transcripts from the transcriptome as log2 fold values. Transcripts indicated in italics and _[RNA]. All DEGs and DAPs represent log2 fold change >2, <0.05 FDR relative to fully hydrated leaves. N=non-senescent; S= senescent; AD= airdry; 12hr=12hrs post-rehydration; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

Antioxidant response is a common strategy used by plants under unfavourable conditions. In *E. nindensis* at the proteomic level and involved in general amino acid metabolism, *L-ascorbate peroxidase* and *glutathione peroxidase* displayed increased abundance in some but not all RWCs, which hints to shuttling between antioxidant response when one considers the overabundance of *superoxide dismutase* (SOD) which was significantly overrepresented in all RWCs except at 40% NST. The aforementioned are principally involved in the removal of reactive oxygen species and shows once more that akin to *glutamate decarboxylase* and proline accumulation *E. nindensis* makes use of common strategies to mitigate oxidative damage.

Thioredoxin reductase is yet another enzyme involved in overall REDOX status by catalysing the thiol-disulphide interchange of target proteins which include a host of proteins involved in carbon assimilation such as carbonic anhydrase, central metabolism such as glucose-6-phosphate dehydrogenase, a range of factors involved in translation and protein degradation, as well as enzymes such as glutathione peroxidase (Gelhaye, et al., 2005). Given the broad range of target proteins, they can be found in numerous locations within the cell but a role for them could be in the mitochondrion where a poplar thioredoxin reductase isoform was shown to activate alternative oxidase (AOX) (Gelhaye, et al., 2004). AOX play an important role in the reduction of ROS (Purvis, 1997) where the reduction of oxygen is coupled with oxidation of ubiquinol which does not lead to increased ATP production but the generation of ROS via mitochondrial activity is slowed. AOX has several other reported functions, for example in regulating PCD (Robson & Vanlerberghe, 2002) but interestingly, loss of function of complex I in *Nicotiana sylvestris* lead to leaves becoming more tolerant to stress through AOX redirection (Dutilleul, et al., 2003). While there were no apparent increases in AOX, given the complexity of *thioredoxin reductase* regulation and the aforementioned reduction of activity at Complex I (Figure 2-12) with associated decline in *inorganic pyrophosphatase* activity (Figure 2-11) it is plausible to speculate that some *thioredoxin reductase* activity could be in favour of AOX activation during dehydration in the NST. Lastly, the significant overabundance of *peroxiredoxin 6* in all tissues irrespective of RWC implies that this protein was switched on and constitutively expressed. Akin to many of the aforementioned antioxidant proteins, *peroxiredoxin* is involved in thiol-based reduction of H_2O_2 (Dietz, 2011) which is one of the more common by-products of general metabolic activity (Niu & Liao, 2016). Apart from its accepted role in a number of plant developmental processes such as delayed leaf senescence when administered as an additive in cut lily (Liao, et al., 2012) and general signalling too much of it can become problematic. SOD is responsible for the production of H_2O_2 from oxide radicles produced via metabolism whereas peroxiredoxin is responsible for removal of it. Thus, it is reasonable to speculate that regulation of H_2O_2 content is very important since *E.*

nindensis like other plants would use H_2O_2 as a signalling molecule to initiate developmental processes but also needs to maintain a safe balance of it such that it does not lead to lipid peroxidation which would otherwise compromise the integrity of the plasma membrane.

The overrepresentation of *aminocyclopropane carboxylate oxidase*, which is involved in the production of ethylene, at most dehydration RWCs suggests once more hormonal regulation. Ethylene has a multitude of roles including development and senescence where it exerts a role in the breakdown of chlorophyll (Iqbal, et al., 2017). Through the catabolism of proteins arginine is processed to produced ornithine and urea by *arginase,* present at 40% NST and 25% ST. Ornithine can be further processed to produce compounds such as glutamate which can be decarboxylated as indicated above. Ornithine itself is an interesting compound and has been shown to play a role in abiotic stress tolerance in Arabidopsis through overexpression of the Arabidopsis homologue (Shi, et al., 2013). Its activity interestingly is greatly enhanced by the presence of methyl jasmonate. It is tempting to speculate that the accumulation of urea in the ST is indicative of the ST being a source site since nitrogen can be remobilised from older leaves to younger leaves in the form of urea (Sultana, et al., 2021).

The overarching influence of amino acid metabolism is to produce compounds that have a positive effect on abiotic stress tolerance while simultaneously limiting the expenditure of ATP during aminoacyl-tRNA synthesis. The final breakdown products from an array of degradation pathways seem to suggest that amino acids are a primary source for acetyl-CoA production. *E. nindensis* maintains a fine balance of H_2O_2 by ensuring that positive effects of it are attained but that detrimental side effects are avoided. The amino acid metabolism of *E. nindensis* illustrates a well-orchestrated process which might not be directly involved in vegetative desiccation tolerance per se but is involved in ensuring oxidative and osmotic stress is mitigated.

2.3.2.5 Lipid metabolism is directed towards degradation and protection.

Lipid metabolism indicated an approximately equal representation of DAPs as significantly accumulating (61 DAPs) and diminishing (68 DAPs). The median fold change across RWCs further demonstrated that generally there was a trend towards decreased abundance at early and moderate desiccation with increased abundances at the late and airdry states for the NST whereas ST showed opposite trends (Figure 2-17). Two data points were removed from Figure 2-17 below to visualise overall data spread better. Those were a -54.51 fold change in the airdry state annotated as *linoleate 9S-lipoxygenase* and a -50.97 fold change in the 25% ST annotated as a *long-chain acyl-CoA synthetase* which was also the least abundant protein in the 60% ST.

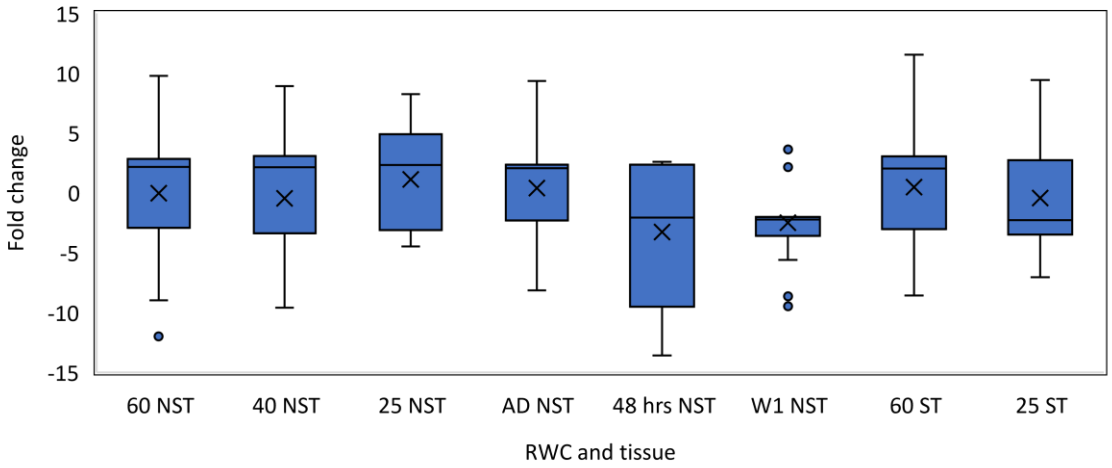


Figure 2-17: Lipid metabolism in *E. nindensis*: Box and whisker plot showing the median fold change for DAPs annotated as belonging to processes involved in lipid metabolism for all RWCs. Two datum points were removed from 25% ST and AD NST for better visualisation. NST=non-senescent; ST= senescent; AD= airdry; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC

Linoleate 9S-lipoxygenase was also the lowest in 48hrs post-rehydration with a fold change of -13.60. An annotated *peroxygenase* had the greatest fold change of 11.54 in the 60% ST but was in fact the highest in all RWCs except for the 1 week post-rehydration which had the greatest fold change of 4.00 annotated as an *adrenodoxin-NADP+ reductase* and lowest fold change annotated as a *12-oxophytodienoci acid* at -9.47. A *plant 3-β-hydroxysteroid-4-α-carboxylate-3-dehydrogenase* had the lowest fold change of -12.00 and -9.59 for the 60% and 40% NST whereas the lowest fold change of -4.49 in the 25% NST was annotated as a *lysophosphatidic acid acyltransferase.*

By considering the selected proteins in Figure 2-18, the overarching lipid metabolic profile was that of favouring lipid breakdown, despite the high accumulation of acetyl-CoA from carbohydrate metabolism which alluded to earlier was most likely required for the formation of malate. For instance, *enoyl-CoA hydratase* which catalyses the second step in β-oxidation of fatty acids was overrepresented in the 40% and airdry NST, and at 25% ST, yet was underrepresented in the 60% ST. Transcriptomic evidence indicated that at least two different genes were upregulated and appeared at very low RWCs and persisted even up to 12hrs post-rehydration. What makes this particularly intriguing is the fact that *enoyl-CoA hydratase* incorporates molecular water during the hydrolysis of fatty acids, and it does this at very low water contents, implying that favouring β-oxidation is very important during desiccation. Although an overrepresentation of one enzyme involved in β-oxidation of fatty acids does not necessarily indicate favouring the process, a thiolase-like domain was annotated through InterPro annotation and had overrepresentation in the airdry and 1 week

post-rehydration. Koala annotation annotated these accessions as being a *chalcone synthase* which is ultimately involved in the biosynthesis of secondary metabolites, more specifically flavonoid biosynthesis. As such, it would be erroneous to suggest that full oxidation was occurring but at least to some extent there was evidence that it was favoured. *Fatty acid α-dioxygenase* was only present in the 40% NST and is involved in α-oxidation of fatty acids. The main differences between the two forms of fatty acid oxidation are that β-oxidation occurs in the mitochondrion and releases acetyl-CoA whereas α-oxidation occurs in the peroxisome and releases CO2 at the end of the cycle, though in plants β-oxidation also occurs in the peroxisome. Alpha-oxidation is also different from β-oxidation in that the first step catalysed by α-dioxygenase incorporates molecular oxygen at the alpha carbon of the fatty acid and many of these are characterised as having peroxidase activity which leads to the production of hydroperoxy fatty acids (Buchaupt, et al., 2014).

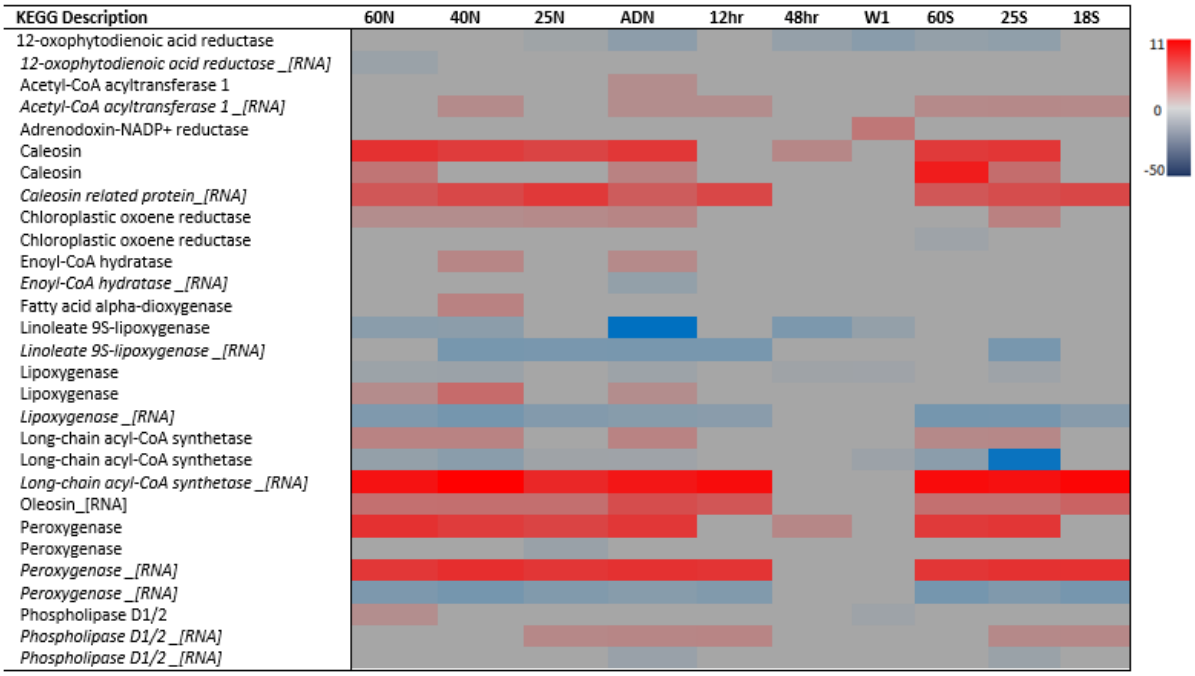


Figure 2-18: Fold change for selected lipid metabolic terms: Heatmap showing a selection of lipid metabolic terms from the proteome with matching transcripts from the transcriptome as log2 fold values. Caleosin was not annotated by Koala but is from InterPro annotation. Transcripts indicated in italics and _[RNA]. All DEGs and DAPs represent log2 fold change >2, <0.05 FDR relative to fully hydrated leaves. N=non-senescent; S= senescent; AD= airdry; 12hr=12hrs post-rehydration; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

As suggested above, the 40% and airdry NST can be seen as tipping points and the presence of the first enzyme involved in α-oxidation of fatty acids at 40% NST suggests that at moderate drought stress the need in the peroxisomes to produce CO_2 is apparently important. Why this might be occurring is counterintuitive since carbon fixation and photosynthesis is all but shut down by this stage which suggests that the enzyme's presence could likely be in favour of

liberating carbon units which can be remobilised elsewhere and leave the hydroperoxy fatty acid to be further processed by peroxygenase. Further evidence of favouring fatty acid breakdown was through overrepresentation of *acyl-CoA oxidase* which is involved in the β-oxidation of fatty acids in the peroxisomes where the substrate specificity ranges from short to very long fatty acids. Beta-oxidation of fatty acids is predominant in plant peroxisomes where the outcome of oxidation not only influences the TAG conversion in seeds for example but can also have impacts on leaf senescence and as mentioned earlier jasmonate production (Arent, et al., 2008). The underrepresentation of it at 1 week post-rehydration and dominant presence during most of the dehydration NST is indicative of it having an important role in the NST where the end-products through both forms of oxidation in the peroxisome exert not only possible hormonal effects but could serve as a means to recycle carbon.

Amino acid degradation pathways involving primarily leucine and isoleucine, can lead to the accumulation of acetoacetyl-CoA and acetyl-CoA where *acetyl-CoA acetyltransferase* catalyses the formation of an acyl-CoA which is longer by two carbon atoms. This was overrepresented in the airdry state and 25% ST suggesting that before the ST reaches the same state it employed the use of this enzyme to recycle degradation products, but perhaps it occured too soon. Its seemingly underrepresentation at the 1 week post-rehydration further implies that the recycling of acetyl-CoA in the airdry state is important and one could speculate that the acetyl-CoA generated through this process could be in favour of producing protective molecules.

Some of the more obscure DAPs identified from Koala annotation as being involved in lipid metabolism included the overrepresentation of *adrenodoxin NADP+ reductase* in the 1 week post-rehydration exclusively. Although annotated as an adrenodoxin involved in the biosynthesis of steroid hormones, it is in actuality a ferredoxin-NADP+ reductase involved in photosynthesis. *Phospholipase D* falls into a broad category of phospholipases whose substrates include a range of phosphatidyl lipids such as phosphatidyl choline and phosphatidylglycerol for instance where the product of catalysis could be phosphatidic acid and choline in the case of acting upon phosphatidyl choline. Over and above its biochemical role, *phospholipase D* can localise to various compartments such as the plasma membrane and have been implicated through the production of phosphatidic acid to have roles in numerous plant developmental and stress-induced processes (Takáč, et al., 2019). Plant phospholipase Ds are further grouped into six subclasses of which the alpha-1 is involved in ABA-mediated seed germination regulation (Choudhury & Pandey, 2016) but has also been implicated in vesicular trafficking (Takáč, et al., 2019). Its overrepresentation at the 60% NST exclusively and underrepresentation at 1 week post-rehydration suggests that at mild water-deficit stress, *phospholipase D* might initiate signalling transduction via production of

phosphatidic acid which ultimately leads to the activation of appropriate responses beyond 60% RWC. The RNA evidence suggested that that there were at least two different genes with differential expression and that it was expressed from 25% RWCs all the way up to 12hrs post-rehydration (Madden, 2019). However, the lack of translated product further illustrates that transcript presence does not correlate to active protein and that phospholipase D activity was only required at 60% RWC.

Sterols form an integral part of membranes where for instance their interaction with other lipids form part of signalling processes such as auxin transport and pathogen-associated molecular patterns and brassinosteroid production in plants (Kim, et al., 2012). Sterol synthesis starts with cycloartenol, a cyclised 4-ringed squalene product, production and in order to be fully functional two methyl groups are removed from it and is accomplished by *plant 3-β-hydroxysteroid-4-α-carboxylate-3-dehydrogenase.* Double and single knockout mutants in Arabidopsis displayed no characteristically detrimental phenotypes but the overexpressing line had growth defects which the authors attributed to alterations of membrane sterol composition which altered auxin response as opposed to alterations in brassinosteroid synthesis (Kim, et al., 2012). The underrepresentation of this enzyme exclusively in the 60% and 40% NST implies that auxin-mediated signalling in the membranes is not activated.

A rational conclusion can be made in terms of increased jasmonate production and how it may either play a role in general plant defence or in leaf senescence development. The enzyme *12-oxophytodeinoc acid reductase,* which had underrepresentation in all tissues except the 60% and 40% NST is involved in the synthesis of jasmonic acid through reduction of unsaturated double bonds from a precursor molecule 12-oxophytodienoic acid and 4,5 dihydro-jasmonic acid (Maynard, et al., 2020). The precursors are produced through lipoxygenase activity and are then further processed in the peroxisome and chloroplasts and are considered to be reactive electrophilic species (RES), which have been implicated in the activation of '*cell survival genes*' (Farmer & Davoine, 2007) but can become hazardous owing to their high reactivity status (Maynard, et al., 2020). The REDOX balance of glutathione and ascorbate for instance are highly regulated by antioxidant enzymes such as glutathione reductase and ascorbate peroxidase which can act as Michael donors in interacting with RES (Michael acceptors) for their removal (Maynard, et al., 2020). The underrepresentation points towards it not having a role in ROS or RES homeostasis as proposed by Maynard, et al., (2020) in favour of more traditional ROS and RES mechanisms involving classical antioxidants. *De novo* biosynthesis of glycerophospholipids was underrepresented in the NST tissue at 40% to airdry through decreased accumulation of *lysophosphatidic acid acyltransferase,* which uses phosphatidic acid as a substrate, despite a lipidomic study on *E. nindensis* identifying accumulation of glycerophospholipids (Madden, 2019). This is likely a

result of phospholipase D activity only being enhanced in the 60% NST where phosphatidic acid is required for signal transduction instead of glycerophospholipid biosynthesis.

Although the overarching profile for lipid metabolism is towards catabolism, there were at least two accessions that showed increased abundance in *long-chain acyl-CoA synthase* (LACS) with one accession showing the lowest abundance at 25% ST. Their role is in the ATP-dependent conversion of long-chain fatty acids into corresponding thioesters which are involved in a number of lipid metabolic processes and their localisation, substrate specificity, and expression in different organs are quite varied (Zhao, et al., 2021). For example, in Arabidopsis, *AtLACS1* to *3* and *6* are expressed in epidermal cells in the leaf whereas *AtLACS2,3* and *9* are expressed in the roots where cutin and suberin are synthesised respectively (Zhao, et al., 2021). Given that they can localise to different organelles it is reasonable to speculate that the underrepresented DAPs are likely chloroplastic in origin whereas the overrepresented DAPs are peroxisomal or ER in origin where for instance ER-derived long-chain fatty acids can be directed towards cuticular lipids and TAGs whereas peroxisome-derived long-chain fatty acids can be directed towards oxidation. Their role further extends to response to environmental stress where knockout mutants have higher susceptibility to drought stress (Zhao, et al., 2021) whereas ectopic expression of an apple LACS1 not only enhanced Arabidopsis but also apple calli tolerance to osmotic-stress and ABA treatment (Zhang, et al., 2018) which the authors attributed to enhanced wax biosynthesis as a means to decrease leaf water loss. As such, the synthesis of these long-chain fatty acids serves several roles. One is possibly in the direction of enhanced cuticular wax deposition for which there is no evidence in *E. nindensis*, the other is to provide substrates for peroxisomal oxidation, and the other is to supply TAGs which can be incorporated into lipid bodies.

Owing to annotation differences between the proteome and transcriptome, no oleosins were identified through either annotation however the transcriptome indicated that there were no apparent differences between ST and NST tissue which is contradictory to another study that looked at the transcript abundance between the desiccation tolerant *L. brevidens* and desiccation sensitive relative *L. subracemosa* (VanBuren, et al., 2018). The observed difference in oleosin abundance when probed on a western blot (Madden, 2019) demonstrated that its appearance in *E. nindensis* was presumed to be likely regulated at the translational level rather than the transcriptional level however, when considering protein evidence of caleosin, a related oleosin, this does not appear to be the case (Figure 2-18). Caleosins are ubiquitous in plant cells, algae, and some fungi (Patridge & Murphy, 2009) and are characterised as calcium binding oil-body surface proteins (Poxleitner, et al., 2006) and like oleosins have widely been accepted to play a role in structural maintenance of oil bodies

(Chen, et al., 1999). In seeds, TAGs which are produced during the maturation of the seed and are stored in oil bodies are catabolised upon germination as a primary source of energy production via β-oxidation until such time that photosynthesis is operational (Poxleitner, et al., 2006). However, it is not strictly partitioned as germination energy supply where for instance during seed maturation in *Brassica napus* at least 10% of the stored TAGs are catabolised during embryo development (Chia, et al., 2005). In Arabidopsis, an ABA-inducible non-seed caleosin isoform was identified (Takahashi, et al., 2000) from drought treatment of Arabidopsis implying firstly that caleosin synthesis and deposition is not a seed-development-only process and secondly that there is great plasticity for their role in plants. In a study using mutant Arabidopsis plants, the authors concluded that caleosin plays an important role in the degradation of stored lipids during seed germination (Poxleitner, et al., 2006) whereas Patridge & Murphy, (2009) suggest that the peroxygenase activity of caleosin is probably involved in oxylipin signalling such as those regulated by ABA. With this in mind, it becomes apparent that distinguishing caleosin from peroxygenase activity is not an arbitrary process. Peroxygenase is part of the peroxygenase pathway where lipoxygenase catalyses the oxygenation of unsaturated fatty acids to produce hydroperoxides. Peroxygenase then transforms these into oxylipins which have been shown to have an impact on plant defence (Meesapyodsuk & Qiu, 2011). In most of the dehydrating NST there was differential accumulation of an annotated *lipoxygenase* suggesting that akin to the sucrose synthase situation, certain isoforms were active whereas other were not. Assuming activity of lipoxygenase it explains activity of peroxygenase, however gene ontology indicated an overall decline in oxylipin biosynthesis (Figure 2-6). The accessions and Koala annotation corresponding to the gene ontology enrichment of oxylipin biosynthesis is *linoleate-9S-lipoxygenase* which had the highest depletion in the airdry state but was also significantly underrepresented in both rehydration time points. The products of this catalysis have been implicated in defence response to fungal pathogenicity (An, et al., 2019) but the evidence presented here seems to suggest that this is not favoured. When casting the net a bit more wider in order to understand the activity of peroxygenase it becomes more evident that peroxygenase is the '*Swiss Army Knife'* for oxyfunctionalisation (Hobisch, et al., 2021) whose promiscuous substrate specificity, like that of the founding member from the fungus *Agrocybe aegerita* of the sub-subclass of oxidoreductase enzymes broadly called unspecific peroxygenase, is quite varied but uses H_2O_2 as both electron acceptor and oxygen donor (Hobisch, et al., 2021). Unlike the cytochrome P450 monooxygenases, peroxygenase is co-factor independent making them thermodynamically more favourable. Thus, it is likely that a two-component process might be occurring; one in which a lipoxygenase isoform delivers hydroperoxide to peroxygenase and one where peroxygenase uses H_2O_2 as substrate directly. It is also plausible then to speculate that caleosin might perform both peroxygenase activity

as well as structural lipid body maintenance. What the end product of peroxygenase activity might be is unknown and warrants further investigation of *E. nindensis* peroxygenases but sufficed to say, it may well be involved in modulating H_2O_2 akin to aforementioned antioxidant enzymes which seems to suggest that *E. nindensis* has significant control over its H_2O_2 pool, ensuring once more that the least amount required for positive signal transduction is there but also limiting unfavourable H_2O_2-induced peroxidation.

Overall lipid metabolism illustrates a complex set of processes with significant crosstalk between carbohydrate and amino acid metabolism. The general trend observed herein was that *E. nindensis* favours oxidation of fatty acids in the peroxisome as a means to increase acetyl-CoA which can either be fed into the TCA or be directed towards malate production. In addition, it seems that lipid-based signalling was important for the early activation of desiccation response in the NST which was absent in the ST further strengthening the argument that translational control is what ultimately differentiates the NST from the ST. The predominant overrepresentation of peroxygenase which in concert with antioxidant enzymes illustrates that *E. nindensis* exerts significant control over H_2O_2. Production of localised long-chain fatty acids could be in favour of enhancing cuticular wax deposition as a physical measure to prevent leaf water loss during desiccation. However, this warrants further exploration in *E. nindensis.*

2.3.2.6 Genetic information processes imply universal stress.

Through Koala annotation, *Genetic information processing* had 53.90%, *chaperones* 70.92%, and *protein processing* which includes folding, sorting, and degradation accounting for 54.28% of DAPs assigned as overrepresented. By considering the greatest and lowest fold changes across the RWCs, summarised in Table 2-7 a range of unique and overlapping proteins illustrate for instance that the appearance of *late embryogenesis abundant* (LEA) proteins is important not only during dehydration but that it is a strategy shared between ST and NST. Note though that LEAs were not annotated through Koala blast but rather through InterPro and original CPGR *E. curvula* annotation. All other proteins presented were from Koala annotation.

The presence of a *cellular nucleic acid-binding protein* (CNBP) as overrepresented in the 60% NST and 1 week post-rehydration indicates that there is active interaction with RNA occurring at these two stages. CNBP is also known as a zinc-finger protein 9 and is a general nucleic acid binding protein with broad substrate (sequence) specificity and in mammalian systems has been implicated in the regulation of immune response under viral infection conditions (Chen, et al., 2021). In plants, their roles are just as varied as in the mammalian system where

they have been associated with a wide range of transcriptional regulation of genes associated with stress response and stress granule formation (Yan, et al., 2022). Reviewed by Yan, et al., (2022), the authors illustrate that activation of what they referred to as RNA binding proteins (which encompasses CNBP and zinc-finger proteins) can be done via ABA-independent and -dependent processes which ultimately leads to nucleic acid processing regulation which can ultimately lead to the formation of stress granules. The overrepresentation at the 60% NST suggests that at the early stages of drought stress transcriptional activation of regulatory genes is initiated and processed to cope with moderate to severe drought stress. The protein also has some increased abundance in the airdry state where one could speculate its principal role is in the facilitation of stress granule formation. The complete lack of it in the ST suggests that only the NST initiates this response and given that it is also ABA-inducible and the evidence that gene ontology enrichment indicated that only the NST had proteins assigned to that GO term (albeit only in the 40% and airdry NST), it is reasonable to speculate that ABA is involved in activating this process in the NST exclusively.

Stress granules (whose relevance is considered elsewhere in this thesis) are composed of a collection of proteins such as chaperones and RNA processing proteins, metabolites, and RNA and their appearance is transient in nature (Kedersha, et al., 2013; Li, et al., 2022c). Chaperone proteins have been heavily implicated not only in their proteostasis role but also in the assembly and protection of stress granules and membraneless compartments. One such group of chaperones are LEA proteins. LEAs are present during the late embryogenesis of seed maturation but their appearance in transcriptomic and proteomic studies on resurrection plants has anchored them as being key players during desiccation stress (Artur, et al., 2019a).

In fact, their importance has been so well accepted that active research within our research group on a range of Xerophyta LEAs is focused on characterising these uniquely thermotolerant and intrinsically disordered proteins on the basis of structure, protective mechanisms, and their ability to undergo liquid-liquid phase separation. Given this intrinsic nature, a number of seed-specific LEA proteins have been shown to undergo liquid-liquid phase separation (Ginsawaeng, et al., 2021) and it is suggested that this process allows LEA proteins to sequester important biomolecules to facilitate either their storage for later mobilisation or to create an environment that enables biomolecular activity such as translation. The presence of LEAs throughout all tissues irrespective of RWC clearly suggests that their role in providing thermo-protection under water limiting conditions is of great importance, where for example the greatest fold change was seen at the airdry state at a 59.82 fold change.

Table 2-7: Summary of highest and lowest fold changes of DAPs assigned to Genetic Information Processing, Chaperone, and Protein Processing for all RWCs as annotated by Koala blast and InterPro

Category	Genetic Information Processing		Chaperone		Protein Processing	
RWC	**Highest**	**Lowest**	**Highest**	**Lowest**	**Highest**	**Lowest**
60 NST	Cellular nucleic acid-binding protein	SCY1-like protein 1	LEA	PAT complex subunit CCDC 47	Serine carboxypeptidase-like clade II	COP9 signalosome complex subunit 3
40 NST	ATP-Binding cassette, subfamily F, member 3			Molecular chaperone DnaK		
25 NST		La-related protein 7		None		Serine carboxypeptidase
AD NST		Ribonuclease T2		PAT complex subunit CCDC 47	Cell division protease FtsH	COP9 signalosome complex subunit 3
48hrs NST		Nucleolar pre-ribosomal associated protein 1				Protein deglycase
W1 NST	Cellular nucleic acid-binding protein	Ribonuclease T2			Phytepsin	
60 ST	Heat shock transcription factor, other eukaryote	SCY1-like protein 1		Trigger factor	Calnexin	COP9 signalosome complex subunit 3
25 ST	YTH-domain-containing family protein	Transcription factor SPEECHLESS		HSP20	Cell division protease FtsH	

In addition, owing to their relatively small size, compared for example with HSP90, the turnover of LEAs likely far exceeds turnover over larger chaperones. The presence of an annotated LEA in the rehydration time points does not invalidate its importance in desiccation tolerance response since the abundance of it is several folds lower than seen in dehydration RWCs but further strengthens the argument for their overall cellular importance. Both the ST and NST accumulated LEAs, though arguably the NST had slightly higher enrichment of them, which illustrates that this is a universal stress response and not necessarily exclusive to ensuring desiccation tolerance. Though it is not only LEAs that impart thermo-tolerance, but heat shock proteins have also been a staple protein/transcript observed in a host of vegetative desiccation tolerant plants where their expression and relative abundance is significantly high, and their importance is considered elsewhere in this thesis. Nevertheless, the occurrence of an HSP70.4, HSP20, and HSP90 in the proteome and transcriptome illustrate these are key players in maintaining proteostasis and at least in the case of HSP70 can be involved in the transient assembly of stress granules.

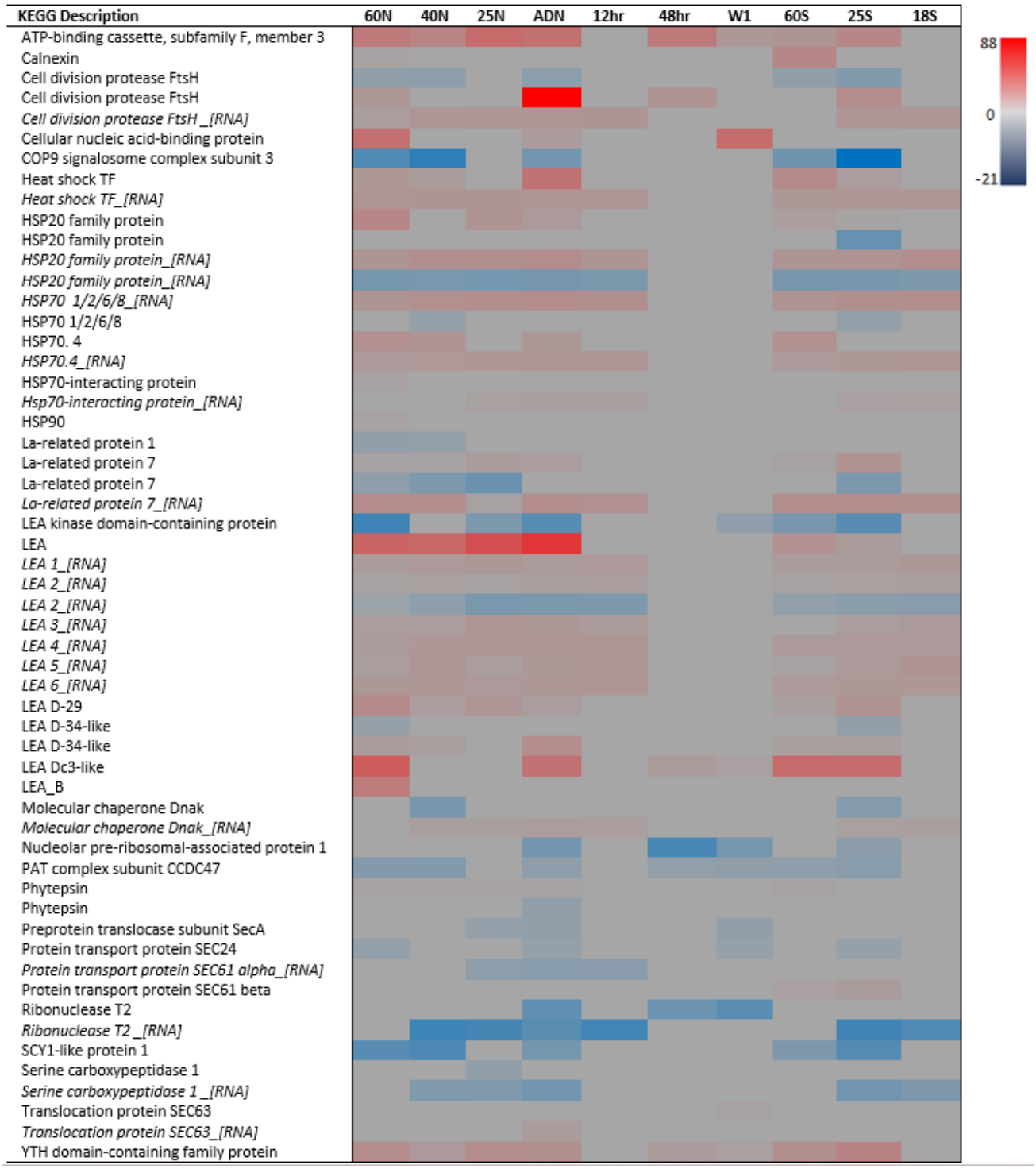


Figure 2-19: Fold change for selected Genetic Information Processing, Chaperone, and Protein Processing terms: Heatmap showing a selection of Genetic Information Processing, Chaperone, and Protein Processing terms from the proteome with matching transcripts from the transcriptome as log2 fold values. LEAs annotated from InterPro annotation. Transcripts indicated in italics and _[RNA]. All DEGs and DAPs represent log2 fold change >2, <0.05 FDR relative to fully hydrated leaves. N=non-senescent; S= senescent; AD= airdry; 12hr=12hrs post-rehydration; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

An annotated *HSP transcription factor* was observed in dehydrating tissues with highest overrepresentation at the airdry state. Although there were other transcription factors which were annotated through Koala blast such as a *homeobox-leucine zipper* (overrepresented in all tissue except 25% NST and ST), a PHD-finger protein ALFIN-like and *upstream activation factor of subunit UAF30* (underrepresented in 60% NST only), a putative transcription factor (overrepresented in 60% and 25% NST and 1 week post-rehydration), and a SPEECHLESS*,* the presence of this particular HSP-related transcription factor seems to suggest that there was greater transcriptional activity occurring in the transcribing of HSP-related genes. Differential expression of RNA transcripts coupled with a underrepresentation of an HSP20 suggests that like many proteins and transcripts encountered thus far, there is preference in favouring the production of particular isoforms over others, implying that there is a high degree of translational regulation occurring. *Calnexin*, exclusive in the 60% ST, is a chaperone that is involved in the ER that has a role in retaining unfolded or unassembled N-glycosylated proteins, releasing peptides only once they have been properly assembled and folded. Its abundance in the 60% ST could be an early sign that the ST undergoes greater misfolding during the early stages of drought.

With regards to the *SPEECHLESS* transcription factor, its primary role is in stomatal development (Chen, et al., 2020) and was significantly underrepresented in most tissues during dehydration tissues with lowest abundance in the 25% ST. As reviewed in Chen, et al., (2020) the cascade of events that regulate its expression is extensive and illustrates a vital role in stomatal development in response to cues. For instance, auxin causes repression of the cascade event whereas osmotic stress can induce the cascade that leads to repression of SPEECHLESS. Repression of it ultimately leads to the repression of protoderm to stomata lineage development which when taken into consideration appears to be a common strategy employed by both tissues as a means to perturb the development of stomata.

The overrepresentation of the *YTH-domain containing protein* is quite interesting as it hints to mRNA processing. Post-transcriptional processing of mRNA involves modifications such as methylation on the N6 of adenosine (m^6A) and accounts for the most common mRNA post-transcriptional modification in eukaryotes (Patil, et al., 2018). The role of mRNA methylation is speculated to be primarily determined by the 'reader' proteins containing a YT521-B homology (YTH) domain which processes it. Two distinct classes of YTH-domain proteins, denoted as YTHDF and YTHDC have been established (Patil, et al., 2018). The former possesses an intrinsically disordered N-terminus with the YTH-domain at the C-terminus whereas the latter based on amino acid sequence is not related to the bigger YTH-domain family (Patil, et al., 2018). Proposed roles for the YTH-domain are primarily centred around the fact that some recognise and bind m^6A whereas others simply do not where they function

somewhat redundantly in mRNA instability through transient interactions with P-bodies or activation of translation through recruitment of eIF3 complex (Patil, et al., 2018; Hernández, et al., 2018). YTHDC by contrast has been associated with m^6A recognition during exon inclusion during splicing and has been shown to recruit a *serine/arginine-rich splicing factor* (SAF) to promote this activity (Roundtree & He, 2016). It is unclear to which sub-family the overrepresented YTH-domain containing protein belongs but given the two distinct roles and knowing that eIF3 was underrepresented in the AD NST while SAF was overrepresented in the 60% NST it is reasonable to speculate that it might be involved in exon inclusion splicing while also being involved in mRNA instability of repressed mRNA. This may present a unique manner in which *E. nindensis* decides which transcripts are required for translation, especially when considering that there are tens of thousands of differentially expressed genes but only a few are eventually translated.

La-related protein (Larp7) is yet another RNA binding protein found primarily in the nucleus where it participates in the binding of a UUU_{OH}-terminal stretch of nascent RNA produced by RNA polymerase III (Dock-Bregeon, et al., 2021; Maraia, et al., 2017). Structurally they comprise a La motif and at least one RNA recognition motif and their overarching role is in the protection of nascent RNA (Dock-Bregeon, et al., 2021). The proteome identified differential representation where one was overrepresented in the 25% and airdry NST and 25% ST and the other as underrepresented in 60% to 25% NST and at 25% ST with absence in the rehydration time points. This suggests that during dehydration there is sufficient protection of nascent RNA and the need for Larp7 expression is perhaps not that necessary but upon reaching drier states the production of Larp7 is increased to enhance the protection of newly transcribed RNA. Its absence in the rehydration time points further corroborates when water status is returned to normal the need to ensure enhanced protection of nascent RNA is not that high and could possibly indicate that under normal conditions it is lowly expressed as part of general nucleolar RNA turnover.

The *ATP-Binding cassette, subfamily F, member 3* (ABCF3), is one of many ABC transporters found in all biological systems which bind and hydrolyse ATP and couple it with transportation of nearly everything produced within plant cells (Do, et al., 2021). ABCF3 has two nucleotide binding domains but lacks transmembrane domains and is also known as *general control non-depressible* (GCN) which is homologous to the yeast GCN20 (Li, et al., 2018). Studies using Arabidopsis mutants indicated a role in translational regulation where it functions in concert with GCN20 (Izquierdo, et al., 2018). Aquaporins denoted as *PIPs* can facilitate the transport of H_2O_2 and it has been shown that overexpression of *ABCF3* enhanced root uptake of H_2O_2 and is suggested to actively regulate the expression of PIPs 1 and 2 (Li, et al., 2018). However, from Table 2-6 and Figure 8-2 the underrepresentation of *PIP* in 60%, 40%, airdry, and 1

week post-rehydration in the NST does not point to ABCF3's role in H_2O_2 transport. Its role however is more likely involved in activation of stress response genes as seen in Arabidopsis under both heat and drought stress (Rizhsky, et al., 2004). Given its general overrepresentation in all tissues it is more plausible to speculate that it is performing a 'housekeeping' role by enabling responsive genes to be activated.

The annotated *SCY1-like protein 1* is involved in the secretory pathway in plants and acts as a route for transporting proteins across the thylakoid membrane. In Arabidopsis two homologues are present, denoted as *AtSCY1* and *AtSCY2.* Loss of function of *AtSCY1* results in albino seedlings whereas loss of function of *AtSCY2* resulted in embryo lethality (Skalitzky, et al., 2010) further corroborated by Liu, et al., (2015). The authors reported that the *scy1-1* mutant failed to develop mature plastids and that when present only a few thylakoids were observed. Transcription of chloroplast-related proteins are activated via retrograde signalling from the chloroplast to the nucleus. Genes related to chlorophyll and components of the light harvesting complexes synthesis were down-regulated in the *scy1-1* (Liu, et al., 2015). Taken together and considering the underrepresentation of this protein in all dehydration RWCs it is safe to conclude that retrograde chloroplastic to nucleus signalling as a result of thylakoid disassembly is primarily responsible for the enrichment of GO terms associated with photosynthetic shut down.

Biosynthesis of certain proteins occurs within the ER and requires a suite of processes in order to translocate nascent proteins. One such translocon is the conserved Sec61 which facilitates protein access through vesicular trafficking where it forms a membrane channel for translocation of unfolded hydrophilic polypeptides (Shao, 2022). *PAT (*protein associated with the ER translocon) and *CCDC47* form a chaperone complex where *PAT* contains transmembrane domains (TMD) which acts as a shield to prevent degradation of the exposed TMD (Culver & Mariappan, 2020). A number of different translocation pathways are present in plants with organelles having their own set of translocation processes (Paul, et al., 2013) and is too complex to warrant full exploration herein, but sufficed to say, the underrepresentation of this chaperone in all RWCs except the 25% NST implies that the biosynthesis of multi-pass membrane proteins are halted or a different strategy for their assembly is used by *E. nindensis.* The other chaperone that showed decreased abundance was annotated as a DnaK chaperone and is the annotation assigned to the bacterial HSP70. Its decreased abundance is unclear, but it could likely be in reference to chloroplastic HSP70 which shares greater amino acid composition with the bacterial DnaK than it does to eukaryotic HSP70 owing to its endosymbiotic origin.

Processing of proteins includes not only their correct folding which is facilitated by a host of cellular chaperones but also the manner in which they are degraded. *Phytepsin*, which predominantly had overrepresentation in the dehydrating NST is an aspartic proteinase with many roles from plant defence, seed development, cellular protein turnover to liberate amino acids, and in degradation of chloroplasts during senescence (Simões & Faro, 2004). Peptidases can be grouped as either endo- or exopeptidase where exopeptidases are further divided into aminopeptidases or carboxypeptidase. Serine carboxypeptidase cleaves C-terminal residues from peptides through its serine-histidine-arginine catalytic site. In Arabidopsis multiple genes encode serine carboxypeptidases (SCPs in clade I) and serine carboxypeptidase-like proteins (SCPL in clade II) but their fundamental biological role remains largely unknown apart from their biochemical characterisation (Chen, et al., 2020). They are involved for example in seed development and germination, brassinosteroid signalling pathways, secondary metabolism, and UV protection to name a few (Chen, et al., 2020). Two of the most widely studied SCPLs are sinapoyl-glucose: malate sinapoyltransferase (SMT) and sinapoyl-glucose: choline acyltransferase (SCT) (Fraser, et al., 2007) and it is believed that many of the other members of this clade have similar enzymatic properties. The activity of peptide hydrolysis or acyltransferase is unknown, but it is tempting to speculate that it might be involved in the glucose: malate transferase process owing to the *malate dehydrogenase* activity noted earlier especially in light of the corresponding expression profile noted for both processes. It is also equally tempting to speculate that the activity might favour SCT activity since this is a process seen in the Brassicaceae seed development (Mock, et al., 1992).

COP9 signalosome complex subunit 3 was originally proposed to be involved in the repression of photomorphic development in plants but mutants of the *FUS11* gene, which encodes subunit 3 of *COP9 signalosome complex* had increased pools of ubiquitinated proteins which the authors concluded was due to regulation of the SCF type E3 ligase (Peng, et al., 2001). Given its great underrepresentation in all dehydration RWCs except at 25% NST, it is reasonable to conclude that protein ubiquitination is favoured and is presumably in the direction of ubiquitin-mediated proteolysis. Ubiquitination of target proteins during abiotic stress is common strategy readily seen as a means to 'switch' target proteins on or off (Doroodian & Hua, 2021) and is also a strategy used by other vegetative desiccation tolerant species such as *S. stapfianus* (O'Mahony & Oliver, 1999) and *X. humilis* (Collett, et al., 2004). Lastly, the great overrepresentation of the chloroplastic FtsH protease, which is a zinc metalloprotease with ATPase and is anchored to the membrane (Kato & Sakamoto, 2018), is likely coupled to various processes involved with the thylakoids. Although there was differential expression between different accessions, the predominant increase in fold change indicates once more preferential activation. The overrepresentation at 60% could likely be in favour of

repairing damaged D1 polypeptides which are degraded and recycled (Kato & Sakamoto, 2018) in an attempt to maintain photosynthetic efficiency during initiation of thylakoid disassembly. It is likely that during the time course of 60% to airdry, it plays a small role in degrading components of the chloroplasts where it can act in concert with SCY in retrograde chloroplastic to nucleus signalling. Its presence at the airdry state, which was the highest of all RWCs and also the highest in the airdry of all DAPs, could likely be due to increased synthesis which could be stored for rehydration as speculated for *X. viscosa* (Ingle, et al., 2007). When considered with the increase in ferritin which sequesters iron which is important for the synthesis of chlorophyll coupled with the associated increase in LEAs and HSP70 which might be involved in the formation of stress granules, it is a reasonable conclusion to make.

The overall genetic information processing, chaperone activity, and protein processing seems to imply that while there were differences between the NST and ST with respect to metabolic processes, processes presented here seem to be universal. Whether it be the accumulation of protective proteins such as LEAs or HSPs or the targeted degradation of ubiquitinated proteins, the ST and NST appear to utilise the same mechanisms and processes. While there were a few unique proteins in each of the categories exclusive to tissues, they were not very common and when present represented subunits of multi-unit complexes such as ribosomes.

Conclusion

The proteomic study done herein presents novel information to our understanding of the drivers of vegetative desiccation tolerance in *E. nindensis*. Having been able to use most of the same tissue from a transcriptomic time course allowed for the direct comparison of transcriptome associated transcripts with proteome proteins without the need to account for experimental set up and growth differences. The proteome was approached first through the lens of understanding the underlying biochemical mechanisms and how this could be reflected in the desiccation response and ultimately senescence. The goal of the proteome study was not to add redundant information on accepted desiccation tolerant features but to focus more exclusively on processes which are not mentioned in the literature which may have an equally valid effect on ensuring desiccation tolerance as do the accepted mechanisms. While there were inherent limitations owing to annotation methods used between the two data sets, the proteome as a whole demonstrated that there was greater translational regulation occurring which was postulated to be the case from the transcriptome which largely could not differentiate the NST from ST (Madden, 2019). The general lack of DAPs in the 48hrs post-

rehydration suggests that at least partial recovery and resumption of core metabolic processes has occurred resembling the full turgor reference. An interesting observation was that there were a significant number of DAPs in the 1 week post-rehydration time point which can only be attributed to the fully hydrated controls used not being fully hydrated and/or up-regulation of processes to increase fitness presumably in preparation for anthesis or perhaps priming for potential drought. Gene ontology enrichment clearly demonstrated the poikilochlorophyllous nature of *E. nindensis* and the importance of photosynthetic shut down, with retrograde chloroplast to nuclease signalling being the primary means of ensuring that during dehydration no new chloroplastic proteins are made. Enrichment analysis also demonstrated in some part that the disassembly of the thylakoids and chlorophylls inherently increases the iron pools. Iron by itself or via the Fenton reaction is problematic and can lead to ferroptosis. *E. nindensis* actively prevents this from occurring by sequestering Fe^{3+} which when taken together with increased FtsH protease activity at the airdry state is highly suggestive of storage of two important components required for rapid reassembly of thylakoids and chlorophyll.

Reprogramming of central carbohydrate metabolism indicated a shift towards pyruvate and malate production with the former being a direct intermediate for the TCA cycle and production of acetyl-CoA which can be shuttled to other processes such as malate production. Malate decarboxylase (oxaloacetate decarboxylating NADP+) had significant overrepresentation in all RWC suggesting at least from what was annotated it plays a very important role in carbohydrate recycling and replenishment of NADPH required by antioxidant enzymes. The underrepresentation of trehalose synthase with concomitant increase in raffinose synthase (especially in the airdry state) clearly demonstrated that *E. nindensis* is a raffinose accumulating resurrection plant which has also been validated through earlier studies on sugar contents in a number of grasses (Ghasempour, et al., 1998). Accumulation of sucrose which is a common strategy used by resurrection plants (Illing, et al., 2005) further supports its pivotal role in desiccation response. Starch production and breakdown appeared to be occurring counterintuitively where starch breakdown occurred early during dehydration and starch synthesis occurring towards the airdry state. The breakdown of starch naturally favours the increase in glucose pools and a hypothesis presented herein is that unused ‘glassy-state-forming’ glucose is re-polymerised and could serve an additional osmo-protectant and/or cell integrity role at the desiccated state. Alternatively, the active synthesis of starch could be likened to what is seen in seed maturation where stored starch is used as carbohydrate fuel source upon germination or in this case rehydration. Either way, a follow up study is needed to determine what is occurring with starch metabolism. The shift in processes that do not use ATP suggests that *E. nindensis* modulates its ATP pool such that ATP-dependent phosphorylation of proteins is favoured over dephosphorylation events. There is evidence to

suggest that jasmonate, salicylic acid, and ethylene have a role during desiccation. Whether they are involved in defence response to pathogens or are activated as a means to regulate cellular senescence remains to be elucidated. A comprehensive study of the hormonal changes occurring in *E. nindensis* is needed to determine whether there are active hormonal signals occurring. The appearance of a chitinase raises an interesting debate as to whether *E. nindensis* might arm itself against potential insect-herbivory or fungal pathogenesis. In an eco-physiological point of view this argument has merit since during dehydration there is enrichment of highly valuable carbohydrates and *E. nindensis* not only needs to protect these but also needs to ensure that the NST which is destined for 'resurrection' is not unintentionally damaged.

Amino acid and lipid metabolism appeared to be centred around degradation pathways. Processes which were favourably abundant were related to oxidative and osmotic stress response by for instance the accumulation of antioxidant enzymes and proline. Through exploration of the various DAPs in these two categories, it became apparent that *E. nindensis* has tight control over the H_2O_2 homeostasis in the cell; ensuring for instance that an adequate supply is present to ensure its positive role in signalling but that its rapid breakdown to prevent lipid peroxidation is accomplished. Further studies on the enzymatic activity of antioxidant proteins from *E. nindensis* are needed to determine if all of the classical antioxidant enzymes are present and to what extent do they match the overrepresentation of the ones presented here. In addition, it would be worthwhile to corroborate the H_2O_2 homeostasis through H_2O_2 assays. Lipid metabolism indicated that long-chain fatty acid synthesis was occurring, but that overall lipid supply was directed towards α- and β- oxidation in the peroxisome. Although lipidomic studies and ultrastructural studies implicated an increase in lipid content, as seen for instance in the maturation of orthodox oil-seeds, the proteome suggests rather that they are targets for acetyl-CoA production. A plausible speculation is that during dehydration lipids may congregate into larger oil bodies primarily due to a concentration effect of having less water. It is also plausible to speculate that lipid-associating proteins such as oleosin and caleosin may sequester some of the already present TAGs, or those derived from long-chain synthesis, where, and according to ultrastructural studies, they assemble along the plasma membrane as a means to stabilise it.

Genetic information processing, protein processing and chaperone activity indicated that there was a core set of abiotic stress responses initiated by both tissue types and illustrates that fundamentally the two tissues use the same method to mitigate stress and ensure continued cellular activity under water limiting conditions. Subtle and small changes between the NST and ST is likely what results in resurrection of the NST and not the ST.

Considering *E. nindensis* in its natural habitat where it is likely to undergo several dehydration cycles in relative short succession, it is tempting to speculate that it simply sacrifices the ST. It initiates similar processes such that the contents within the ST are not irreversibly damaged and may then act as a source of nutrients for the rehydrating NST and formation of new tissues or even inflorescence development. Both phenomena are observed post-desiccation. Considering this hypothesis that the ST is sacrificed to act as a nutrient source for the NST might provide *E. nindensis* with an evolutionary and selective advantage in being able to remobilise nutrients more rapidly once rain occurs. It light of this, it would be worthwhile to investigate the source-to-sink dynamics of *E. nindensis.*

To conclude, the phenomenon of vegetative desiccation tolerance is a fascinating one and will for many decades to come intrigue researchers to unravel the key mechanisms that underly it. The proteome of *E*. nindensis illustrated some well-accepted processes seen in many other resurrection plants. Attempts were made to delve deeper into the proteome beyond these accepted processes to determine whether there were unique processes exclusive to *E. nindensis*. Despite that there were a host of other DAPs not considered due to space limitations, this to some extent this was achieved since unique processes in the NST were identified. However, further exploration of lesser known proteins and their potential role in driving vegetative desiccation tolerance is needed. The proteome however significantly reduced the number of differentially expressed genes from the transcriptome which has a direct impact on reducing and enriching the selection of candidate genes for potential genetic transformation of important crops such as maize and more importantly, orphan crops such as the close relative *E. tef.*

Chapter 3: A consolidated omics view of desiccation tolerance in *Eragrostis nindensis.*

Introduction

Omics-based studies are arguably the bread-and-butter of gene discovery, especially in light of understanding stress adaptation. With the addition of multiple layers of omics data from genomics to metabolomics our understanding of the core set of processes which in the case of *E. nindensis* results in vegetative desiccation tolerance are enriched and the complexity of making sense of these processes is reduced. The untargeted proteome presented in Chapter 2 illustrated that there were a set of metabolic processes which underwent significant reprogramming to facilitate early drought response and later desiccation response. Chapter 2 illustrated that general stress response was common between the senescent and non-senescent leaves but the question of whether senescence is promoted or repressed respectively was not fully explored. In Chapter 2 the idea that small and subtle changes of proteins is probably responsible for differentiating the NST from the ST, but for the most part, these could not be fully explored or discussed in detail. In this Chapter, a small subset of proteins which have some relationship to senescence will be explored in the 25% RWC state between the two tissues to address the question of whether there are proteomic signatures which could explain the phenotype differences between the NST and ST. In addition, the core findings from Chapter 2 will be presented in three models exploring the proteomic changes occurring between early drought stress and moderate drought stress, changes occurring from moderate stress to the desiccated state, and changes occurring upon rehydration. Lastly, the results obtained from the proteome will be used to evaluate to what extent it validates the transcriptome.

Methods

For the comparison between the ST and NST 25% RWC, accession numbers from the proteome were separated for those belonging to proteins whose abundance was less than the fully hydrated control and those whose abundance was greater than the fully hydrated control for each tissue type and RWC. Duplicate accessions were removed from each list. These were then uploaded to Jvenny (https://bioinfogp.cnb.csic.es/tools/venny/index.html) (Bardou, et al., 2014) to find overlaps and unique features. For each category (NST up; NST down; ST up; ST down etc) the list of accessions was then uploaded onto UniProt using the ID mapping function to retrieve associated protein names. Proteins whose names were returned as ‘uncharacterised’ were first manually annotated using the *E. curvula* FASTA database. In

instances where that database also had it annotated as ‘uncharacterised protein’ the amino acid sequence was blasted using pBlast on NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi) by excluding *E. curvula (taxid:38414)*. Blast results were then sorted based on highest percent coverage and the associated protein was then selected. In instances where neither the FASTA database nor Blast could annotate the sequence, the associated accession was uploaded onto DAVID (https://david.ncifcrf.gov/) selecting the InterPro option to retrieve InterPro names. Protein names were then manually cleaned by removing redundant terms such as ‘domain containing protein’, EC numbers, N- and C-terminus, as well as duplicate protein annotations. Proteins annotated as ‘DUF’ were removed as were proteins which could not be annotated. The filtered and cleaned lists per tissue per RWC were then reuploaded onto Jvenny to generate a Venn diagram of protein names. Proteins selected and discussed were selected by considering their overall known function using databases such as NCBI, UniProt, and InterPro. Selected proteins does not represent an exhaustive list of senescence related proteins.

Results and discussion

3.3.1 Differential senescence-associated proteins hints at suppression of senescence in the NST

At 25% RWC tissues are in the late stages in the dehydration response where one can argue based on literature that commitment to desiccation tolerance or senescence has occurred. By this stage tissues will have laid the foundation in order to ensure the ability to enter this quiescent state. Ultrastructural studies indicated that at this stage there are great differences between the two tissue types (Madden, 2019) whereas the proteome in Chapter 2 indicated that there were very few changes occurring in the NST at this stage. A hypothesis for this observation is that during drying from 40% to 25% RWCs in the NST, very few new or enriched proteins are being produced as most of the proteins required for desiccation, and by extension senescence suppression, would have been made, and in further drying to the airdry state proteins required for rehydration are made. It should once again be noted that there are hundreds of proteins to consider but owing to space limitations, a selection of differentially abundant proteins (DAPs) which have some association with senescence will be considered for the 25% RWC state exclusively since the 60% RWC represents early drought response and is not at a relative water content which imposes commitment to desiccation response as made evident in Chapter 2. In the following section, senescence is not used in the sense of chlorophyll and thylakoid breakdown but rather the association it has with programmed cell death (PCD). As such, when it is stated that a protein suppresses senescence it is meant to

be interpreted as upon rehydration resumption of metabolic growth occurs and the tissue does not undergo PCD.

The protein deglycase DJ-1/PfpI-domain (Figure 3-1) has been extensively studied in animal systems owing to its close relationship with several human cancers. In plants however they have not garnered as much attention. In Arabidopsis three homologues to the mammalian DJ-1 gene have been identified and their proposed roles vary from oxidative stress response to glutathione metabolism to name a few (Lin, et al., 2011). Studies on one particular homolog *AtDJ1C* have shown that it is expressed more in younger leaves than in older leaves where in the younger leaves a gradient of expression was noted with the least towards the apex of the leaf indicating that it has an important role in overall leaf development (Lin, et al., 2011). Furthermore, disruption to the gene via T-DNA insertions resulted in albino seedlings with chloroplasts being smaller and devoid of thylakoids which when taken with chloroplast localisation strongly suggests a role in chloroplast development (Lin, et al., 2011). In another study the authors demonstrated that loss of function of *AtDJ-1a* caused accelerated cell death of aging plants and interacted with superoxide dismutase and glutathione peroxidase (Xu, et al., 2010). Its underrepresentation in ST in light of age-dependent development is understandable, however its overrepresentation in the NST is curious because at this stage all thylakoid membranes have been dismantled. This becomes more interesting when considering that in all the NST exclusively an annotated *CHLOROPLAST DEVELOPMENT AND BIOGENESIS 1* protein was identified through Blast. As the name suggests it is involved in chloroplast biogenesis and mutations to the gene result once more in an albino phenotype with severely compromised chloroplasts (Chen, et al., 2022). It is clear that there is very little activity occurring within the chloroplasts since chloroplastic-derived proteins are noticeably absent with the general absence of the 50s and 30s ribosomal units which ultimately make up the final 70s prokaryotic ribosome though interestingly the 30s is overrepresented in the 25% ST (Table 8-3). The appearance of these two proteins is suggestive of two possible scenarios; one, these could simply be residual proteins from thylakoid breakdown which have accumulated but this lacks substantiation owing to the observation of increased protein degradation via ubiquitinated-mediated proteolysis. The other possibility is that these are not degraded but rather protected in the NST exclusively as a means to ensure that thylakoid reassembly can occur rapidly upon rehydration. Its underrepresentation in the ST demonstrates for the first time that senescence in the ST is actively promoted by not having the required proteins to facilitate chloroplast biogenesis and thylakoid reassembly.

Glutamine amido-transferase type 2 according to InterPro description has a multitude of roles in the cell where the principal catalysis is involved in the removal of an ammonia from glutamine and subsequent transfer to create a new carbon-nitrogen group. Proteins with this

catalytic ability are involved for instance in the synthesis of purines, amino sugars, asparagine, and glutamate. Local alignment of the protein accessions associated with glutamine amido-transferase, namely A0A5J9VPI (NST) and A0A5J9SHW1 (ST) showed 35% sequence similarity. Blast results indicated that the NST protein is related to ferredoxin-dependent glutamate synthase (Fd-GOGAT) in the chloroplast whereas the ST protein is an NADH-dependent glutamate synthase 1(NADH-GOGAT) in the chloroplast. Both proteins have been implicated in regulating senescence through their role in nitrogen remobilisation. As mentioned in Chapter 2 senescence is a complex process and age-dependent and stress-induced onset of senescence is often indistinguishable from one another since the same hormonal signals and accompanying transcriptional activation occur during both processes. The balance between glutamine and glutamate is tightly controlled through activity of glutamine synthetase and GOGAT where the latter can either use ferredoxin or NADH as a reductant. Glutamine synthetase provides glutamine through the incorporation of NH_4^+ derived from recycled nitrogen (such as for example from protein catabolism and photorespiration) in an ATP-dependent reaction. GOGAT incorporates the glutamine with addition of 2-oxogluterate (from the TCA) to produce glutamate. This cycle can be further expanded with the synthesis of aspartate from asparagine. The increased conversion of glutamine to glutamate (Gln: Glu) and asparagine to aspartic acid (Asn: Asp) ratios has been shown to occur in Arabidopsis during onset of senescence (Watanabe, et al., 2013) and the conversion between these forms has further been implicated as key players in nitrogen transport. It has been proposed that Fd-GOGAT and NADH-GOGAT are involved in two different processes; the former is primarily involved in primary nitrogen assimilation and photorespiration whereas the latter is more involved in nitrogen recycling and remobilization (Lee & Masclaux-Daubresse, 2021). When NADH-GOGAT's expression is enhanced in rice, an increase in nitrogen remobilisation from older leaves to younger leaves was noted (Lee, et al., 2020) whereas mutations in Fd-GOGAT led to premature senescence (Zeng, et al., 2017). Strictly speaking, GOGAT is not restricted to the chloroplast. Given that *E. nindensis* is a C4 plant it is less likely to undergo photorespiration. In addition, at 25% RWC all Rubisco has been depleted (Table 8-3). Thus, the Fd-GOGAT presence is likely present in the mitochondria as a means to recycle nitrogen derived from catabolic processes. The absence of it in the ST as well as underrepresentation of the NADH-GOGAT suggests a few things. Firstly, it could imply that NADH/Fd-GOGAT is simply not translated in the ST which secondly suggests that glutamine is accumulating in the ST which thirdly corroborates the hypothesis that the ST is sacrificed as a means to provide nutrients for the rehydrated NST since glutamine is readily transported. The Fd-GOGAT presence in the NST is clearly in favour of glutamate production for antioxidant responses.

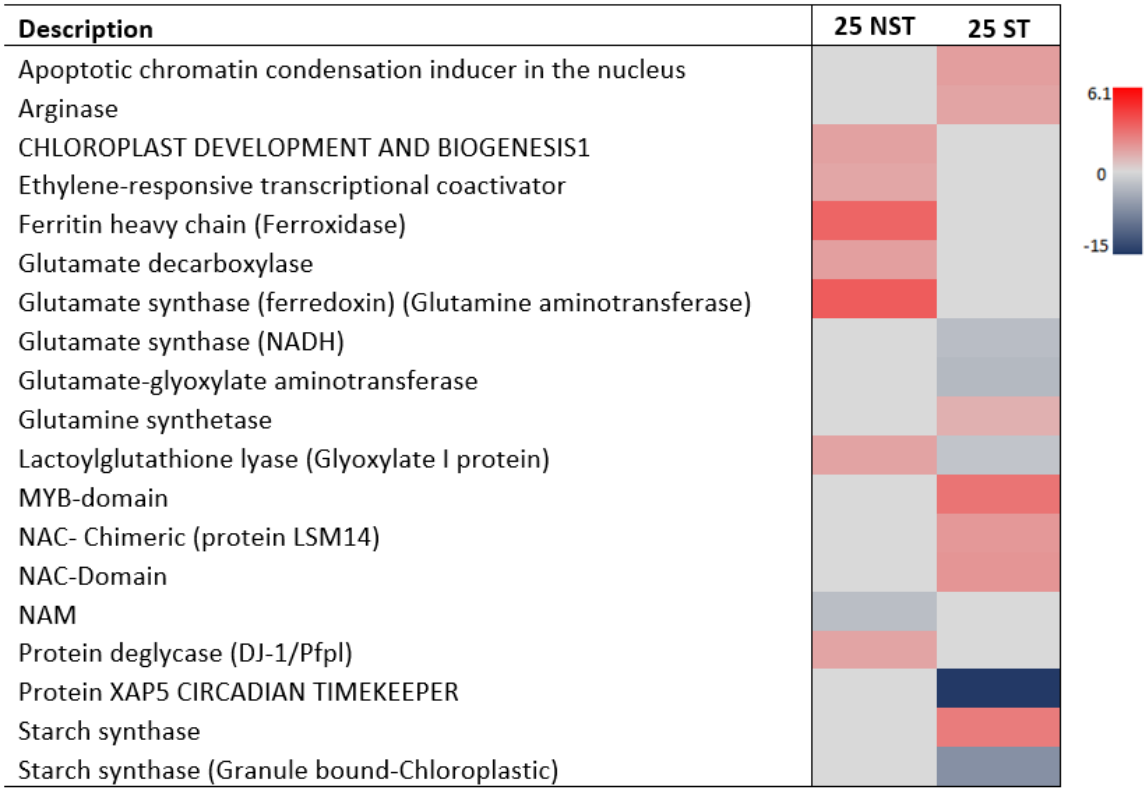


Figure 3-1: Selected senescence related terms for 25% RWC: Selection of senescence-related terms identified in the 25% RWC range for ST and NST. DAPs represent log2 fold change >2, <0.05 FDR relative to fully hydrated leaves.

Transcription factors (TFs) play a vital role in activation of a number of stress-induced processes. The problem however, with TFs is that their abundance is often so little that they tend to be completely missed in untargeted proteomics such as the one done herein. The transcriptome identified a number of TFs belonging to the WRKY, bZIP, and MYC families (Madden, 2019) but the proteome was only able to identify a fraction of them. For instance, a putative EcWRKY21 was only abundant at 40% RWC (Table 8-3). At 25% RWC, there were four differentially abundant TFs (TF-domain containing) proteins identified (Figure 3-1). Those being a bZIP (present in both tissues), MYB and NAC in the ST only, and NAM present in the NST. Basic leucine zipper (bZIP) is ubiquitous and have been shown to be activated through an array of pathways where their role in plant development is even more expansive (Lorenzo, 2019). MYB, like bZIP TFs, are involved in innumerable processes especially under abiotic stress where for example they can activate ABA-signalling pathways or enhance production of flavonoids (Wang, et al., 2021) and their role in senescence is likely via intermediatory (in)activation of associated pathways. A *MYB-domain* was identified as significantly abundant in the ST but not in the NST. This small occurrence hints to differential activation of MYB TFs

but overall role in senescence remains to be further elucidated. *NAC* (**N**AM **A**TAF 1/2 **C**UC2) is one of the largest TFs found in plants and has been shown to have differential effects on ABA-induced, age-dependent, abiotic-induced, and dark-induced senescence in rice (Lee & Masclaux-Daubresse, 2021). Without knowing which type of NAC this is, it is difficult to state what its role might be, but one can speculate that its overrepresentation in ST could be in favour of promoting senescence in the ST. *NAM* (No Apical Meristem) is yet another gene encoding a TF involved in development. As the name suggests, it is involved in meristematic processes but has been shown in rapeseed that it positively enhances progression of senescence by promoting the generation of ROS and PCD (Wang, et al., 2022). It is noticeably underrepresented in the NST which when taken into consideration with *NAC* and *MYB* in the ST demonstrates differential activation and inactivation of TFs involved in driving senescence.

One of the greatest underrepresented proteins identified in the 60, and 25% ST was a protein annotated as Protein FAM50 through Kolala annotation and was later elucidated as being the *XAP5 CIRCADIAN TIMEKEEPER* (XCT). This is a particularly interesting protein as its absence suggests several aspects of the ST which is not evident in the NST. XCT plays an active role in regulation of plant circadian clock (Zang, et al., 2023) but has also been implicated in regulating DNA damage response (Kumimoto, et al., 2021) and has been shown to have negative regulation downstream of the ethylene incentive response pathway (Ellison, et al., 2011). The ethylene pathway proceeds from the conversion of methionine to s-adenosylmethione via s-adenosylmethione synthase (SAM). This product is then converted to aminocyclopropane carboxylate (ACC) by ACC synthase. ACC through catalysis by ACC oxidase produces ethylene. Ethylene then in turn exerts regulatory effects on the ethylene signalling pathway where it ultimately leads to the activation of ethylene insensitive 3 (EIN3) and ethylene insensitive like 3 (EIL 3) (Dolgikh, et al., 2019) which ultimately lead to the activation of several genes related to an array of plant development processes including senescence. SAM is noticeably underrepresented in most dehydration RWCs, ACC synthase is not present as differentially abundant and ACC oxidase is only present at 60 and 25% ST (Table 8-3). The transcriptome indicated up-regulation of EIN3 (Figure 2-14) suggesting that the ACC produced in the ST is being converted to ethylene for ethylene response activation of downstream target genes. Knock-out mutants further showed enhanced ethylene responses (Ellison, et al., 2011) so thus having no XCT is clearly in favour of promoting ethylene response, but in the same breath mutants lacking XCT have pre-mRNA splicing defects (Zhang, et al., 2023) and are hypersensitive to UV-C and gamma radiation (Kumimoto, et al., 2021). The great underrepresentation in the ST and when taken in light of what seems to be up-regulation of EIN3 seems to suggest activation of EIN3 genes which are most likely in favour of promoting senescence. It should once again be made clear that at 25% RWC all

pigmentation is lost which implies that any senescence driven processes are not directed to thylakoid disassembly but could more likely be directed towards PCD. A contradiction to the difference between the ST and NST is that the NST has an ethylene responsive transcriptional co-activator as overrepresented. The closest Arabidopsis homologue to this protein is At2G42680.1 which on Tair is described as a Multiprotein Bridging Factor 1A (MBF1a). One study showed that triple knock-out mutants of MBF1 are much more sensitive to oxidative stress (Arce et al., 2010). The authors also indicate that MBF1 regulates an ethylene response factor (ABR1) where they concluded that *AtMBF1* plays a role in regulating the cross talk between ABA and ethylene. MBF1c by contrast has been shown to regulate thermotolerance in Arabidopsis (Suzuki, et al., 2008) and overexpression of the wheat *TaMBF1c* in ryegrass enhanced thermotolerance of ryegrass (Huang, et al., 2022). Thus, its appearance in the NST and not in the ST suggests better transcriptional (and subsequent translational) response to stress.

As mentioned in Chapter 2, the breakdown of thylakoids increases iron pools which can lead to iron-induced PCD. Having ferritin around to sequester the iron is an important step in preventing iron-induced PCD, which the ST completely lacks (Figure 2-7). A further argument can be made when considering protein glycation. Glycation is a natural process that occurs between reducing monosaccharides such as glucose and fructose and disaccharides such as maltose to name a few, where the reducing sugar interacts with N-terminal and lysine residues of proteins and can be likened to a form of post-translational modification. Glycation is a common by-product of regular cellular activity such as during the photosynthesis (Takagi, et al., 2014) and has further been shown to increase in an age-dependent manner in Arabidopsis (Bilova, et al., 2017). Naturally any form of protein modification has some effect on the protein function. Rabbani, et al., (2020) showed that certain glycating agents modified functional sites on proteins retendering them inactive. Increased glycation exasperated by abiotic stress leads to the formation of dicarbonyl stress which can lead to premature or accelerated senescence. To overcome the accumulation of glycated proteins, *lactoylglutathione lyase* is used. This protein incorporates methyl glyoxylate and glutathione (GSH) to produce lactate which can be further processed through electron transfer to pyruvate. In the NST, this protein was exclusively overrepresented whereas in the ST it was completely absent suggesting once more that senescence, and in particular most likely age-dependent senescence is being favoured in the ST and not in the NST. Lastly, and by no means exhaustive, the ST had overrepresentation of *apoptotic chromatin condensation inducer in the nucleus* (AIF) which was absent in the NST. Apoptosis differs between mammalian and plant system owing the fact that mammalian cells undergoing apoptosis forms apoptotic bodies which are then engulfed by phagocytes but in plants this of course does not occur. Plant cells will typically

still have structural components such as cell walls present when undergoing ‘apoptosis’ and can of course be a natural cellular differentiation event such as the differentiation of parenchymous cells to sclerenchymous cells or in the case of stress lead to PCD. Though *sensu stricto* fundamentally different from another, mammalian and plant cell apoptotic events have the hallmark signature of chromatin condensation (Sevrioukova, 2011; Latrasse, et al., 2016) which ultimately leads to apoptosis which seems to suggest that once more senescence is promoted in the ST but not in the NST.

By considering only a small subset of the entire proteome for the 25% RWC it is very clear that there are fundamental processes that differ between the ST and NST. Proteins presented here seem to suggest that senescence is suppressed in the NST despite it undergoing processes such as thylakoid disassembly. Redirection of metabolism favourably enhances the production of glutamate in the NST while glutamine production is favoured in the ST is highly suggestive of the ST becoming a source site for nitrogen remobilisation upon rehydration. This is further corroborated by the underrepresentation of *starch synthase* which was shown in Chapter 2 to be overrepresented in the air-dry NST. Not having it present demonstrates firstly that photosynthesis is likely not going to resume owing to underrepresentation of *DJ-1* and absence of chloroplast biogenesis proteins, and secondly that new starch is not going to be synthesised in the ST. As the primary carbohydrate storage molecule, a lack of starch synthesis is highly suggestive of a tissue which is not necessarily destined for metabolic reactivation.

3.3.2 Proteomic model for desiccation tolerance in *E. nindensis*

The vegetative desiccation tolerant response is often described as undergoing two unique processes. The first is classical drought response and is seen in many plant species since most plants can lose a fair amount of water before it becomes problematic. These responses as indicated in Chapter 1 often involve remodulation of morphological and physiological processes to mitigate further water loss. It is thus not surprising that at the 60% RWC in both tissue types a plethora of differentially abundant proteins are noted (Figure 2-5 for example), all of which are predominantly involved in stress mitigation and maintenance. At approximately 40-50% RWC, most, if not all, sensitive plants simply do not recover, presumably because they lack regulatory mechanisms which are not ‘switched on’. One classical example is not shutting down photosynthesis. As alluded in Chapter 1, photosynthesis is a very dangerous process under water-limiting conditions. It is at this 40% RWC stage when the desiccation tolerant response is initiated in resurrection plants where the primary goal is to ensure that tissues have the required resources to become fully desiccated and is reflected in the

proteome where a suite of processes are initiated at this stage which persists until the airdry state is reached. At the airdry state, desiccated plants presumably lay down important mRNA and proteins required for rehydration and can stay in this state for several months. Upon rehydration the remobilisation of stored RNA and proteins assists in the rapid recovery of important metabolic processes such as photosynthesis. With this in mind the importance of the processes involved in the transition from drought response to desiccation response will be considered.

In Figure 3-2, the first process to undergo significant changes occurs within the chloroplasts. Chlorophyll content from physiological studies indicated that at 40% RWC almost no chlorophyll was detected (Madden, 2019). This demonstrates that the desiccation response is reached when total disassembly of thylakoids and chlorophyll occurs. During this stage, latent starch from photosynthesis from higher RWCs is rapidly broken down to glucose where most of the glucose is converted to sucrose. Sucrose as mentioned in Chapter 2, plays a pivotal role in not only removing glucose but in providing osmoprotection where it could along with other organic acids and carbohydrates form NaDES, as proposed by Du Toit, et al., (2021).

Various hormones exert regulatory activity on the chloroplasts to initiate and maintain this breakdown. For instance, jasmonate derived through fatty acid metabolism, ethylene through amino acid metabolism, and salicylic acid-induced proteins ensure disassembly of the thylakoids and release of chloroplastic proteins. This disassembly inherently increases the iron pools which is sequestered by ferritin as a means to reduce ferroptosis.

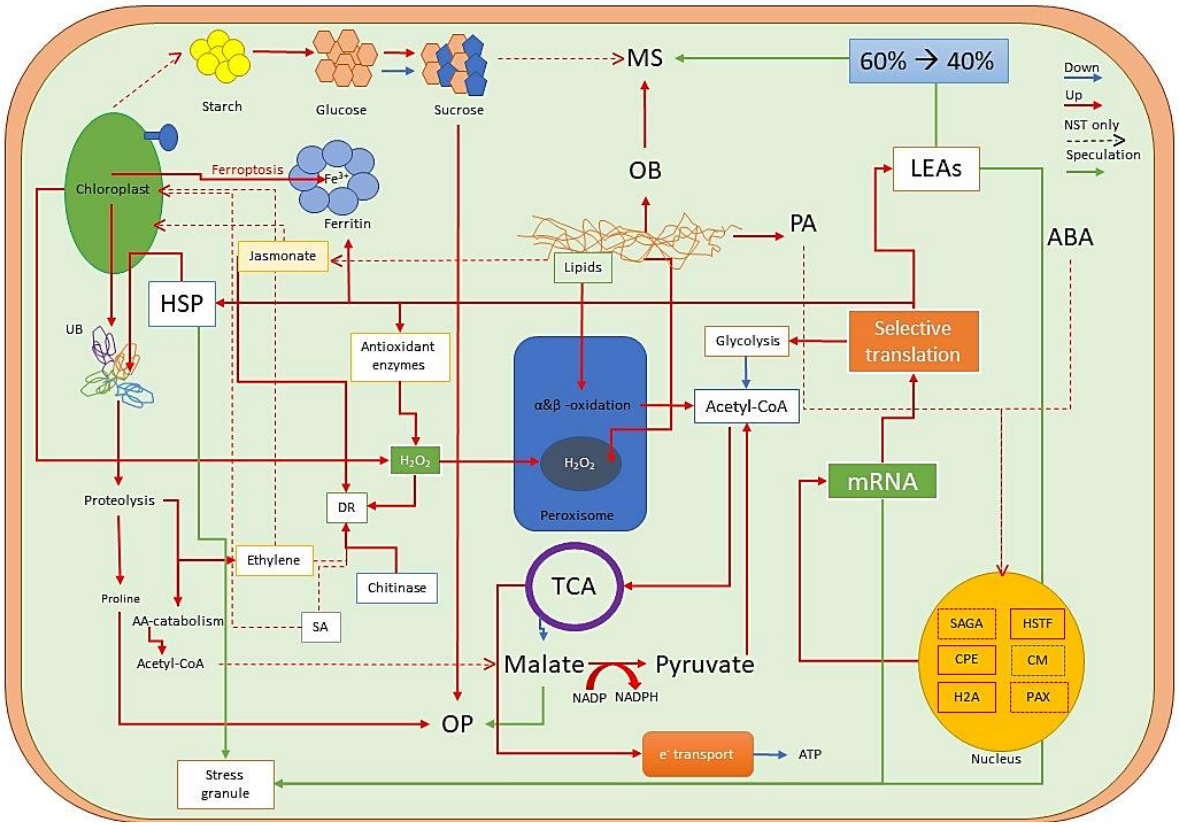


Figure 3-2: Drought to desiccation response in *E. nindensis*: Schematic summary of some proteomic changes occurring in the switch from drought response to desiccation response. Arrow in red represents processes which were overrepresented, blue arrows indicate underrepresented processes, arrows with dotted lines (irrespective of colour) were present in the NST only and green arrow represents speculation. MS= membrane stabilisation, OB= oil bodies, SA= salicylic acid; OP= osmoprotection; SAGA= stress response activation; HSTF= heat shock transcription factors; CPE; centromeric protein E; CM= chromatin remodelling; H2A= histone H2A; PAX= complex associated with DNA repair and histone methylation. All proteins were significantly different at FDR<0.05 with fold-change of 2 or more.

Proteins liberated from the chloroplasts are presumed to undergo ubiquitination owing to the presence of an E3 ligase and thus are likely to be targeted for ubiquitinated-mediated proteolysis. The constituents of this proteolysis such as the liberation of proline, are in favour of recycling carbon and providing osmoprotection. Proline, like sucrose, is one of the more commonly reported constituents of NaDES which further adds validity to the proposition of NaDES at lower RWCs. As proposed by Du Toit, et al., (2021) the formation of the NaDES is akin to proteins which undergo liquid-liquid phase separation to generate membraneless compartments where NaDES could accomplish something similar by sequestering important proteins. The accumulation of amino acids from proteolysis could also be in favour of directing acetyl-CoA towards production of malate, yet another commonly encountered NaDES constituent. As evident from the proteome there appears to be a great degree of malate to

pyruvate conversion, and this could be in favour of providing pyruvate for TCA purposes where the central role is in the recycling of $FADH_2$ since pyruvate is not produced via glycolysis. The conversion of malate to pyruvate is additionally in favour of replenishing NADPH pools which are required by several antioxidant enzymes. ATP production is halted or at least reduced significantly where central glycolytic enzymes using ADP are favoured. The full extent of lipid metabolic processes is primarily in the formation of oil bodies proposed to be derived from chloroplasts (Madden, 2019) which can act as membrane stabilisers along with sucrose and towards oxidation in the peroxisome to liberate acetyl-CoA. Selective translation of antioxidant enzymes coupled with peroxygenase activity from lipids assists in maintaining cellular H_2O_2 homeostasis. HSP synthesis is directed towards movement of chloroplastic-derived proteins to ubiquitination complexes and nascent peptide folding. HSPs along with LEAs could also be involved in providing thermo-protection and by extension osmoprotection where they may undergo phase separation to create an environment to either facilitate biochemical processes or sequester biomolecules in stress granules. Phosphatidic acid from lipid metabolism and appearance of ABA could have regulatory roles on gene expression in the nucleus where for instance heat shock transcription factors (HSTF) and SAGA (stress associated complex involved in transcriptional activation of stress response genes) are noticeably enriched. The activation of PAX (complex associated with DNA damage repair and histone methylation) and deactivation of CM (chromatin remodelling) points to a high degree of regulation occurring with the genome as a means to ensure that transcription is occurring and that damage to chromosomes is mitigated. The processes activated here lay the foundation for *E. nindensis* to enter desiccation response and while the majority of the processes occurred in both tissue types, reaching the airdry state appears to be predominantly regulated through hormonal and transcriptional activation of key genes in the NST.

Upon reaching the airdry state, a number of processes are noticeably diminished (Figure 3-3). The production of malate from amino acid-derived acetyl-CoA is no longer favoured. Sucrose production is halted but conversion of it to raffinose is highly favoured as well as the observation of a starch synthesis enzyme. Two possible explanations for this could be that whatever remaining glucose is present is re-polymerised to act as carbohydrate storage or it is simply laid down in preparation for rehydration and resumption of photosynthesis. The latter is further supported by the chloroplast to nucleus retrograde signalling which during early stages of dehydration informs the nucleus not to transcribe any new photosynthetic proteins but during the latter stages could be in favour of having these proteins pre-made in preparation for rehydration.

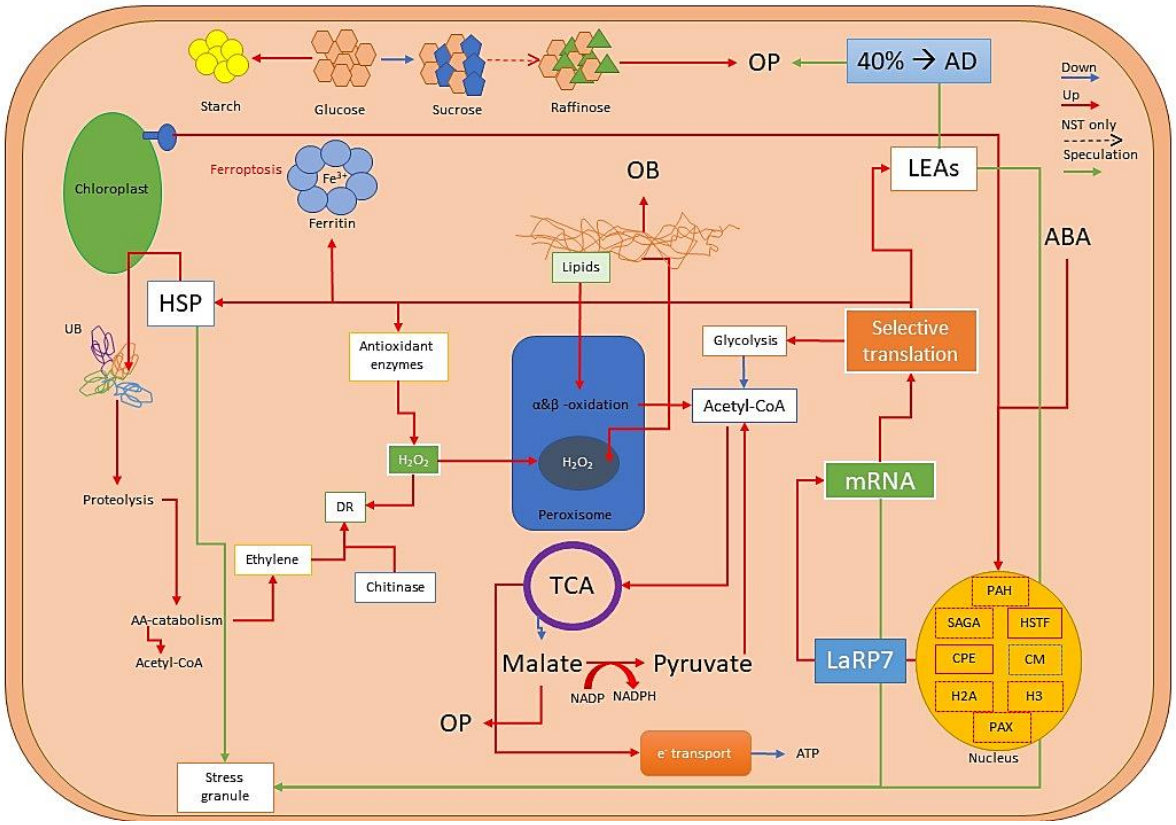


Figure 3-3: Desiccated state in *E. nindensis*: Schematic summary of some proteomic changes occurring in the final stages towards the desiccated state in preparation for rehydration. Arrows in red represent processes which were overrepresented, blue arrows indicate underrepresented processes, arrows with dotted lines (irrespective of colour) were present in the NST only and green arrow represents speculation. MS= membrane stabilisation, OB= oil bodies, SA= salicylic acid; OP= osmoprotection; PAH= paired amphipathic helix protein Sin3a; SAGA= stress response activation; HSTF= heat shock transcription factors; CPE; centromeric protein E; CM= chromatin remodelling; H2A and H3= histone H2A and H3; PAX= complex associated with DNA repair and histone methylation; LaRP7= La-related protein 7. All proteins were significantly different at FDR<0.05 with fold-change of 2 or more.

Similar constitutive processes are still evident such as the accumulation of HSPs and LEAs where their role is likely in the sequestration of these newly made proteins or transcripts. The emergence of LaRP7, involved in nascent mRNA protection, demonstrates that at this stage active transcription is occurring and that these transcripts are not required for desiccation but for rehydration. Antioxidant enzymes are still required but their role is probably more related to maintaining H_2O_2 homeostasis rather than ROS alleviation since this is more likely to be occurring at the early stages of dehydration. The presence of a great number of DAPs in the airdry state clearly supports the notion that the tissue is still metabolically active but is quiescent in nature and is preparing for rehydration. At this stage the biomolecules produced between 40% and airdry are vital at maintaining overall cellular viability.

Upon rehydration a noticeable underrepresentation of processes seen during dehydration is noted (Figure 3-4). The sequestered iron is remobilised into the chloroplasts where retrograde signalling once more informs the nucleus that it can resume production of chloroplastic proteins. Chloroplastic proteins which were protected by HSP70 can be redistributed to allow for rapid reassembly of chloroplasts and thylakoid membranes. This is further evident with the noticeable underrepresentation of SA-induced regulation over chloroplast degradation. The malate to pyruvate shuttle which proved important as a primary source of pyruvate for TCA is shut down. The conversion of sucrose to raffinose is shut down but synthesis of starch is favoured. Since photosynthesis is active once more, the appearance of antioxidant enzymes and lipid-derived peroxygenase illustrates not only their importance during dehydration but also in regular cellular ROS homeostasis since ROS is readily produced via both energy organelles.

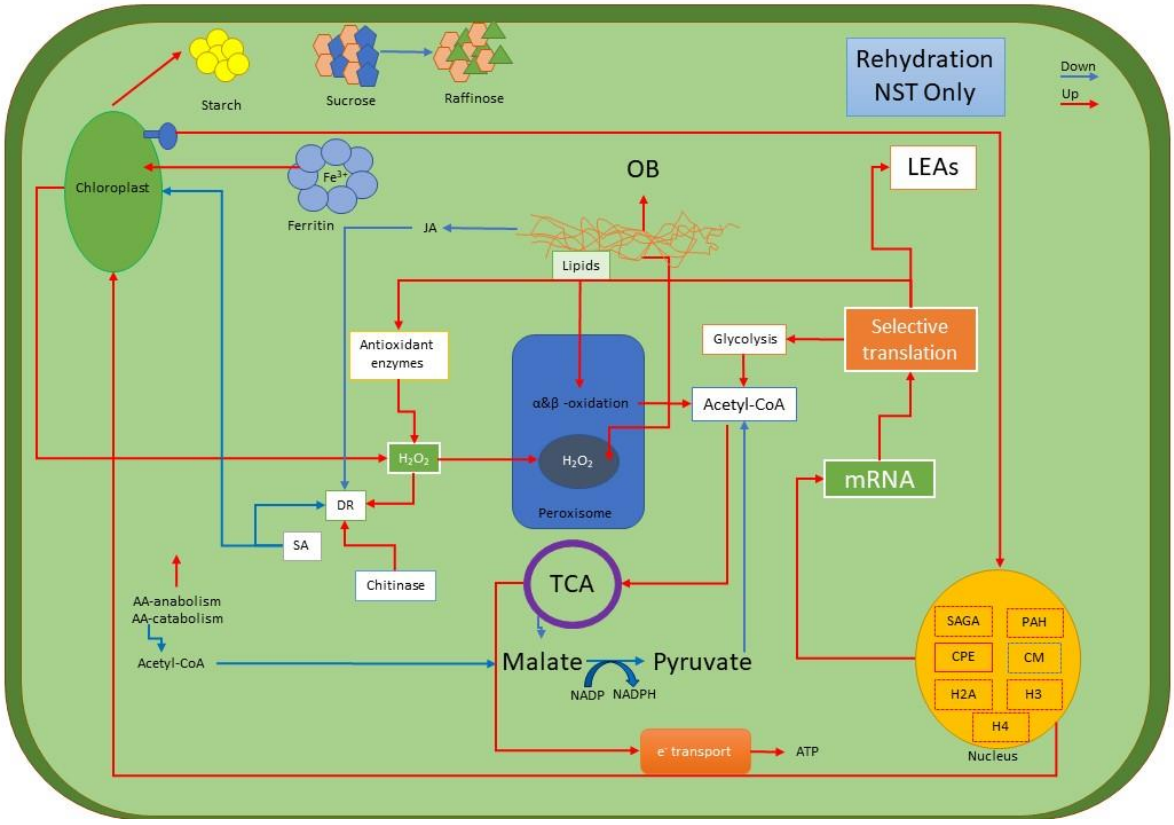


Figure 3-4: Rehydrated state in *E. nindensis*: Schematic summary of some proteomic changes occurring from the desiccated to rehydrated state. Arrows in red represent processes which were overrepresented and blue arrows indicate underrepresented processes. JA = jasmonic acid; OB= oil bodies, SA= salicylic acid; PAH= paired amphipathic helix protein Sin3a; SAGA= stress response activation; CPE; centromeric protein E; CM= chromatin remodelling; H2A, H3, and H4= histone H2A, H3, and H4;. All proteins were significantly different at FDR<0.05 with fold-change of 2 or more.

The sequestration of iron and photosynthetic related proteins by ferritin and chaperones respectively are indicative of *E. nindensis* using the airdry state as a means to ensure that upon rehydration photosynthesis can rapidly resume to replenish lost resources and to return to metabolic normality.

The overall process clearly demonstrates that like many other resurrection plants *E. nindensis* initiates responses at the early, moderate, and late desiccation stages that, firstly, ensures that it is able to enter a desiccated state and that, secondly, it has the required resources to ensure rapid recovery. To the best of my knowledge, this is the first instance where the role of ferroptosis is highlighted in a poikilochlorophyllous plant and the potential role of hormones in regulating overall desiccation and potentially suppression of senescence in the NST. There is also an argument to be made that cross talk between the hormones ensures correct activation of stress-related genes.

The transcriptome, which was accompanied by ultrastructural microscopy proposed the following model:

During drought response the NST undergoes vacuolation and chloroplasts shrink. Metabolic reprogramming occurs with additional emergence of oleosin-coated lipid droplets along the periphery of the cell while transcription and translation are maintained. In the ST however, vacuolation is insufficient but a large accumulation of starch is noted. The thylakoids are unstacked and the accumulation of enlarged osmophilic bodies is noted. Like the NST, transcription and translation are still occurring. During the desiccation response, the NST is mechanically and oxidatively stable. Complete photosynthetic shutdown has occurred while transcription and translation are still maintained, however, the suppression of transcriptional senescence is evident. In the ST by comparison, the cells are compressed with a lack of mechanical stabilisation. Like the NST, photosynthesis is shut down and translation and transcription are maintained but senescence is not supressed. At the desiccated state, the NST protects RNA and proteins are processed with accumulation of transcripts required for rehydration whereas in the ST the cells have completely collapsed with little evidence of RNA protection and inefficient translation. Upon rehydration the NST fully recovers structurally, and the chloroplasts are reassembled whereas in the ST, primarily due to complete cellular collapse, no recovery is evident.

The proteome largely corroborates and validates the observations made in the transcriptome and fully illustrates that there are translational control differences between the two tissues. The proteome demonstrated that there were significant changes occurring with proteins throughout dehydration and rehydration and between tissue types with some proteins displaying what can be considered as constitutively overexpression in response to stress, for

example HSPs and LEAs, whereas some were clearly induced under stressed conditions, for example sucrose and raffinose synthase. Although there were not that many differentially abundant proteins that differed between the two tissue types, akin to what is seen in the transcriptome, there were instances of certain proteins being differentially abundant in the NST. For example, the high enrichment of hormone-related proteins, raffinose synthase, and proteins associated with transcriptional regulation demonstrated that it is more likely that smaller changes occur that lead to differences between the two rather than overtly obvious changes. As such, there were proteomic differences between the two tissues but the contribution they have in establishing desiccation tolerance remains to be further elucidated. From terms that could be matched between the two omics data sets, there did not appear to be a great deal of temporal or spatial variation from transcription to translation. There were of course some examples of temporal delay, for example with raffinose synthase, but for the most part, many of the transcripts which could be matched, were translated at the same time, which clearly demonstrated that *E. nindensis* exerts great translational control presumably through differential mRNA methylation owing to the presence of proteins associated with RNA methylation. The relative expression (in terms of fold change) of matched transcripts and proteins appeared to be very similar. Lastly, coupling proteome with transcriptome for the same tissues greatly improves our understanding of what drives desiccation tolerance in the non-senescent leaves of *E. nindensis*. Prior to delving deeper into the proteome, it appeared as if *E. nindensis* was not conforming to expected strategies seen in other vegetative desiccation tolerant plants and it was presumed that it simply ‘does its own thing’. However, the proteome when looked through the lens of the transcriptome illustrates that it does not ‘do its own thing’ and that it indeed undergoes a well-orchestrated set of processes to ensure desiccation survival and complete and total metabolic reactivation upon rehydration.

Chapter 4: Functional characterization of the heat shock protein 70.4 from *Eragrostis nindensis*

Introduction

The selection of candidate genes for genetic engineering for the sake of improving a stress response can be a lengthy process. It often starts with an analysis of the transcripts from a stress tolerant species associated with that particular stress and comparing it to a relative which is sensitive. The global transcriptome of *E. nindensis* as noted in Chapter 2, displayed an array of differentially expressed genes between the drought tolerant non-senescent tissue (NST) and the drought sensitive senescent tissue (ST). One of the highest differentially expressed genes from the transcriptome was the heat shock protein 70 (HSP70) whose transcript abundance emerges at ~60% RWC and persists throughout the dehydration and rehydration process. The proteome from Chapters 2 and 3 further illustrated that HSP70 plays a significant role during dehydration stress. This Chapter is thus aimed at exploring an HSP70 from *E. nindensis* in more detail.

Heat shock proteins: a review

As the name implies, HSPs were initially identified as proteins that accumulate in response to heat (Ritossa, 1996) however, the presence of HSPs under drought, salinity, pathogenicity, and insect herbivory (Park & Seo, 2015) is indicative of a much greater role in general response to stress over and above their accepted role as chaperone proteins involved in a number of cellular protein dynamics including folding and degradation (Park & Seo, 2015; Berka, et al., 2022). The HSP nomenclature is based on the molecular weight of the protein and are thus grouped as either HSP100,90,70,60 and the small HSPs. The roles for the different HSPs are quite varied. HSP100 for example, which is most commonly localised in the cytoplasm and nucleus is involved in processes that target protein aggregates such as re-solubilization. HSP70 by contrast can localise to the chloroplast, mitochondria, cytoplasm, endoplasmic reticulum, and the nucleus and is primarily involved in the unfolding of misfolded peptides and the translocation of unfolded proteins. Phylogenetic analysis of HSP70 have indicated that the cellular localisation of the four major HSP70s are monophyletic (Krenek, et al., 2013; Boorstein, et al., 1994) and that the mitochondrial and chloroplastic HSP70s are more related to the DnaK homologues in cyanobacteria (Boorstein, et al., 1994) consistent with endosymbiotic origin of these plastids.

HSP70 comprises a large family of chaperone proteins and is evolutionary conserved from prokaryotes to eukaryotes with the bacterial DnaK being approximately 50% similar to the

eukaryotic HSP70. Interestingly, some archaea lack genes encoding for HSP70 (Macario, et al., 1999) but those that have been studied which have genes encoding them are more related to the bacterial DnaK than to the eukaryotic HSP70 suggesting that either the gene was lost or acquired (Gribaldo, et al., 1999). The expression pattern of HSP70 (constitutive and inducible), appearance in all forms of life, and high degree of sequence conservation demonstrates that it is not only ancient in origin, but 'a critical physiological trait' (Bettencourt, et al., 1999). Eukaryotes have multiple genes encoding HSP70 proteins which most likely arose from multiple duplication events as seen in *Paramecium caudatum* (Krenek, et al., 2013), *Saccharomyces cerevisiae* (Boorstein, et al., 1994) and *Physcomitrella patens* (Tang, et al., 2016) and evolved by mechanisms such as purifying selection. For example, humans have at least eight unique members differing in amino acid composition and cellular localisation (Duagaard, et al., 2007; Tavaria, et al., 1996). The rationale for having multiple genes is a difficult one to deduce. In the case of humans Hsp70-1a, Hsp70-1b, and Hsp70-6 are stress induced whereas the others namely Hsc70, Hsp701t, and Hsp70-2 are not, suggesting from this perspective that they serve two different roles; one of stress response and one of proteostasis housekeeping (Duagaard, et al., 2007).

Amino acid sequence composition can distinguish cellular locations from one another, for example, the cytosolic EEVD motif and RARFEEL (Yu, et al., 2021), mitochondrial GDAWV and YSPSQL in *Physcomitrella patens* (Tang, et al., 2016), endoplasmic reticulum HDEL/KDEL (Yu, et al., 2021), and chloroplastic DVIDADFTDSK in *Physcomitrella patens* (Tang, et al., 2016). The DnaK archetype has four structural motifs, a nucleotide-binding domain (an ATPase domain) at the N-terminus of approximately 45 kDa, a 15 kDa and 10 kDa substrate-binding domain denoted as beta and alpha respectively, and a C-terminal disordered region with a high degree of variation (Berka, et al., 2022; Sarkar, et al., 2012).. The two domains are linked by a short linker region that facilitate the interactions between the N- and C-terminal domains. The linker region is much less conserved which has implications for the flexibility of the domains and ultimately conformational arrangement of them and could potentially elucidate different or unique functions between species under various stress conditions (Yusof, et al., 2022).

HSP70 is one of the more researched HSPs in animal studies, in particular in reference to cancer research and general disease immunity but as Berka, et al., (2022) suggest, they are often overlooked in the world of plant stress response. HSP70 is by definition a ~ 70 kDa protein with a high degree of conservation as reported by Boorstein, et al., (1994).

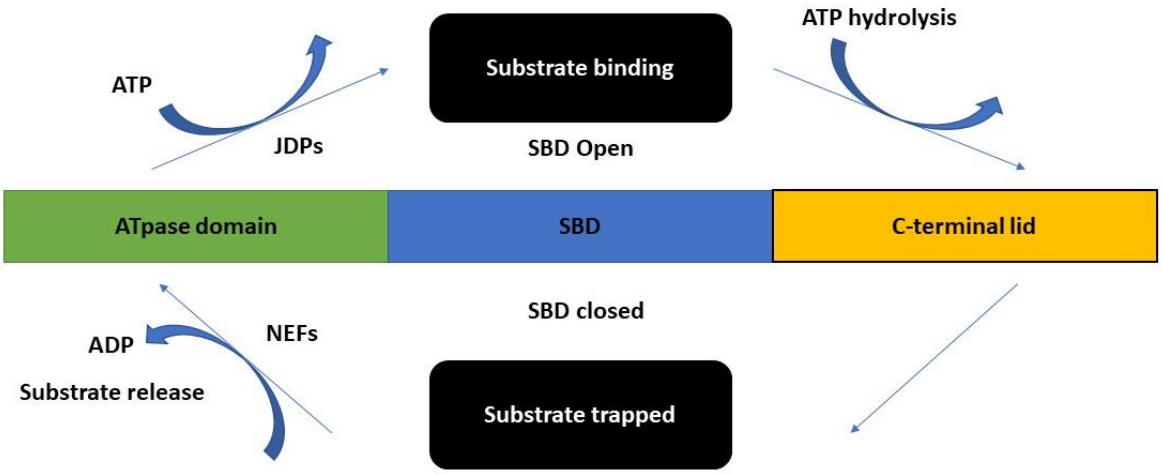


Figure 4-1: HSP70 Chaperone model: graphical representation of HSP70 chaperone activity. JDPs = J-domain proteins, SBD = substrate binding domain, NEFs = nucleotide exchange factors. Binding of JDPs to ATpase facilitates ATP binding causing the SBD to adopt an open conformation allowing substrates to bind. Hydrolysis of ATP results in the substrate being trapped in the SBD. NEFs exchange ADP with ATP resulting in the release of the substrate. Figure adapted from Berka, et al., (2022) and Sarkar, et al.,(2012)

Through recent advances in optical tweezer technology, the canonical model for how HSP70 functions, which has been widely accepted, has been extended. The accepted mode of action requires an orchestrated cycle of binding and release of peptides involving ATP and co-chaperones (Wruck, et al., 2018). Early studies revealed that the peptide-binding region can be highly specific for certain amino acid residues (Fourie, et al., 1994) though hydrophobic residues tend to be the preferred residues. The preference for hydrophobic residues for mis-folded recognition demonstrates that for instance the cytosolic HSP70 is able to recognise that there are proteins displaying residues which they should not be. The mechanism by which HSP70 performs its chaperone duty is best studied and understood in the bacterial system where HSP70 is denoted as DnaK, HSP40 as DnaJ (a J-domain protein), and a nuclear exchange factor termed GrpE. Briefly and adapted from Imamoglu, et al., (2020) and Lu, et al., (2021), HSP70 can, depending on the ATP/ADP status of the NBD, exist in two conformations. When it is in the ATP-state, it exists in an open conformation such that the inner hydrophobic region of SBD is exposed. In its ADP-state, the NBD and SBD are loosely associated, and it adopts a closed conformation where the α-helical 'lid' is closed. The association and dissociation dynamics differ between these two states. When it is open, the rate of substrate binding and release is much greater compared to the closed state. The binding of DnaJ accelerates the hydrolysis of ATP which is bound to the NBD which results in the formation of the ADP-state. An exchange between ADP and ATP is facilitated by the binding of GrpE to the NBD. This exchange subsequently results in the release of the

substrate. This well-orchestrated allosterically regulated cycle from open- to closed-state conformation can not only protect the integrity of unfolded proteins but can also assist in refolding the misfolded protein (Lu, et al., 2021) which would otherwise aggregate (Imamoglu, et al., 2020) and cause severe perturbations to overall proteostasis.

Overexpression of *HSP70* has been implicated in enhancing tolerance to a number of abiotic stresses. For example, overexpression of *NtHSP70-1* in tobacco demonstrated that it was a drought-/abscisic acid-inducible gene (Cho & Hong, 2005). The authors further demonstrated that overexpression resulted in tobacco seedlings maintaining leaf turgidity under water deficit stressed conditions when compared to the vector only and antisense transgenics despite a gradual decrease in leaf water potential. Interestingly the authors also reported that the expression of the early response to dehydration (ERD15) was reduced in the overexpressing plants (Cho & Hong, 2005). In another study, where the *HSP70* from *Brassica campestris,* a plant shown to be an intermediate in response to drought (Ashraf & Mehmood, 1990), was overexpressed in tobacco improved physiological parameters were observed when compared to the wild type (Wang, et al., 2016). The authors reported higher chlorophyll content across the treatment time, a significant reduction in the amount of malondialdehyde (an indicator of lipid peroxidation) and increased activity of superoxide dismutase and peroxidase in the overexpressing plant compared to the wild type (Wang, et al., 2016). Enhanced tolerance to salinity and drought stress of transgenic sugarcane (*Saccharum* spp. Hybrid) overexpressing *HSP70* from a drought tolerant perennial grass *Erianthus arundinaceus* has also been reported (Augustine, et al., 2015). When the *HSP70* from *Chrysanthemum morifolium* is overexpressed in *A. thaliana* similar trends of enhanced tolerance to drought stress were reported (Song, et al., 2014). Overexpression of *HSP70* from *Macrolyloma uniflorum*, a drought tolerant legume, in *A. thaliana* resulted in improved tolerance to drought, salinity, heat, and cold stress (Masand & Yadav, 2016). These results when taken together imply that HSP70 has significant effects on the tolerance of sensitive plants to abiotic stress and may most likely be due to the chaperone function of HSP70 whereby overexpression of it results in a greater turnover of damaged or misfolded proteins.

Liquid-liquid phase separation

The cellular environment is a complex and dynamic one and becomes even more complex under an array of conditions. In the traditional view of a cell, functions are often compartmentalised in membrane bound organelles for instance photosynthesis within the chloroplast. However, the existence of membraneless compartments such as p-bodies and stress granules raise a number of questions about their origin, assembly, stability, and function. In 2009 the idea of phase separation was proposed and upon further investigation

from the scientific community it was revealed that many of these membraneless compartments undergo liquid-liquid phase separation (LLPS) in their formation. Phase separation can in its simplest form be described as the propensity of biomolecules to spontaneously separate into phases that can be distinguished from one another (Xu, et al., 2021). The ability for biomolecules to phase separate is dependent on a number of factors, including but not limited to the concentration, presence of other molecules, pH, and salt. Proteins that are able to phase separate are often characterized as having what are described as intrinsically disordered regions (IDRs), so called for the lack of propensity of these regions to produce folded domains. Low-complexity sequence regions, such as tandem repeats of asparagine and glutamine are often found within these IDRs.

Chaperone proteins such as HSP70 and the LEAs have been implicated in regulating protein homeostasis of membraneless organelles (Li, et al., 2022c) and assist in retaining the function of stress granule components. Stress granules form an integral part of stress response by sequestering important mRNA, preinitiation factors, and RNA binding proteins (Kedersha, et al., 2013). (Li, et al., 2022c) demonstrated that HSP70 is able to undergo LLPS with increasing protein and PEG 3350 concentration and that both termini of the protein are able to phase separate though truncation of the protein resulted in a lesser degree of LLPS. Although their study focused on the role of HSP70 in the context of amyotrophic lateral sclerosis, their results nevertheless demonstrate that HSP70 can phase separate *in vitro* and plays a role in stress granule homeostasis by preventing the liquid-to-solid formation of stress granules. The formation of liquid condensates is well known in the context of transcription where transcriptional components either in tandem or independently form discreet condensates at transcriptional activation sites where it is hypothesised that the formation of these condensates improves the efficiency of transcription for example as seen in tumorigenesis (Lu, et al., 2020). In the context of abiotic stress, one can argue that the propensity of chaperone proteins to undergo phase separation might improve their chaperone ability over and above the ‘removal’ of misfolded proteins from the cytosol since the accumulation of misfolded proteins can have a negative impact on the proteostasis of the cell. It is thus feasible to speculate that HSP70 from *E. nindensis* might undergo phase separation in an attempt to maintain proteostasis while the cell undergoes reshuffling of metabolic activity to enter this quiescent state. It is also plausible to speculate that chaperone proteins might encase important rehydration transcripts or protein during this phase separation and have these at the ready upon rehydration.

Chapter aims and objectives:

This Chapter is primarily aimed at gaining a better understanding of a selected HSP70 from *E. nindensis*. The Chapter specifically is aimed at answering the following:

1. What is the effect of overexpression of EnHSP70 in *A. thaliana*?
2. Is this selected EnHSP70 able to undergo liquid-liquid phase separation? If so what conditions facilitate the formation of condensates?
3. Where in the cell might this EnHSP70 localise and what could be the implications in relation to *E. nindensis* and desiccation tolerance?
4. Can EnHSP70 improve stress tolerance to high salt and/or mannitol in BL21 cells?
5. Can EnHSP70 improve the thermotolerance of lactate dehydrogenase?

Methods

4.3.1 Bioinformatic analysis of EnHSP70

The coding sequence for *EnHSP70* was obtained by blasting an *Arabidopsis thaliana HSP70* from https://www.arabidopsis.org/ against the genome of *E. nindensis* id 54689 PacBio unmasked vV2.1 on https://genomevolution.org/coge/ website using the function 'GoGeBlast' with parameters set at E-value of 1.0 x 10-5, Gap cost with existence at 5 and extension at 2. The highest scored alignments were then selected and the FASTA files submitted onto NCBI blast with parameters set for 'highly similar' sequences. Once it was verified that the sequence obtained through genome blasting returned hits for other *HPS70* sequences, the sequence was then further analysed from the RNA transcriptome FASTA file to verify its appearance in the *de novo* assembly of the transcriptome (Madden, 2019). The RNA sequence was then translated into amino acid sequence using the online webpage https://web.expasy.org/translate/. The amino acid sequence was then further scrutinised by blasting the sequence using protein blast on NCBI, this time selecting the newly added experimental data set option. All parameters for nucleotide and peptide blasts were set to default. Amino acid sequences from the top hits with the highest E-value score were then selected and aligned to EnHSP70 using ClustalOmega multiple sequence alignment online tool (https://www.ebi.ac.uk/Tools/msa/clustalo/). The multiple sequence alignment and phylogenetic tree analysis was calculated from default setting on the website. Protoparam (http://web.expasy.org/protparam/) was used to evaluate the amino acid composition as well as retrieving information regarding the stability, aliphatic index score and grand average of hydropathicity. Signal peptide predictions were done using SingalP v.4.1 server (http://www.cbs.dtu.dk/services/SignalP/), PrediSi (http://www.predisi.de/), and TargetP 1.1 Server (http://www.cbs.dtu.dk/services/TargetP/). The online webserver, MoRFpred was used to predict the number of amino acid residues of molecular recognition features (MoRFs) (http://biomine.cs.vcu.edu/servers/MoRFpred/). MoRFs are small segments of a protein which has the ability upon binding to target molecules to undergo a disorder-to-order transition and are considered to be a unique class of intrinsically disordered regions (Vacic, et al., 2008). Furthermore, EnHSP70 was modelled using the online software, trRosetta developed by Jianyi Yang's research group at Nankai University, https://yanglab.nankai.edu.cn/trRosetta/. Their software NucBind was used to predict nucleic acid binding, COACH-D was used for protein-ligand complex prediction, CoABind was used for CoA-binding site prediction, and PepBind was used for peptide-binding site prediction. Disorder prediction was done using PONDR (http://www.pondr.com/) using the VLXT and VSL2 predictors. In addition, disorder was predicted using MFDp2 (Mizianty, et al., 2013). Subcellular localisation of EnHSP70 was

done using the following online tools: BUSCA (Savojardo, et al., 2018), DeepLoc – 1.0 (Armenteros, et al., 2017), Localizer (Sperschneider, et al., 2017), Plant-mSubP (Sahu, et al., 2019), pLoc-mPlant (Cheng, et al., 2017), LocTree3 (Goldberg, et al., 2014), and TargetP – 2.0 (Armenteros, et al., 2019). The sequences of the up-regulated HSP70 from the transcriptome (Madden, 2019) were also aligned to the chosen sequence to determine sequence similarity.

4.3.2 Molecular cloning

The following section describes the various methods and techniques used for the molecular cloning of the *HSP70* sequence from *E. nindensis* into pB2WG7 for overexpression of *EnHSP70* in *A. thaliana* ecotype Col-0 and transient expression in *Nicotiana benthamiana*. *EnHSP70* was also cloned into the bacterial expression vector RY8686 for analysing phase separation. For localisation, *EnHSP70* was fused downstream of mCherry and cloned into pB2WG7. Schematic representation of the three constructs is shown in Figure 4-2.

4.3.2.1 Amplification of EnHSP70

The cDNA sequence for *EnHSP70* was synthesised by GenScript. The synthetic gene was cloned into pUC18 at GenScript and sent as a lyophilised pellet. The pellet was resuspended in 20µL nuclease-free water and transformed in competent DH5α cells (NEB) and plated onto 100 µg.mL^{-1} ampicillin plates. Colonies were selected and inoculated into 5 mL LB media and plasmid isolated using the Zyppy mini-prep kit from Zymo research as per the manufacturer's recommendations. The purified plasmid was then used as template DNA for subsequent PCR and digests. Primers as outlined in Table 5-1 below were designed to add a PacI and SpeI site on the 5'-end of *EnHSP70* and adding a BamHI site to the 3' end. In addition, the 6X His-tag was added to the 3'-end of the *EnHSP70* for protein purification. *EnHSP70* was amplified using Q5 Hot Start Hi-fi 2x ready mix (NEB) using an annealing temperature of 60°C and extension for 1.30 mins. PCR products were visualised by electrophoresing products on a 1% (*w/v*) agarose ethidium bromide-stained gel at 100 V for 30-45 mins and visualised using a UV illuminator.

Table 4-1: Primers used in study. Underlined regions indicate restriction enzyme sites. Lowercase letters indicate cleavage overhangs.

Primer Name	Sequence (5' → 3')	T_A (°C)	Purpose
EnHSP70 FOR	cccc**TTAATTAACTAGT**aaaATGGCGGCCTCCCAGAGCC	60	Overexpression
EnHSP70 REV	ggcc**GGATCCC**TAGTGATGGTGATGGTGATGGGAC		
EnHSP70_mC FOR	caggt**GGATCC**ATGGCGGCCTCCCAGAG	60	Phase separation
EnHSP70_mC REV	gagt**GCGGCCG**CCTAGGACTTCCTGTCTTC		
MCLOC FOR	ggaaa**CCCGGG**AAAATGGTGTCTAAAGGCG	60	Subcellular localisation
MCLOC REV	gccat**GGATCC**ACCTGCGGCAGAACCCGCA		
EnHSP70_LOC REV	gacc**ACGCGT**CTAGGACTTCCTGTCTTCCC	60	
BAR FOR	AAGTCCAGCTGCCAGAAACC	55	Arabidopsis screening
BAR REV	GAACTGACAGAACCGCAACG		
qEnHSP70 FOR	GTGTTTGATCTTGGCGGTGG	60	qPCR Arabidopsis screening
qEnHSP70 REV	CCCATCATGTAGCGAGCACT		
SAND FOR	CAGACAAGGCGATGGCGATA	65	qPCR Reference
SAND REV	CTTTCTCTCAAGGGTTT CTGGGT		

4.3.2.2 Overexpression construct

For pB2WG7, sequential digests with BamHI and SmaI were done. A total of 4 µg of plasmid DNA was digested. The *EnHSP70* amplicon using the *EnHSP70* forward and reverse primer pair was digested singly with BamHI. The backbones and amplicons were gel excised and purified using the Wizard SV gel purification kit as per manufacturer's recommendations. Digested backbones were transformed into competent DH5α cells to screen for successful digestion. Ligations were set up using a 3:1 ratio of digested EnHSP70:pB2WG7. T4 DNA ligase (Sigma) was used at 1.5-fold excess and incubated overnight at room temperature. A total of 5 µL of the ligation reaction was transformed into DH5α cells (NEB) and plated on 100 $\mu g.mL^{-1}$ spectinomycin supplemented LB plates. Colonies were picked and colony PCR performed to screen for successful ligations and to remove any re-ligated plasmids. Successful colony PCR colonies were then grown overnight in 5 mL spectinomycin supplemented LB, and plasmid DNA isolated as before. Isolated plasmids where then digested with HindIII (which yields two distinct bands if the ligation was successful and only a single band in the empty vector) and double-digested with EcoRI and BamHI (which yields two distinct bands if the ligation was successful and three bands in the empty vector). PCR was also done on the isolated plasmids. Those that showed correct banding pattern were sent for Sanger sequencing at SUNCAF using the *EnHSP70* qPCR primer set above.

Figure 4-2: Maps of EnHSP70 final constructs: EnHSP70 cloned downstream of CAM35s promoter in pB2WG7 vector using SmaI and BamHI sites for overexpression. EnHSP70 cloned downstream of mCherry using BamHI and NotI sites for phase separation. mCherry and EnHSP70 cloned into pB2WG7 for localisation using SmaI, BamHI, and MluI. Map generated using SnapGene viewer. Full length sequences for constructs can be found in section 8.1.

4.3.2.3 Liquid-liquid phase separation construct

For cloning into the bacterial expression vector RY8686 for liquid-liquid phase separation, the vector was double-digested with NotI and BamHI using a 2-fold excess of NotI. The *EnHSP70* amplicon using the *EnHSP70*_mC forward and reverse primer pair was digested with BamHI and 2-fold excess of NotI. Ligation and isolation of plasmids was done as described above. Validation of successful ligation was done as before.

4.3.2.4 Subcellular localisation construct

For cloning *EnHSP70* downstream of mCherry into pB2WG7 for subcellular localisation, the primer pair MC LOC forward and reverse were used to amplify mCherry from RY8686 with a SmaI site at the 5'-end and a BamHI site at the 3'-end. The *EnHSP70*_MC forward and *EnHSP70* LOC reverse primer set was used to amplify *EnHSP70* such that a BamHI site was added to the 5'-end for ligating with mCherry and a MluI site was added at the 3'-end for ligating with pB2WG7. All methods described above were used to generate the fusion product.

4.3.3 *A. tumefaciens* transformation

Clones for pB2WG7_*EnHSP7*0 were bulked by midi-prep using the Invitrogen Midi-prep as per the manufacturer's recommendations and transformed in *A. tumefaciens* strain GV3101 (with helper plasmid pMP90RK) via electroporation. First, competent *A. tumefaciens* was made by streaking GV3101 *A. tumefaciens* onto LB agar plates supplemented with 30 $\mu g.mL^{-1}$ kanamycin and 50 $\mu g.mL^{-1}$ rifampicin and allowed to grow at 28°C for 2 days. A single colony was picked and inoculated into 5 mL LB media supplemented with kanamycin and rifampicin as above and grown with continuous shaking overnight at 28°C. One millilitre of the overnight culture was then inoculated into 100 mL LB media and grown overnight at 28°C with continuous shaking. The culture was centrifuged at 2 460 *x g* for 10 mins at room temperature. The supernatant was removed, and the cell pellet was resuspended in 2 mL sterile water. The SS34 tubes were then filled with sterile water and the resuspended cells was then centrifuged as above. The water rinsing was repeated once more. The resultant water-rinsed pellet was resuspended in 10% (*v/v*) glycerol, centrifuged as before, and finally resuspended in 5 mL 10% (*v/v*) glycerol, aliquoted and stored at -80 °C. For transformation, up to 400 ng of plasmid DNA was added to 100 µL of competent *A. tumefaciens* in a 0.1 cm electroporation cuvette and kept on ice for 5 mins. The cells were then electroporated at 1.8 kV, 2 µF, and 200 Ω. The cells were then removed and placed in a clean tube to which 900 µL LB was added and incubated at 27°C for 2 hours with gentle inversion one hour into incubation. This was then centrifuged for 5 mins at 4000 rpm using a benchtop centrifuge, 900 µL of the supernatant was removed, and the pellet was resuspended in 100 µL of which 90 µL or 10 µL was plated on LB agar plates supplemented with kanamycin and rifampicin as indicated above and spectinomycin at 100 $\mu g.mL^{-1}$ for pB2WG7. Plates were then incubated at 27°C for two days. Colony PCR was performed on a selection of colonies by picking a colony and resuspending it in 50 µL nuclease-free water and incubating at 95°C for 30 mins. From this, 2 µL was used as template DNA for colony PCR. The *mCherry*::*EnHSP70* construct was transformed into Agrobacterial strain GV3101 as described above and selected as above.

4.3.4 *Arabidopsis thaliana* floral dipping and selection of transformants

For transformation of Col-0, seeds were sown onto a 3:1 soil mix containing Sphagnum peat (Jiffy-7®) and vermiculite, covered with plastic film, and stratified at 4°C in the dark for two days before being placed in a growth room at 22°C, 16 h light, 8 h dark, and 55% relative humidity. They were watered 2-3 times a week and supplemented with phostrogen at 1.4 $g.L^{-1}$ once during their growth stage and grown until the emergence of flowers. *A. tumefaciens*

harbouring the expression constructs were grown overnight in 5 mL LB supplemented with 100 µg.mL^{-1} spectinomycin, 30 µg.mL^{-1} kanamycin and 50 µg.mL^{-1} rifampicin and incubated with continuous shaking at 28°C. The 5 mL overnight culture was then inoculated into 500 mL LB media with the aforementioned antibiotics and grown once more overnight at 28°C with continuous shaking. On the day of floral dipping, the culture was centrifuged at 3 500 *x g* for 15 mins at room temperature. The pellet was resuspended in 250 mL 5% (*w/v*) sucrose and 0.05% (*v/v*) Silwet L-77 surfactant. The aerial parts of *A. thaliana* were submerged into the Agrobacterial solution for 20 seconds and laid onto a paper towel. The plants were covered with plastic wrap and left overnight in the dark. The following day, the plastic wrap was removed, and plants placed up right. Five days after the first dip, the flowers were dipped once more to increase the transformation efficiency. Following the second dip, aracons were placed around the dipped plants. Plants were allowed to resume growth and were fertilized one week after dipping. T_1 seeds were collected from these plants and screened as follows. Seeds were first left at 4°C for one week to break residual dormancy before being sown onto same soil mix as above and allowed to germinate as before. One week after germination, they were sprayed with BASTA (Beyer) at 1:200 dilution. One week thereafter, plants were sprayed once again, and the surviving plants were transplanted into individual baskets and allowed to self-fertilize. The presence of the bialophos resistance gene pB2WG7 ensures that only successfully transformed ovaries that developed into seeds are able to grow in the presence of BASTA. Once the leaves had reached approximately 1 cm in length, one leaf per plant was removed and DNA extracted using a modified CTAB method. Each leaf was gently ground using 450 µL TES buffer (0.1 M Tris pH 8.0; 0.01 M EDTA; 2% *w/v* SDS) to which 25 µL proteinase K (Invitrogen) was added and incubated at 60˚C for 1h. Following incubation, 160 µL 5 M NaCl and 70 µL 10% (*w/v*) CTAB (hexadecyltrimethylammonium bromide) was added and incubated at 65˚C for 1 h. A total of 780 µL of chloroform: isoamylalcohol (24:1) was added and samples incubated on ice for 30 min. The samples were then centrifuged at 12 074 x *g* for 10 min. The top aqueous layer was transferred to a new tube to which 345 µL isopropanol was added and centrifuged for five mins. The supernatant was discarded, and the pellet was washed twice with 70% (*v/v*) ethanol via 3 min centrifugation at the same speed as above. The pellet was airdried in a fume hood and resuspended in 50 µL 1x TE buffer to which 5 µL RNase (1 mg.mL^{-1}) was added and incubated at 37˚C for 1 h. Genomic DNA was then diluted 1:10 and PCR done on the genomic DNA using either the BAR primer set, or the qPCR primer set for *EnHSP70* indicated above. Upon T2 seed setting, seeds were collected and stored as above. Of the successfully identified transgenic lines, a selection was sown onto soil as above and allowed to germinate. One week after germination the seedlings were sprayed with 1:200 dilution of BASTA and then again one week thereafter. Upon sowing the T_2 seeds it became evident that the overexpression of *EnHSP70* had a negative effect on overall germination. To

assess the degree of retardation to germination, T_2 transgenic seeds were imbibed with 1 $mg.mL^{-1}$ gibberellic acid and sowed as before, but it had little effect in assisting with increasing germination. Owing to the delay in obtaining T_3 seeds due to retarded germination of T_2 seeds as well as results from other assays described below the study of the overexpression of *EnHSP70* in Arabidopsis was not taken further. Nevertheless, expression of *EnHSP70* was determined as described below (section 4.3.9).

4.3.5 Liquid-Liquid phase separation assay

For LLPS, *EnHSP70* was cloned into the bacterial expression vector RY8686 as described above. Successful clones were then transformed into BL21 cells as previously described. One colony was picked and grown overnight at 30°C with continuous shaking in LB media supplemented with 50 $\mu g.mL^{-1}$ kanamycin. A total of 5 mL of the overnight culture was then inoculated into 500 mL kanamycin-supplemented LB and grown with continuous shaking at 37°C until the OD_{600} measured 0.6. IPTG at a final concentration of 0.4 mM was added and the culture was induced for a further 4 hours with continuous shaking at 37°C. The culture was then centrifuged in pre-weighed tubes at 10 000 *x g* for 10 mins. Cells were lysed using a lysis buffer (50 mM Tris-HCl pH7.5, 500 mM NaCl, 4 mM β-mercaptoethanol, 2 mM PMSF supplemented with one tablet of c0mplete Mini EDTA-free protease inhibitor cocktail (Roche) per 10 mL lysis buffer) at 5 mL lysis buffer per gram of cells. The cell lysate was then centrifuged at 30 000 *x* g for 20 mins to pellet the insoluble fraction. The soluble fraction was collected and stored at -80°C until purification. The mCherry::EnHSP70 fusion protein was purified using the HIS-Pur Ni-NTA resin (Thermo Scientific). The resin was equilibrated with equilibration buffer (20 mM Na_2HPO_4, 300 mM NaCl, 10 mM imidazole, pH 7.5), washed several times with wash buffer (20 mM Na_2PO_4, 300 mM NaCl, 25 mM imidazole, pH 7.5), and the protein eluted with elution buffer (20 mM Na_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 7.5). Fractions were collected and visualised on a Coomassie-stained 10% resolving SDS-PAGE. To confirm correct sizes, an additional 10% resolving SDS-PAGE was transferred onto a nitrocellulose membrane as described in Chapter 5 methods and probed overnight using an anti-6XHIS horse radish peroxidase-conjugated primary antibody (1:4000) at 4°C. Detection was done by using equal volume of reagent 1 and 2 from the ECL Western Detection kit (Advanta) and incubating in the dark for 30 mins before being visualised. For phase separation, an initial test was done by varying the protein concentration (0-75 µM), polyethylene glycol (PEG) source (3000 to 20 000), and PEG% (0-20%) was done to determine optimal conditions for phase separation. Once establishing the optimal conditions for phase separation, the effect of increasing KCl concentration (75 – 750 mM) and addition of 1,6-hexandiol (0-10%) was investigated to elucidate which factors influenced the formation

of condensates, i.e., electrostatic, or hydrophobic respectively. For the assays, an equal volume of protein to buffer (comprising 25 mM Tris-HCl, 1 mM DTT, 10% PEG,) was prepared by pipetting the protein into the buffer several times to ensure complete mixing of the two. The composition was altered to account for inclusion of KCl and 1,6-hexandiol. For the KCl treatments the protein was diluted such that the effect of the relatively high NaCl concentration in the elution buffer could be diluted. For all assays, a total of 3 µL of this was then loaded on a microscope slide and the coverslip secured with tape. The slide was then inverted and left at room temperature for 30 minutes before being visualised under a Nikon Ti-E inverted microscope at 100x magnification using a low-fluorescence emersion oil. All images were processed using the accompanying software where the *Texas red* channel was selected for mCherry excitation. Images taken by Abigail Russels (Molecular and Cell Biology Department, University of Cape Town).

4.3.6 Subcellular localisation

For subcellular localisation, *EnHSP70* was cloned downstream of mCherry in the plant expression vector pB2WG7 as described above. The construct was then used to transfect *Ricinus communis* protoplast. All screening and preparation were done as before. A total of 2 µg of plasmid DNA was used to transfect protoplast isolated by Dr Renato Delmondez De Castro who at the time was experimenting with castor protoplast isolation and transfection. Protoplasts were transfected using the PEG-mediated transformation as described by Yoo, et al., (2007). Fluorescent microscopy was done as described above.

4.3.7 Bacterial abiotic stress assays

BL21 cells expressing the mCherry::EnHSP70 protein were used to assess whether expression of EnHSP70 in bacterial cells could confer improved tolerance towards abiotic stresses. For this, BL21 cells harbouring the mCherry::EnHSP70 construct used for phase separation and non-transformed BL21 cells were grown overnight in 5 mL LB supplemented with 50 $\mu g.mL^{-1}$ kanamycin for EnHSP70 and 50 $\mu g.mL^{-1}$ chloramphenicol for BL21 cells. The following day, 0.5 mL of the culture was subcultured into 50 mL LB and grown for 2 hrs at 37°C. One flask was then allowed to continue growing whereas the other flask was supplemented with 1 mM IPTG. All cultures (BL21 cells only, EnHSP70 induced, and uninduced) were grown to an OD_{600} of 0.4-0.6. For assays, two sets of experiments were done. One onto which the cells were spread plated and one during which the OD_{600} was measured for growth curves. For spread plating, the three cultures were diluted to OD_{600} of 0.1 and a serial dilution was then made. Of the serial dilution, a total of 100 µL of the 1:10 and 1:100 was spread plate onto LB agar supplemented with 350 mM NaCl for salt stress, 350 mM

mannitol for osmotic stress, and regular LB for controlled conditions. For heat stress, the diluted cultures were incubated at 50°C for 30 minutes then cooled at room temperature for 10 mins and plated as before. The presence or absence of growth was recorded by incubating the plates at 37°C overnight. For growth curves, 5 mL LB was inoculated with glycerol stocks of pre-induced EnHSP70 and uninduced such that the starting OD_{600} measured 0.01 and further supplemented with 350 mM NaCl or 350 mM mannitol. Heat stress was not applied to liquid cultures. OD_{600} measurements were taken every hour for 8 hours by removing 100 µL of culture from two replicates. The experiment was repeated twice with the same number of technical replicates per cell line. Significant differences were determined by taking the linear portion of the growth curve and comparing the slope between control and treatments using the Student t-test. In addition, the doubling rate of each cell line was determined by plotting the log2 of mean absorbance and using the most linear part of the curve to generate a straight line. The slope of the equation was then used to compare the doubling rate between cell lines under different treatments. Preliminary results indicated that the uninduced cells and BL21 untransformed cells had the same growth curves. As such, for subsequent replicates the BL21 cells were not used. .

4.3.8 Lactate dehydrogenase enzyme activity assay

To investigate whether EnHSP70 could enhance LDH activity under heat stress the following was done. LDH was selected as it was readily available in the laboratory and was based on the methods described in Artur, et al., (2019) who used LDH to investigate the thermotolerance of LEA proteins from *X. schlechteri*. LDH, BSA, and mCherry::EnHSP70 was diluted to a stock solutions of 500 nM in 25 mM Tris-HCl pH 7.5. For heat stress and control reactions, an equimolar ratio of LDH and protectant (EnHSP70 or BSA) was incubated at 42°C for 30 mins whereas control conditions were kept at room temperature for the same duration. For the enzymatic assay, a total of 400 nM protein was used in a reaction mix containing 25 mM Tris-HCl, 4 mM sodium pyruvate, and 2 mM NADH. Absorbance at 340 nm was measured as a decrease in NADH concentration every 5 seconds for up to 2 minutes or until the absorbance plateaued. For quantification, the linear range of the enzymatic reaction was taken, and the slope used for comparison between control and treated conditions and significance determined using Student t-test.

4.3.9 EnHSP70 expression detection in T_2 lines

Total RNA was isolated using the Tri-reagent (Sigma) as per the manufacturer's recommendations. The protein fractions were processed as described in section 2.2.2. RNA was visualized by agarose gel electrophoresis. A total of 1 µg of RNA was used for cDNA synthesis. The RNA was DNase treated using DNase I (Invitrogen) before cDNA was made using the RevertAid First Strand cDNA synthesis kit (ThermoScientific) as per the manufacturer's recommendations. For PCR, the cDNA was diluted to 1:10 and 1 µL used as template using the SAND (At2G28390.1) primers in Table 4-1 to evaluate condition of cDNA before using the *qEnHSP* and/or the *EnHSP70* primer set using the Q5 Hot Start HiFi 2X ready mix (NEB). The protein samples were then run on 10% resolving gel and transferred onto a nitrocellulose membrane as previously described. An anti-6XHis horseradish peroxidase conjugated primary antibody at 1:4000 dilution was used to probe for the presence of EnHSP70. Detection was done by using equal volume of reagent 1 and 2 from the ECL Western Detection kit (Advansta) and incubated in the dark for 30 mins before being visualised.

Results

4.4.1 Bioinformatic analysis of EnHSP70

EnHSP70 contains 484 amino acids with a theoretical isoelectric point of 5.05 and overall nett charge of -15.0 and predominantly comprised alanine, leucine, and glutamine with tyrosine and tryptophan accounting for the least (Figure 8-4). Instability index score was 35.10 which classifies EnHSP70 as stable (Gasteiger, et al., 2005). Aliphatic index score was 90.89 and the grand average of hydropathicity was -0.200 (Gasteiger, et al., 2005). MoRF prediction indicated that 6 residues at the N-terminus comprising APAIGI with an average MoRFpred probability of 0.5405 , a methionine residue towards the centre of the protein with probability of 0.504, and two sections towards the C-terminus namely VTDD and DRKS with average probability of 0.5527 and 0.5910 respectively had MoRF (Disfani, et al., 2012). Signal peptide predictions all returned low scores indicating that software programmes were not able to identify signal peptide sequences. Protein-peptide binding residues as predicted by PepBind (Zhao, et al., 2018) indicated three regions in EnHSP70 with predicted protein-peptide binding. Those being residues 354, 355, and 360 (alanine, aspartic acid, and threonine) with predicted propensity greater than 0.51 for all, residues 377 to 380 (phenylalanine, serine, threonine, and tyrosine) with predicted propensity greater than 0.77 for all, and residues 386 to 388 (glutamic acid, valine, and leucine) with predicted propensity greater than 0.92 for all. EnHSP70 was predicted to not bind CoA (Meng, et al., 2018). COACH-D modelling indicted a 0.95 C-score for ATP binding with lower C-scores assigned to phosphate, sodium, potassium, and peptide binding (less than 0.3 for all) (Wu, et al., 2018; Yang, et al., 2013). Multiple sequence alignment (Appendix section 8.2) results indicated that EnHSP70 shares sequence similarity and conserved regions across many of the other plant HSP70s including *A. thaliana*, *Z. mays*, and close relative *E. curvula*. There were regions where the majority of the selected species had amino acid sequences to which the EnHSP70 did not align. Owing to no universal annotation between the selected species and general lack of clarity differentiating HSP isoforms and cognates areas of mismatch could likely be due to these reasons. Interestingly, EnHSP70 and its close relative *E. curvula* and another grass *Digitalis exilis* have sequences to which none of the other species aligned, suggesting that the grasses, which are all C4 species, might have an amino acid composition which might be unique to the grasses. Motif prediction was done using InterPro and revealed that motifs such as an ATpase binding domain at the N-terminus and a substrate/peptide binding domain towards the C-terminus was predicted. Conserved site motif prediction indicated that between the two domains a HSP70 conserved area is found.

Online prediction tools are powerful in that they provide on the basis of algorithms a general idea of where a protein might localise. The problem, however, is that one cannot solely rely on one prediction tool. When looking at Table 4-2, the various online tools had varying results for where it predicted EnHSP70 to localise. The different tools use slightly different algorithms based on different machine learning trained in different ways. It is thus not surprising that one tool might return a result that contradicts another. Nevertheless, using multiple tools assists in the potential elimination of where it might not be. From the prediction results, it is unclear where EnHSP70 might localise but the lack of EEVD at the C-terminus confirms it is non-cytosolic.

Table 4-2: EnHSP70 subcellular location predictions

Online tool	**Predicted site**	**Likelihood (%)**	**Reference**
BUSCA	Nucleus	70.0	(Savojardo, et al., 2018)
	Other	30.0	
DeepLoc-1.0	Cytoplasm	80.0	(Armenteros, et al., 2017)
	Mitochondrion	6.0	
	Peroxisome	6.0	
	Plastid	4.0	
	Nucleus	2.0	
Localizer	Nucleus	100.0	(Sperschneider, et al., 2017)
Plant-mSubP	Cytoplasm/Nucleus	37.0	(Sahu, et al., 2019)
	Plastids	23.0	
	Nucleus	16.0	
	Cell membrane	8.9	
	Cytoplasm	6.6	
	Mitochondrion	2.2	
Plant-PLoc	Mitochondrion	100.0	(Cheng, et al., 2017)
TargetP -2.0	Other	99.0	(Armenteros, et al., 2019)
	Mitochondrion	0.0	
	Chloroplast	0.0	
LocTree3	ER	89.0	(Goldberg, et al., 2014)

EnHSP70 was fused downstream on mCherry and cloned into the plant expression vector pB2WG7 for localisation studies. Protoplasts from *Ricinus communis* were used to visualise the localisation of EnHSP70. The protoplasts were isolated by Dr Renato Delmondez De

Castro who at the time was investigating optimised conditions for protoplast isolation and transfection of castor and agreed to use the localisation construct generated herein as proof of concept for transfection efficiency. Evident in Figure 4-3 where intact protoplasts are shown in panel A and ruptured protoplasts in panel B is a strong localisation towards the chloroplasts. Lack of fluorescent signal to any other part of the cell or free-floating demonstrates that it localises to the chloroplasts. Of the prediction software used only DeepLoc-1.0 and Plant-mSupP indicated localisation to the chloroplast/plastid (Table 4-2).

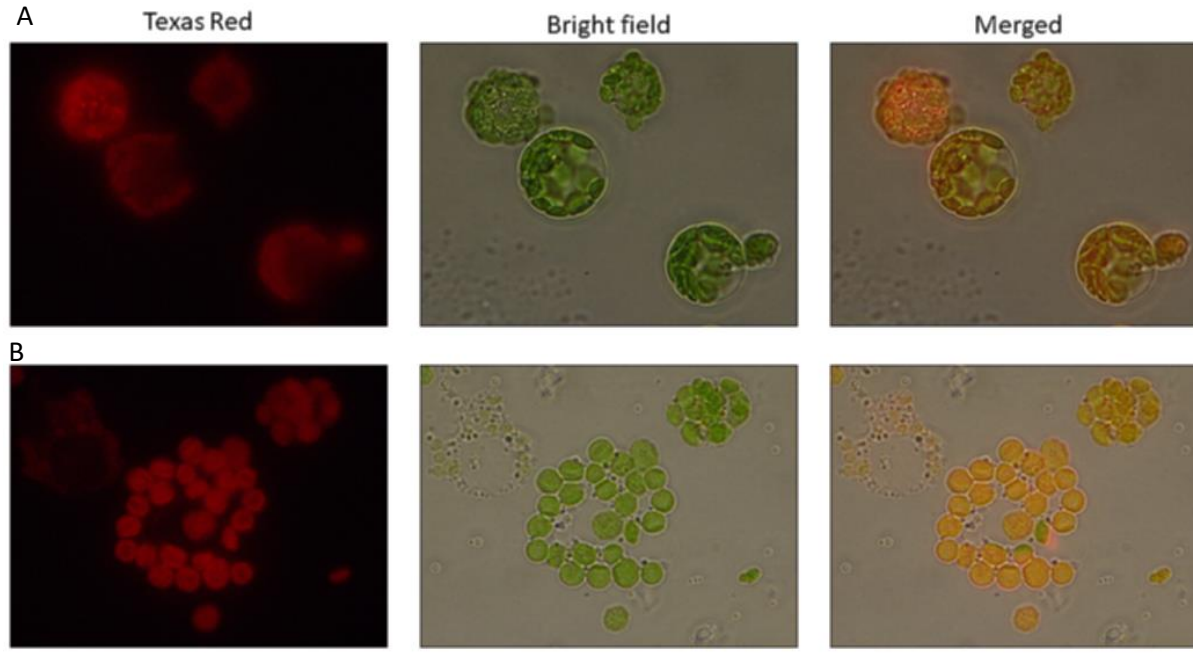


Figure 4-3: EnHSP70 localisation in castor protoplasts: Localisation of EnHSP70 in isolated Ricinus communis protoplasts. Protoplasts isolated and transfected by Dr Renato Delmondez De Castro. Magnification at 100X. Panel A shows intact protoplasts whereas panel B shows ruptured protoplasts. Scale bar 10 µM. Texas red indicated mCherry fluorescence, brightfield showing chloroplasts, and merged image showing overlap and enriched areas of mCherry fluorescence around chloroplasts.

Upon further investigation into the transcriptome and comparison to the selected sequence it was discovered that the sequence that was selected and had been used for all other studies in this Chapter, did not show up as significantly upregulated. A multiple sequence alignment was done with the sequences which were significantly upregulated and based on the percent identity matrix generated (Figure 4-4) by Clustal Omega, the studied sequence (denoted as En_hsp70) had highest similarity to gene IDs En_0110764 and En_0032451 and the least amount of similarity with gene ID En_0045201.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1: En_0097662	100.00	88.01	88.39	22.83	30.60	29.35	29.52	30.71	30.98	33.70	25.22	35.23	34.26	33.71	38.10	34.38	34.09
2: En_0099290	88.01	100.00	91.69	23.21	29.98	28.62	29.03	29.96	30.52	33.70	26.09	35.23	34.26	33.71	37.57	34.38	34.09
3: En_0077187	88.39	91.69	100.00	23.40	30.29	28.70	29.21	30.47	30.75	33.97	26.09	35.51	34.54	33.99	38.10	34.67	34.38
4: En_0045201	22.83	23.21	23.40	100.00	28.24	27.80	27.77	27.30	25.38	27.97	33.62	27.72	27.31	25.93	21.55	27.68	27.04
5: En_0078515	30.60	29.98	30.29	28.24	100.00	99.41	80.75	80.57	45.79	48.55	56.76	49.04	48.32	49.68	48.30	49.68	49.15
6: En_0082428	29.35	28.62	28.70	27.80	99.41	100.00	74.73	78.83	45.45	48.45	56.76	48.51	47.80	48.65	45.75	49.15	47.51
7: En_0033930	29.52	29.03	29.21	27.77	80.75	74.73	100.00	99.56	44.49	49.48	57.66	49.79	49.05	49.69	47.04	50.64	48.96
8: En_0071329	30.71	29.96	30.47	27.30	80.57	78.83	99.56	100.00	44.49	49.28	57.66	50.00	49.26	48.49	46.89	50.43	48.86
9: En_hsp70	30.98	30.52	30.75	25.38	45.79	45.45	44.49	44.49	100.00	61.90	56.03	65.03	64.63	65.12	68.18	65.54	64.77
10: En_0084348	33.70	33.70	33.97	27.97	48.55	48.45	49.48	49.28	61.90	100.00	84.48	88.39	87.75	90.02	88.75	90.59	89.63
11: En_0110763	25.22	26.09	26.09	33.62	56.76	56.76	57.66	57.66	56.03	84.48	100.00	87.07	87.07	87.93	-nan	92.24	89.66
12: En_0003437	35.23	35.23	35.51	27.72	49.04	48.51	49.79	50.00	65.03	88.39	87.07	100.00	99.19	93.10	92.33	93.69	92.51
13: En_0102140	34.26	34.26	34.54	27.31	48.32	47.80	49.05	49.26	64.63	87.75	87.07	99.19	100.00	92.91	92.33	93.69	92.31
14: En_0089637	33.71	33.71	33.99	25.93	49.68	48.65	49.69	48.49	65.12	90.02	87.93	93.10	92.91	100.00	92.28	95.71	92.48
15: En_0110764	38.10	37.57	38.10	21.55	48.30	45.75	47.04	46.89	68.18	88.75	-nan	92.33	92.33	92.28	100.00	98.73	98.77
16: En_0032451	34.38	34.38	34.67	27.68	49.68	49.15	50.64	50.43	65.54	90.59	92.24	93.69	93.69	95.71	98.73	100.00	98.78
17: En_0079394	34.09	34.09	34.38	27.04	49.15	47.51	48.96	48.86	64.77	89.63	89.66	92.51	92.31	92.48	98.77	98.78	100.00

Figure 4-4: EnHSP70 percent identity matrix: Matrix indicating percent sequence identity among the upregulated HSP70 amino acid sequences from the transcriptome. Highest similarity of the sequence used in this study was found with gene IDs En_0110764 and En_0032451. Percent identity matrix generated following multiple sequence alignment using Clustal Omega online sequence alignment programme set at default parameters.

Further investigation into the upregulated sequences using NCBI Blast (Table 4-3) revealed that 37.5% of the sequences were predicted to be isoform 4 with the two sequences with highest similarity to the sequence used in this study being isoform 4.

Table 4-3: Predicted isoforms of the upregulated *HSP70* sequences from the global transcriptome. Where more than one isoform is shown the order indicates degree of similarity with the first isoform being the highest

Gene ID	Predicted isoform
0110763	Partial cognate
0110764	70.4
0099290	70.8; 70.1
0084348	Cognate protein-like
0077187	70.8; 70.1
0097662	70.8; 70.1
0089637	70.4
0032451	70.4
0079394	70.4
0078515	Mitochondrial
0045201	70.15; 70.14
0082428	Mitochondrial
0003437	70.3; 70.4; 70.2
0033930	Mitochondrial
0071329	Mitochondrial
0102140	70.3; 70.4

It is also interesting to note that 4 unique sequences were identified as mitochondrial with none of the sequences annotated as chloroplastic. The relatively high degree of sequence similarity led us to decide to continue with the study.

The MFDp2 disorder prediction tool predicted an overall disorder of 12.40% with two regions of disorder at the N- and C-terminus of 16 residues and 44 residues respectively. From PDONR under the VL-XT model, 11 disordered regions were predicted with an overall disorder prediction of 32.23% (Figure 4-5). Under the VSL2 model, an overall disorder score of 18.60% was given with 9 predicted regions. VL-XT and VL3 models gave the strongest prediction for high disorder at the C-terminus. All the models predicted some disorder at the ~200-250 amino acid site which might be the linker regions between the nucleotide binding domain and substrate binding domain. Disorder enhanced phosphorylated sites were predicted to be at several tyrosine, serine, and threonine residues (Figure 4-5).

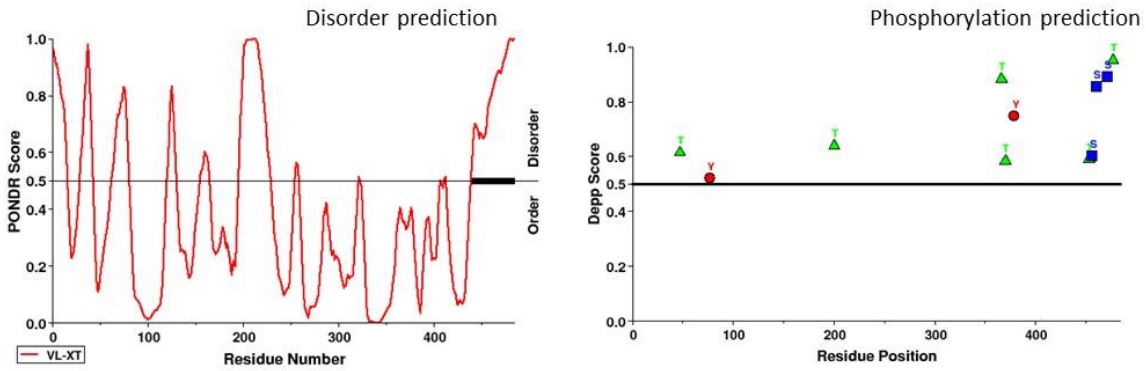


Figure 4-5: EnHSP70 disorder and phosphorylation prediction: Predicted disorder under the VL-XT model on PONDR and disorder enhanced phosphorylation prediction.

The three-dimensional structure of EnHSP70 is depicted in Figure 4-6 displayed in rainbow from N-terminus in violet to C-terminus in orange and was modelled using trRossetta. The overall structure comprises a string at the N-terminus that leads into the nucleotide binding domain (blue hues). The cyan helical region is most likely the α-helical clamp whereas the lime green section between the orange and violet domains is likely the small linker region. The region in lime green is most likely the substrate binding domain. Using the NucBind option, DNA and RNA-binding prediction was done, and evidence of some nucleic acid binding is present in Figure 4-6 predominantly at both binding domains.

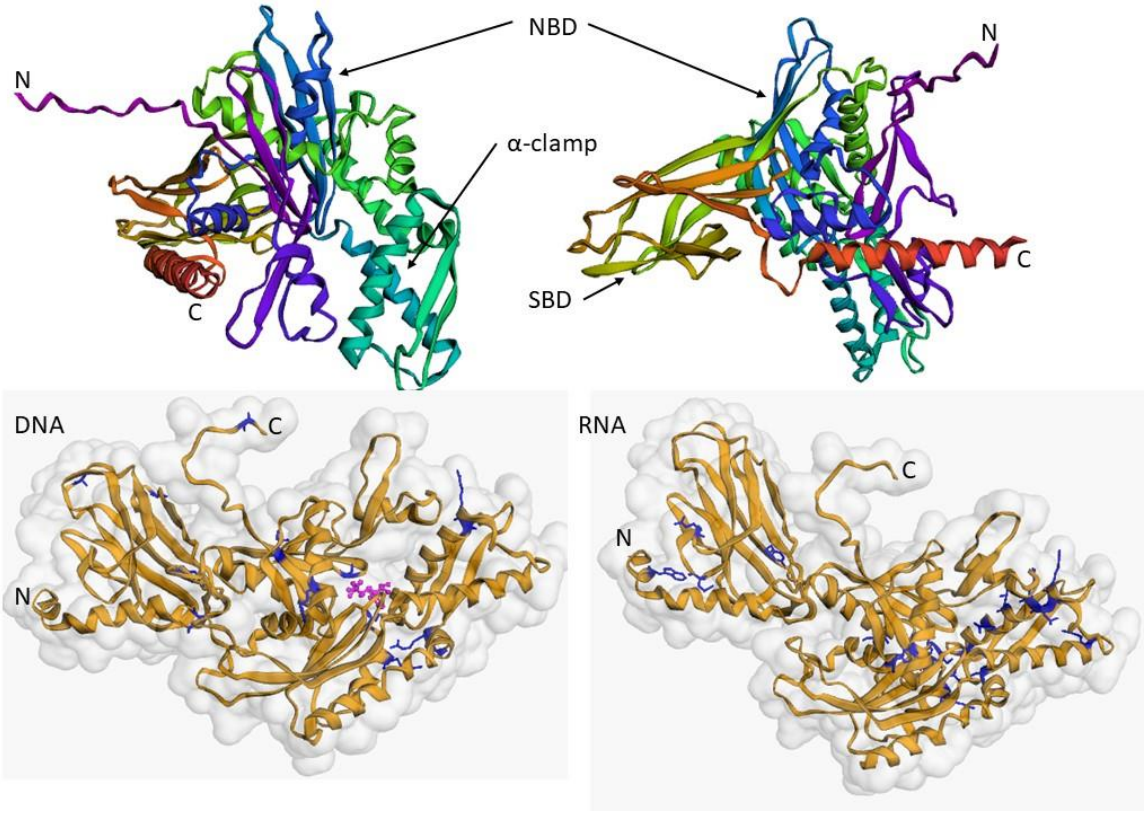


Figure 4-6: Three-dimensional model for EnHSP70. 3-D modelling of EnHSP70 using trRossetta online software (Du, et al., 2021; Wang, et al., 2022; Su, et al., 2021) with default parameters. Protein displayed in rainbow from N- to C-terminus Confidence score for predicted model is 0.906. 3-D modelling with predicted DNA and RNA binding sites (Su, et al., 2019). Consensus binding residues shown in blue.

4.4.2 Molecular cloning

Successful ligations were identified using several methods. First, colony PCR was done on a large number of colonies to isolate putative clones. Plasmid DNA was then isolated and subjected to PCR and diagnostic digests to distinguish between false positives from colony PCR. Those that had passed the digest and PCR screening were then sent for Sanger sequencing to confirm insertion. In Figure 4-7, examples of the digest strategy used to distinguish false positives from true clones are shown. For the overexpressing clones in pB2WG7 HindIII digest was effective at distinguishing the two. For the phase-separation however, the most optimal choice of enzymes unfortunately resulted in small fragments which could only be visualised by overexposing the agarose gel. Nevertheless, successful clones were identified via this three-step screening method. Results for PCR and sequencing not shown.

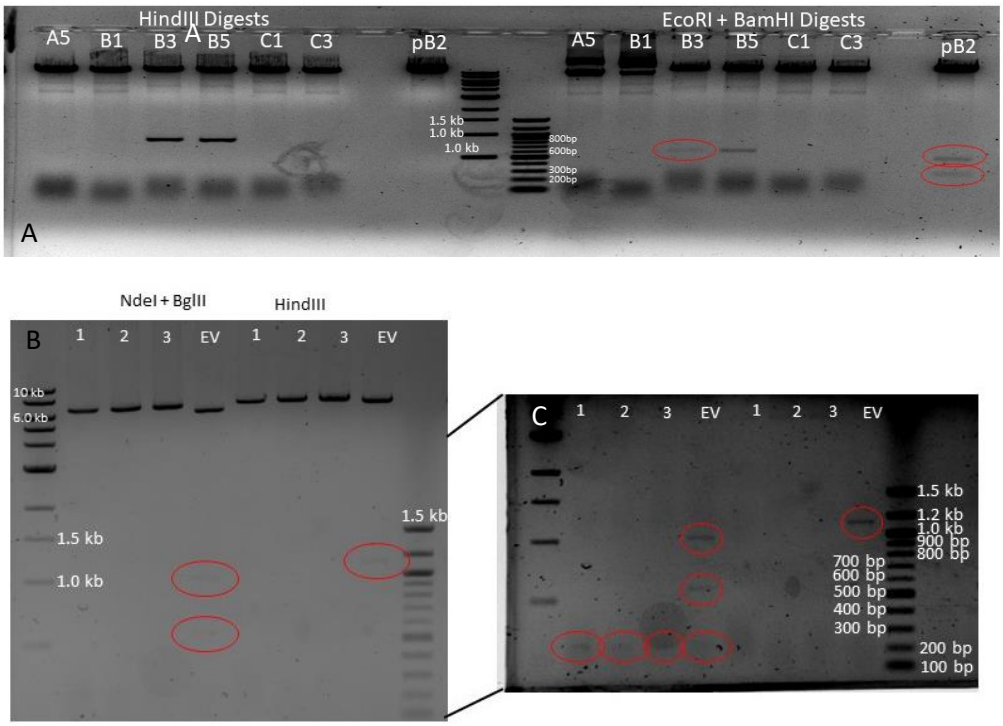


Figure 4-7: Diagnostic digests of EnHSP70 clones: (A) Digests from a selection of clones for EnHSP70 cloning into pB2WG7 using HindIII single digests and EcoR1 and BamHI double digests. Successful ligation yields two bands in HindIII digestion but a single band in empty vector (pB2). For Double digests two bands indicate successful ligation whereas three bands are seen for empty vector. (B) Digests from a selection of clones for mCherry::EnHSP70 in RY8686 using NdeI and BglII double digests and HindIII single digest. Double digests result in two bands for successful clones and three for the empty vector. HindIII single digests results in a single band observed for successful ligations and two for the empty vector. Image C indicates higher exposure of image B. For all gels, 1% agarose was used and stained with ethidium bromide. Molecular weight ladders are 1kb and 100bp from NEB

4.4.3 *Arabidopsis thaliana* transgenic screening

Arabidopsis thaliana ecotype Col-0 were transformed using the floral dip method and allowed to self-fertilize. The seeds from these plants represented the T_1 generation and were sown for selective screening by applying 1:200 BASTA. Genomic DNA from surviving *EnHSP70* overexpressing and pB2WG7 empty vectors plants which was isolated using a modified CTAB method was diluted 1:10 and used as template for the amplification of a portion of the *EnHSP70* transgene in the *EnHSP7*0 overexpressing plants using the qEnHSP70 primer set in Table 4-1 and the *BAR* PCR primer set for the pB2WG7 empty vector plants.

Figure 4-8 shows the 1% agarose gel images of a selection of transgenic plants following aforementioned PCR. A product of ~300 bp is observed in the positive control and in some of the transgenic T_1 *EnHSP70* overexpressing plants which is absent in the Col-0 WT plants. An amplicon of ~ 700 bp is observed for the positive control and in some of the pB2WG7 empty

vector plants. *EnHSP70* overexpressing plant lines 9,43,18,96, and 54 were selected for T_2 seed production whereas empty vector plant lines 30,52,66,20, and 4 were selected for T_2 seed production.

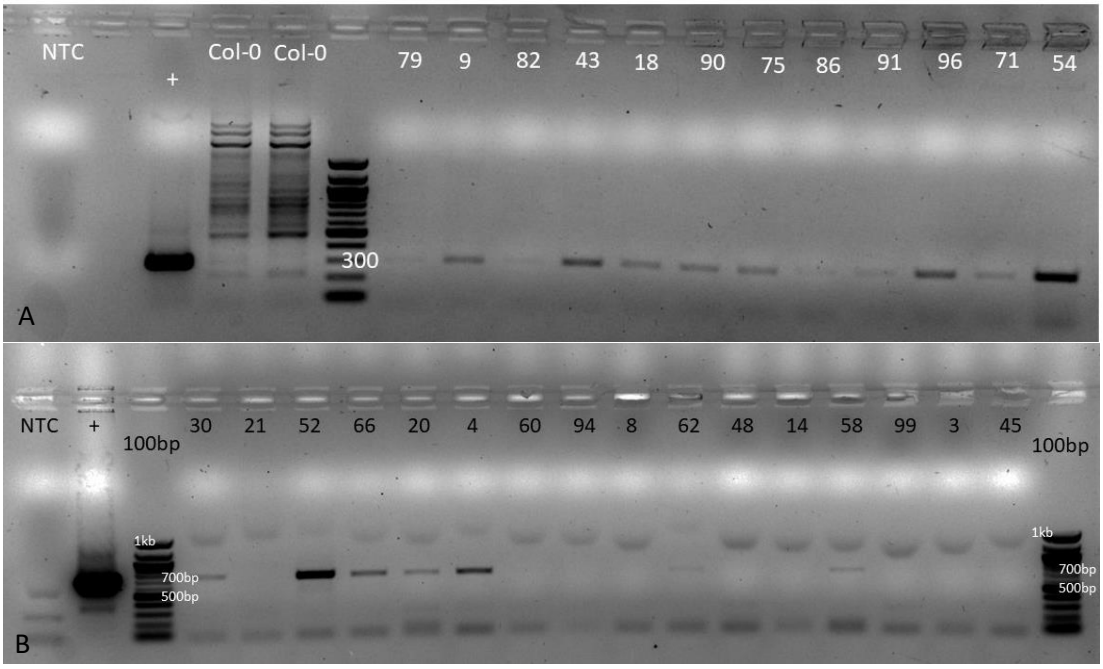


Figure 4-8: Arabidopsis thaliana transgenic screening*: Agarose gel electrophoresis images of EnHSP70 overexpressing T1 transgenic plants (A) and pB2WG7 empty vector T1 transgenic plants (B) after PCR amplification on 1:10 CTAB extracted genomic DNA using the EnHSP70 qPCR primer set (A) indicating a product of ~300 bp and BAR PCR primer set (B) indicating product of ~700 bp. Molecular weight ladder is NEB 100 bp.*

During the initial stages of germination of T_2 plants, a delay in germination phenotype was observed for the overexpressing lines which was absent from the EV plants. Germination efficiency of T_2 was on average 44% between the selected T_2 plants and treatment with gibberellic acid did not improve the efficiency. In addition, seeds germinated 2-3 weeks post-sowing and post-screening which suggested that the overexpression of *EnHSP70* had some effect on germination since this was not noted with the empty vector T_2 plants. Nevertheless, DNA, RNA, and protein was isolated from surviving T_2 *EnHSP70* overexpressing and empty vector plants. DNA was used for BAR PCR screening, RNA was used to convert to cDNA which was then used in an end-point RT PCR to evaluate expression of the transgene, and protein was used to determine whether translation of *EnHSP70* was occurring. illustrates the BAR PCR screening of T_2 *EnHSP70* overexpressing lines and pB2WG7 empty vector plants showing isolation of putative T_2 transgenic plants for both following BASTA treatment. All plant lines shown below had survived initial and subsequent BASTA selection but only some i.e., 54.2, B1.1, B1.3 for *EnHSP70* overexpressing lines and 4.4, 66.2, and 52.2 for empty vector

plants had the BASTA resistance gene. Figure 4-10 shows an RNA electrophoresis (A) used for cDNA synthesis, SAND PCR on synthesised cDNA (B), and RT PCR using the qEnHSP70 primer set (C). A band in the SAND PCR of approximately 246 bp is indicative of viable cDNA. A band of approximately 300 bp in the qEnHSP70 was noted for lines B1.1, B1.2, B1.3, 18.6, and 54.4, however the relative intensity of the bands is suggestive of low levels of expression.

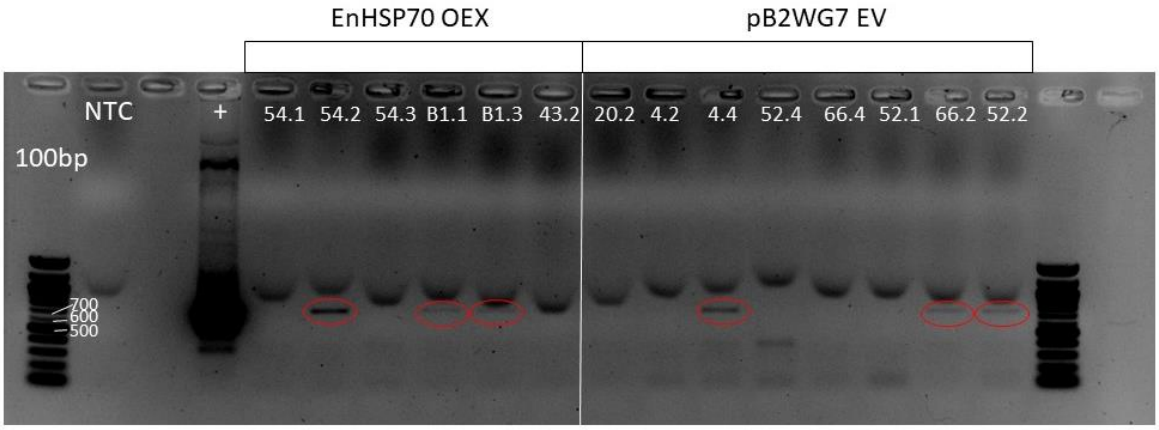


Figure 4-9: T_2 PCR Screening: Agarose gel indicating selection of T_2 transgenic EnHSP70 overexpressing lines and pB2WG7 empty vector plants using the BAR PCR primer set. Product of ~700 bp indicates transgenic plant line. Molecular weight ladder is NEB 100 bp.

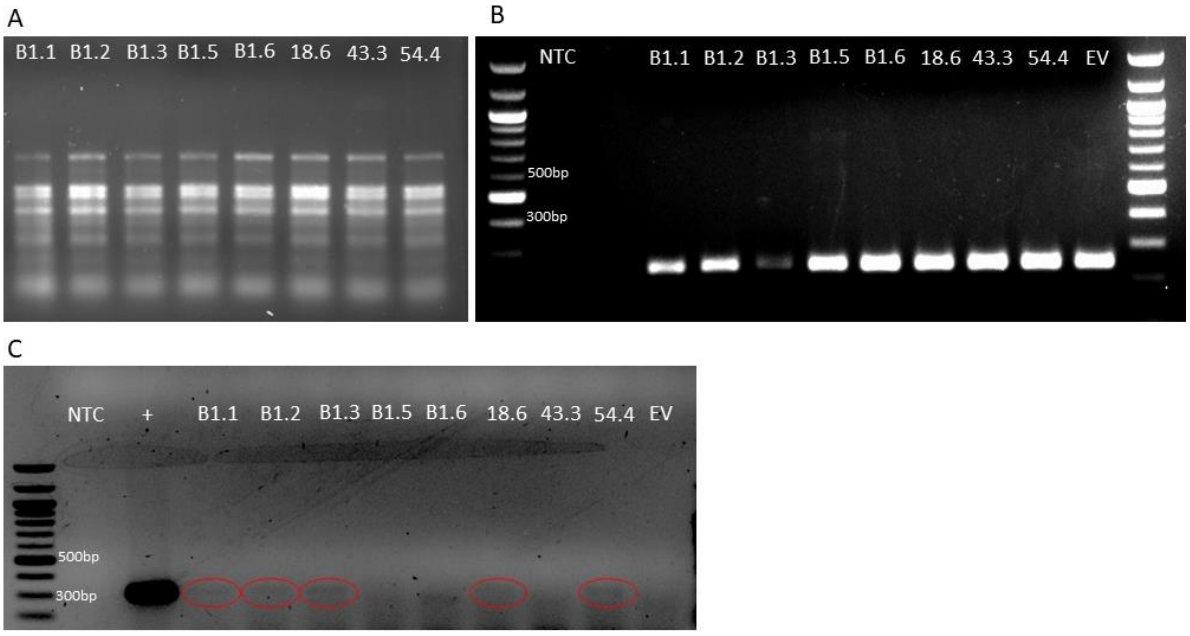


Figure 4-10: EnHSP70 expression in T_2 plants: Agarose gel showing RNA (A) extracted from T_2 plants using the Tri-reagent (Sigma) which was subsequently used for cDNA synthesis. Viability of cDNA was checked using the SAND PCR primer sets (B) indicating a band of roughly 246 bp. Image C shows the amplification of a portion of the EnHSP70 transgene in T_2 overexpressing plants. Molecular weight ladder is NEB 100 bp.

4.4.4 mCherry::EnHSP70 expression and purification

EnHSP70 which had been cloned downstream of mCherry in the bacterial expression vector RY8686 was transformed into BL21 *E. coli* cells and expression induced in 500 mL LB culture. Purification of mCherry::EnHSP70 (Figure 4-11) did not prove to be problematic other than the appearance of non-specific bands during elution fractions. Upon detection using an anti-6XHis horse radish peroxidase conjugated primary antibody, it was revealed that the band at ~34 kDa was in fact 6XHis-tagged mCherry. The largest band above the 72 kDa marker is the fusion protein at a predicted MW of 83.2 kDa. The large band just below that is presumed to be the bacterial DnaK (HSP70) which is a common contaminant. The band at ~60 kDa is presumed to be EnHSP70 lacking the mCherry fusion. The various bands observed from the purification could have arisen possibly by the linker region between mCherry and Enhsp70 undergoing some cleavage during the preparation of the samples before loading onto the protein gel. It is plausible that the high temperature and presence of β-mercaptoethanol in the sample application buffer may have inadvertently cleaved this region. The linker region (GAPGSAGSAAGGS) contains two serine residues shortly after the start site of *EnHSP70*. The reduction of the thiol-bonds could have weakened the integrity of the linker region. Nevertheless, literature implies that mCherry is unable to undergo phase separation so the presence of 'free' mCherry following denaturing conditions was not a concern. A three-dimensional modal of this fusion protein can be seen in Figure 8-5.

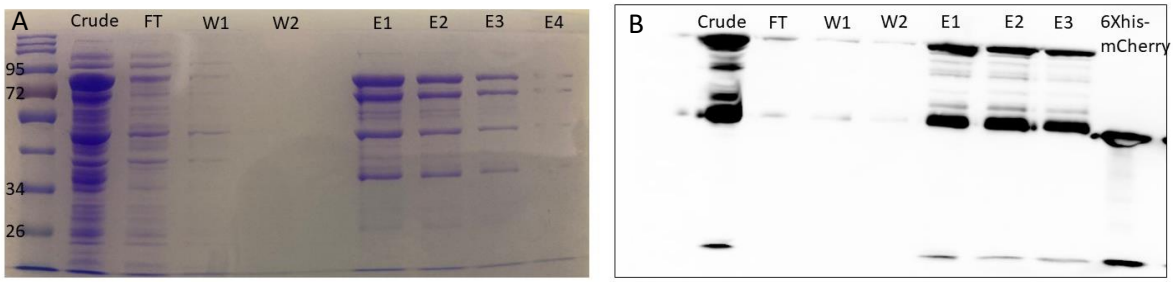


Figure 4-11: mCherry::EnHSP70 purification: (A) SDS PAGE image of mCherry::EnHSP70 fusion protein purification using HIS-PUR Ni-NTA resin indicating various fractions and (B) anti-6XHis horse radish peroxidase-conjugated primary antibody western blot following ECL detection. ECL detection was able to distinguish non-6XHis-tagged proteins seen on Coomassie-stained gel.

4.4.5 EnHSP70 liquid-liquid phase separation

Initial investigations revealed that EnHSP70 was able to phase separate under 10% and 15% PEG 4 000 where at 10% PEG 4 000 condensates formed only at 25 µM protein. Increasing the protein concentration resulted in the formation of condensates but they appeared to be aggregates and lacked a spherical shape often associated with LLPS condensates. At 15% PEG 4 000, the concentration of the protein did not affect the propensity of EnHSP70 to phase separate. In a follow-up experiment, PEG 3 000 and PEG 20 000 with varying protein concentration was investigated. To this end as shown in Figure 4-12, PEG 20 000 yielded greater condensate formation by having a condensate fraction of approximately 0.07 at 66.5 µM protein whereas PEG 3 000 at the same protein concentration had a condensate fraction of approximately 0.06. This further corroborated early investigations that phase separation is protein concentration and PEG percentage dependent with greater condensate formation at higher protein concentrations. For subsequent evaluation of protein interactions as either inter-static or hydrophobic, 33 µM of protein and PEG 20 000 was used. To investigate what effect salt concentration has on the formation of condensates, a KCl gradient was done. Shown in Figure 4-12 is a decrease in condensate formation with increasing KCl concentration along with a rapid decline in condensate fractions from 0.04 to below 0.01 at <500 mM KCl. Owing to the relatively high NaCl concentration in the elution buffer used, the protein was diluted such that the effect of increasing KCl could be investigated. To investigate whether condensate formation was driven by hydrophobic interactions, 1,6-hexandiol was added as it is known to disrupt condensate formation through interfering with weak hydrophobic bonds (Düster, et al., 2021). Inclusion of 1,6-hexandiol did not appear to have an effect on the condensate formation or shape of the condensates at either percentage. Condensate fractions decreased slightly at 5% 1,6-hexandiol addition but at 10% it was comparable to 0%. Phase separation propensity of EnHSP70 appears to thus be primarily driven by protein concentration, increasing molecular crowding, and KCl.

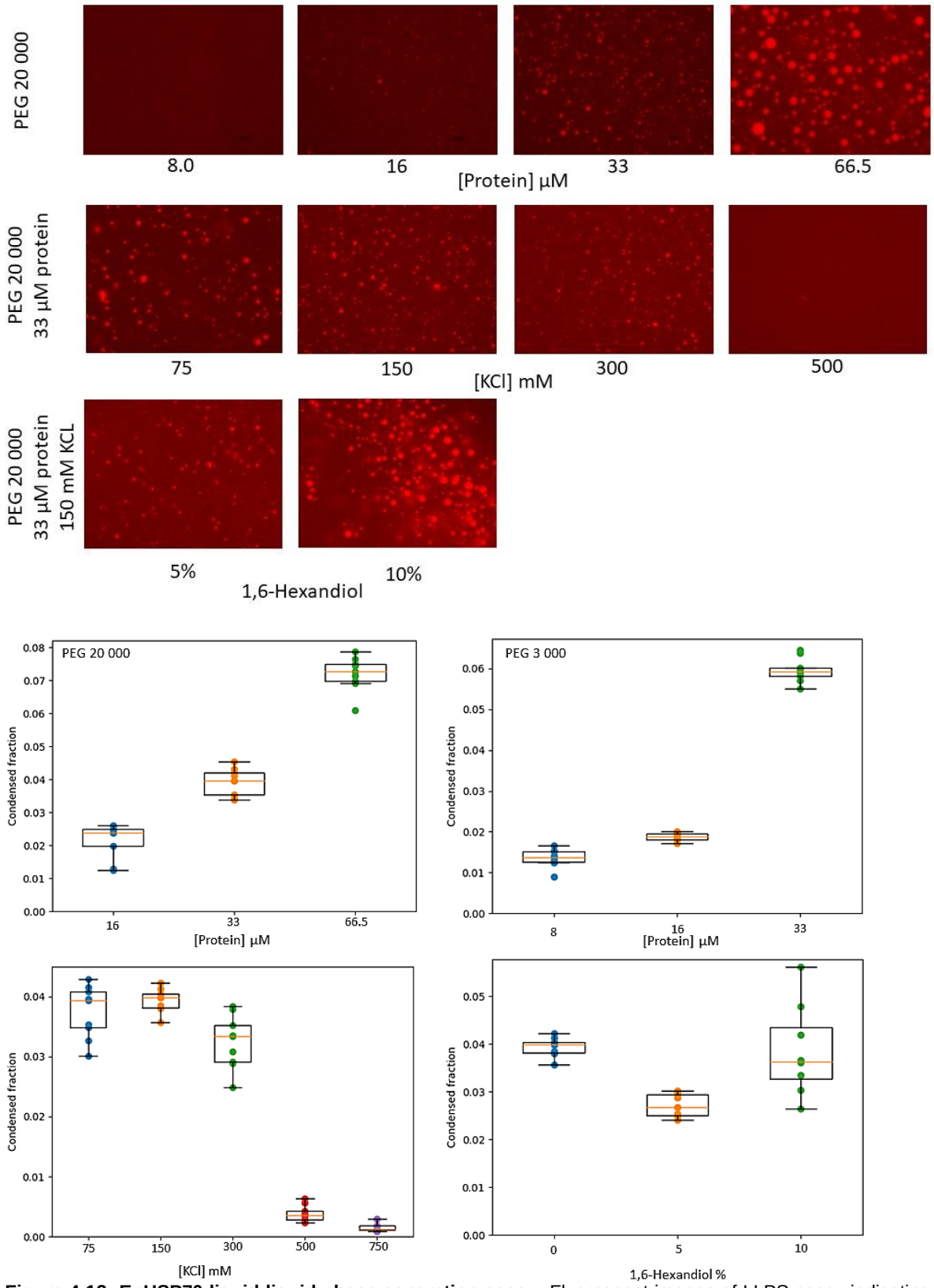


Figure 4-12: EnHSP70 liquid-liquid phase separation assay: Fluorescent images of LLPS assay indicating the effect of increasing protein concentration at PEG 20 000, increasing [KCl] using 33 µM protein, and addition of 1,6-hexandiol at 150 mM KCl using 33 µM protein. All images taken at 100x magnification and scale bars at 10 µm. Graphs indicate the condensate fraction for various treatments where 10 images per treatment were used with a standard deviation of 2. Graphs generated from Python code sources from https://github.com/leetheflee/droplet-analysis. Images taken by Abigail Russel. Scale bar 10 µM.

4.4.6 Bacterial abiotic stress assay

BL21 cells expressing mCherry::EnHSP70 were used to investigate whether EnHSP70 could impart any protective properties when cells were exposed to stress. Stressed conditions tested included high salt (350 mM), high osmotic stress (350 mM), and heat stress (50°C) and were initially determined by absence and presence of growth between induced, uninduced, and BL21 cells following overnight incubation at 37°C on LB media supplemented with respective stress. This data (not shown) did not yield any conclusive results as it represented the end point of growth and all cells irrespective of treatment grew well. Spread plating however can be misleading as it represents the end point of growth and colonies forming may not necessarily be in direct contact with the agar, in the case of salt or mannitol-supplemented media and could therefore be 'shielded' by cells which are in contact with the media. In addition, the rate of diffusion between the agar and cells might have been affected by the percentage of agar used in the media. As such, a growth curve was done for salt and mannitol stress only on BL21 expressing EnHSP70. Expression was pre-induced using 1 mM IPTG and cultures were sub-cultured into 5 mL LB such that the starting OD_{600} was at 0.01. Cultures were then grown for 8 hours and sampled every hour.

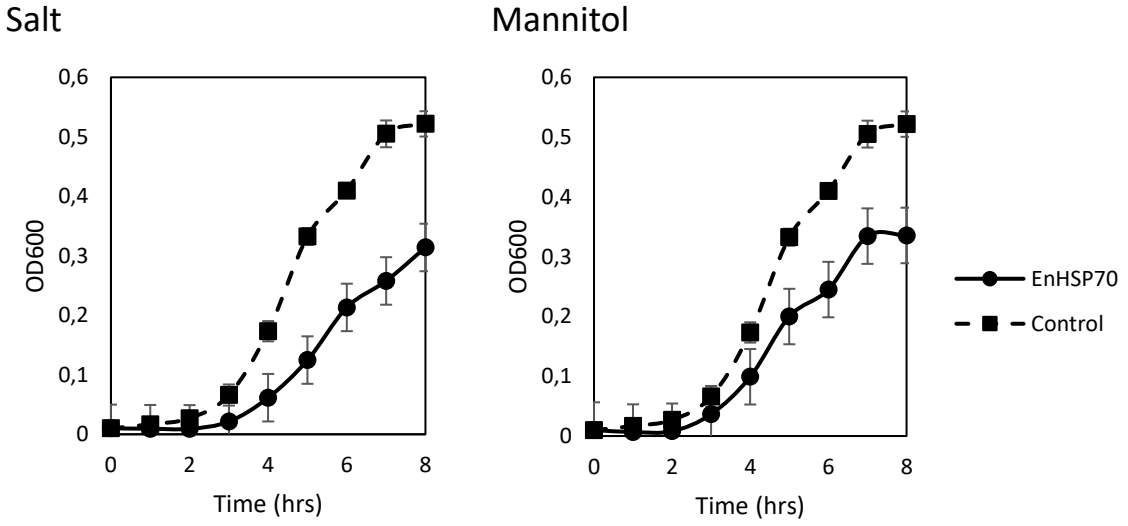


Figure 4-13: Bacterial abiotic stress assay: 8 hour growth curves for pre-induced EnHSP70-expressing BL21 cells (circles) in 350 mM salt and 350 mM mannitol with square dashed line showing growth of pre-induced EnHSP70-expressing BL21 cells under unstressed conditions.

Under both salt and mannitol stress, shown in Figure 4-13, the expression of EnHSP70 did not enhance the tolerance towards these stresses as the overall growth of the EnHSP70 expressing cells was much slower than the control. These results seem to suggest that the

selected EnHSP70 may not be involved in stress response and has some negative effect on overall establishment of growth since those with pre-induced EnHSP70 took marginally longer to reach exponential growth phase than those which were uninduced.

4.4.7 Lactate dehydrogenase assay

Following incubation of LDH with EnHSP70 in equimolar ratios and then subjecting the protein mixes to 42°C heat stress and subsequent 340 nm measurement of depletion of NADH, EnHSP70 did not appear to enhance the thermotolerance of LDH compared to LDH alone (Figure 4-14). Under heat stress conditions LDH activity was comparable to control conditions while EnHSP70 presence resulted in an overall decrease in percent enzyme activity ($p<0.05$).

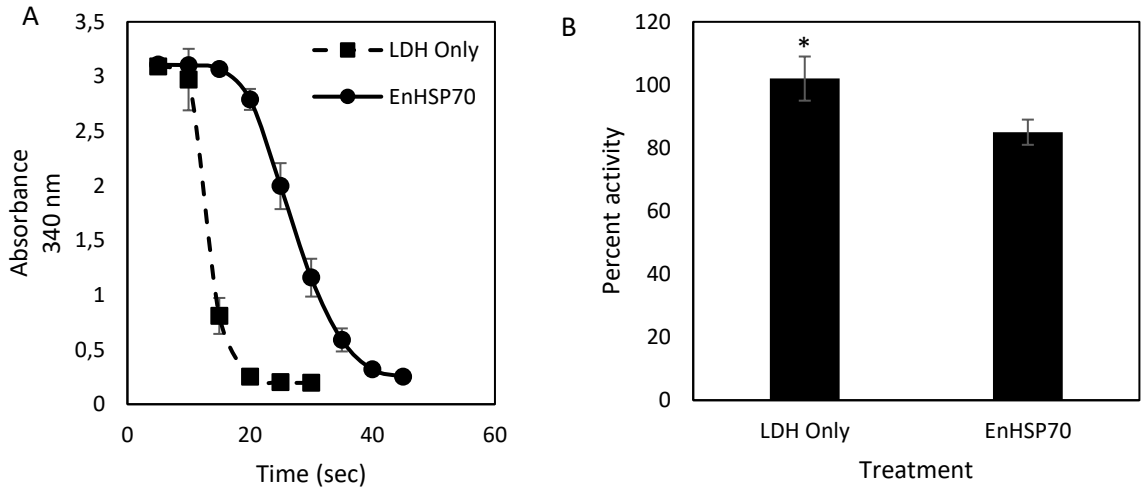


Figure 4-14: Lactate dehydrogenase assay: Enzymatic curves for LDH activity following heat stress at 42°c for 20 minutes. Panel A indicates the absorbance at 340 nm as depletion of NADH through LDH activity where square line is LDH alone and circle is LDH with EnHSP70. Panel B shows percent activity as determined from linear range of enzyme activity by dividing the experimental conditions' slope with the slope of the control conditions. Significance determined via student t-test with $p<0.05$ with N=3.

Discussion

The transcriptome (Madden, 2019) indicated that one of the most upregulated genes was HSP70. Given this and the fact that it is a universal stress-associated protein with high phylogenetic conservation, this Chapter sought to characterise some aspects of an HSP70 from *E. nindensis*.

The multiple sequence alignment revealed a high degree of sequence conservation with areas throughout the sequence aligning to all the species selected. Those areas are the N-terminal NBD and the C-terminal SBD. High variation in sequence composition was noted between these domains and is presumably the variable linker region. There was one particular region to which EnHSP70 only aligned to the other two grass species used in the alignment and comprised of 6 amino acids (DRGEDI). Why this might be case could be an indication of a grass specific amino acid composition at that site. Nevertheless, the sequence alignment demonstrated a high degree of conservation and further exemplifies the universal conservation of HSP70. Further analysis of the selected sequence identified the selected sequence as an isoform 4 owing to highest sequence similarity to the Arabidopsis homologue.

Localisation predicting software failed to give a conclusive answer as to where it is predicted to localise which emphasises the need for more robust predicting software. Through collaborative work with Dr Renato Delmondez De Castro it was determined that this particular EnHSP70 predominantly localised to the chloroplasts in castor protoplasts which raises a few interesting questions. In hindsight knowing that the selected sequence was not differentially up-regulated from the transcriptome, it is only plausible to speculate that given its location near the chloroplasts it might simply be performing a general housekeeping chaperone role in maintaining proteostasis during thylakoid disassembly. It is also valid to speculate that its absence in the transcriptome might simply also be due to thresholds of what was deemed significant. Additional real-time qPCR analysis of the different HSP70 forms, including the sequence used herein is warranted to justify any conclusions but was beyond the scope of the work presented herein.

The discovery that the selected *EnHSP70* sequence, while present in the RNAseq FASTA file, was not significantly upregulated presented a unique problem in that the majority of the work presented herein had already been initiated and there would have been no additional time allotted to study the upregulated sequences in as much detail as presented here. The relatively high sequence similarity albeit it only in the 60% range, indicated that the studied protein had highly conserved areas between the two upregulated proteins and that it is most likely another gene copy of isoform 4. Eukaryotes have multiple *HSP70*s compared to prokaryotes. The

higher number of *HSP70*s is mostly attributed to genome duplication events throughout evolutionary history. Resurrection plants with the exception of a few, are well characterised as having polyploidy and it has been suggested that genome duplication events in association with rewiring of networks involved in seed-desiccation pathways led to some species such as the desiccation tolerant *Lindernia brevidens* acquiring vegetative desiccation tolerance compared to its close relative *L. subracemosa* (VanBuren, et al., 2018). The genome duplication and ploidy nature of *E. nindensis* (Pardo, et al., 2020) presents an argument to be made of a much greater redundancy system compared to desiccation sensitive relatives.

Three-dimensional modelling indicated that the N-terminus was disordered. This was further corroborated with disorder prediction software that indicated that there was some degree of disorder. Other prediction tools used herein demonstrated that there was some nucleic acid binding which was validated as ATP through protein-ligand predictions. MoRF prediction indicated that there were regions within EnHSP70 which are uniquely intrinsically disordered. Disordered regions are generally considered to be drivers of liquid-liquid phase separation owing to their ability to not adopting stable helices or pleats. To date, only one other study has demonstrated that an HSP70 was able to undergo phase separation in a PEG percentage and protein concentration manner (Li, et al., 2022c). Given the general role that HSP70 performs in maintaining proteostasis it is suggestive that it might play an important role in either establishing condensates or facilitating their formation. Initial tests of LLPS indicated that EnHSP70 was able to undergo phase separation in a PEG percentage and protein concentration manner, though at higher PEG percentages the concentration of protein was negligible. However, the shape of the condensates was not completely spherical and appeared to be aggregating. To further investigate this, PEG 3 000 and PEG 20 000 was used in a protein concentration manner and illustrated that the PEG 20 000 facilitated condensate formation better as evident in Figure 4-12 both visually and quantitatively. In addition, the more protein there was the greater the formation of condensates. PEG is used as a molecular crowder in phase separation assays and in the field of artificial drought, can be used to simulate a water-deficit state. As such, and in the context of this thesis, PEG can be likened to a cell undergoing dehydration. The increase in protein concentration used for the assays can be likened to increased translation within the cell. Though naturally cells would not accumulate protein to such a high concentration but as an *in vitro* system it works well. The LLPS assay demonstrated that the formation of condensates is highly affected by salt concentration where greater KCl concentrations resulted in a decline in condensate formation. This observation in relation to desiccation response in *E. nindensis* suggests that its primary purpose could be in facilitating condensate formation to remove chloroplastic-derived proteins but may diminish towards lower RWC states as the cellular salt concentration increases.

Although there is no evidence for the changes in salt concentration in *E. nindensis* during dehydration, it is plausible to speculate that as water becomes limiting the concentration of salts might increase locally. When coupled with the disassembly of the thylakoids and the observation that at 40% RWC commitment to complete breakdown is made, it is likely that this particular HSP70 is only present during the early to mid-stages of dehydration. Further testing on the binding capacity of EnHSP70 to chloroplastic proteins would elucidate whether this proposed role is valid. This Chapter demonstrated for the first time that an HSP70 from a resurrection plant is able to undergo phase separation in a PEG and protein manner which can be likened to both increased molecular crowding during dehydration and increased translation respectively. The discovery that EnHSP70 is able to phase separate has great implications to our understanding of stress granule formation in plant cells and could lay the foundation for further exploration into these transient membraneless compartments.

Selection of T_1 plants using the herbicide resistance gene found within the T-DNA region was successful enough to identify a number of putative transgenics. Selection at the T_2 stage however presented a problem. Upon sowing the T_2 seeds it became apparent that germination was greatly impacted. For instance, after sowing 50 T_2 seeds, only half would germinate which complicates the selection of transgenics since the expected proportion of wild-type plants is a quarter with heterozygous and homozygous transgenics accounting for half and a quarter respectively. This was further validated when the second lot of seeds were sown and again showed the same trend. This led us to speculate that perhaps overexpression of *EnHSP70* might be causing suppression of germination or delaying it. HSPs are quite common in seeds in the desiccated state and is one of the characteristic signatures of vegetative desiccation tolerance. During the seed collection, storage, initial sowing, seeds are kept at 4°C. Ashraf, et al., (2021) demonstrated that an HSP70-16 and a voltage-dependent anion channel jointly facilitated the suppression of Arabidopsis seed germination at 4°C. In the crustacean *Daphnia magna DmHSP70-A*'s expression was shown to be higher leading up to and during diapause (Chen, et al., 2021) and has been hypothesised that some HSP70s play a role in developmental suppression during diapause (Darlinger, et al., 2001). Overexpression of transgenes inherently imposes an unnatural state under normal conditions and raises an interesting debate on whether transgene investigations should be done under overexpression or inducible conditions. Thus, what could have potentially led to the stunted germination in the T_2 lines is a combination of cold-induced suppression of germination which was not alleviated through gibberellic acid treatment and an arrest in development owing to possible role of *EnHSP70* overexpression. While expression was detected in the T2 lines, the RT PCR suggested that in the leaves the expression was not as great as the housekeeping SAND gene. Owing to limited supply of T_2 seeds it was not possible to validate if the expression in

seeds were at the same baseline level or higher but sufficed to say, the expression was not what was expected. Anecdotally, the successfully confirmed T_2 plants had much greener leaves during the drying of plants for seed setting compared to the empty vector plants and also took marginally longer to reach a dried state.

The cauliflower mosaic virus promoter is routinely used to drive constitutive expression of transgenes. The expression vector used herein is meant to be used with the Gateway Cloning technology but instead was used as expression vector for cloning the *EnHSP70* using traditional restriction enzymes. There is roughly a 500 bp region that separates the 35S promoter from *EnHSP70* and it is likely that this may have caused decreased expression in the leaves. This is further supported by the observation that by using up to 30 µg of protein for western blot detection, no transgenic protein could be detected when probed with the anti-6X His antibody. Transient expression in tobacco leaves which were vacuum infiltrated with Agrobacteria harbouring the overexpression construct, indicated that the protein was able to be expressed when probed on a dot blot using the anti-6X His primary antibody (Figure 8-6), demonstrating that the protein is capable of transient expression. The low level of expression noted in T_2 plants and clear impact on germination lead us to conclude that pursuing stable evaluation of *EnHSP70* expression would not be feasible. This was further corroborated by the bacterial assays and LDH assays discussed next.

The LDH assay results indicated that when LDH was incubated with EnHSP70, it did not provide any enhanced thermotolerance to the overall activity of LDH when compared to LDH only. This results disproved the hypothesis that it could provide enhanced protection against heat stress. In fact, BSA outperformed EnHSP70 in these assays (data not shown). Even though bioinformatic analysis indicated that EnHSP70 is thermostable, it was not able to confer any thermostable protection to LDH. The bacterial assay results indicated that when EnHSP70 is constitutively expressed in BL21 cells, the growth is negatively impacted when compared to control. This in and of itself is an interesting observation and could likely be due to several factors including but not limited to over accumulation of foreign eukaryotic proteins interfering with regular stress response. The hypothesis was that expression of EnHSP70 in *E. coli* would enhance tolerance to stress relative to control conditions was thus disproved and provided supported evidence for its overexpression effect on cellular growth. In fact, during induction of the protein when compared to uninduced cultures, the EnHSP70 expressing cultures took much longer to reach an OD_{600} equivalent of log phase. While this may be deemed as a negative result, it simply points to a different potential functional role of this selected HSP70. Nomenclature would suggest that HSPs are induced under heat, but as alluded to in the introduction, their induction can be caused by various stresses. In addition, there are several HSPs that perform what one might deem as ‘housekeeping’ roles and are

not stress-induced. The bacterial assay and the retarded germination of T_2 Arabidopsis plants seems to suggest that the selected HSP70 is not a stress-inducible HSP70 but could rather be involved in housekeeping. When taken into consideration with where it localises this becomes much more plausible. The disassembly of the thylakoids and release of chloroplastic proteins is a controlled and regulated process which requires the assistance of a chaperone to remove these proteins to ensure general housekeeping during dehydration. However, an equally valid point would be that HSP70 does not function in solitude and requires a host of co-factors. It is plausible to speculate that the observed lack of improved thermotolerance may well be due a limiting number of co-factors relative to the overabundance of EnHSP70.

Conclusion

This Chapter set out to explore some functional aspects of a heat shock protein 70 from *E. nindensis*. It set out to overexpress it in the model organism *Arabidopsis thaliana* and compare any fitness gains against the empty vector controls. To this end it was established that overexpression of EnHSP70 in Arabidopsis led to a delayed or retarded germination which impacted the evaluation of its overexpression. Three-dimensional modelling indicated that it had a disordered N-terminus. The appearance of the disordered region led us to speculate whether it might play a role in LLPS as this could have great impact on the formation and/or stabilisation of stress granules or in the establishment of membraneless organelles in which certain biochemical processes could occur. The observation that it is able to undergo LLPS illustrates that it is able to form discreet condensates which when taken with postulates of NaDES formation during dehydration could be an important requirement for desiccation response in resurrection plants. Its localisation to the chloroplast could have great impacts on drought sensitive plants which typically do not switch photosynthesis off upon water-deficit stress. In *E. nindensis* the appearance at the chloroplasts could be in favour of chaperoning degraded proteins from there to ubiquitination and proteolysis sites or as proposed in Chapter 3 could also be in favour of protecting certain thylakoid reassembly proteins in non-senescent leaves. In the scope of genetic improvement, a range of other studies have demonstrated that overexpression of HSPs in general can have great fitness gains in enhancing tolerance towards an array of abiotic stresses which further corroborates the evolutionary importance of HSPs. The selected protein used herein could not confer thermotolerance to bacterial cells nor enhance the activity of lactate dehydrogenase which is highly suggestive that it is not a stress-inducible protein but may likely be involved maintaining proteostasis during early to mid-stages of desiccation in *E. nindensis*. Further exploration of the up-regulated EnHSP70s is required to elucidate what their role in desiccation tolerance might be.

Chapter 5: Developing a method for generating transgenic *Eragrostis tef.*

Introduction

With the rapid advancement of biotechnology over the past two-to-three decades, scientists have been able to accurately alter or manipulate the expression of certain genes to ensure a desired phenotype. This Chapter will focus on developing a method for transgene expression in the Ethiopian orphan crop *E. tef*. To this end, a method of successful transgene delivery into *E. tef* will be investigated with particular emphasis on establishing a somatic embryogenesis callus induction and plant regeneration method.

5.1.1 Transgenic plants

A transgenic plant is conventionally made via the insertion of foreign gene(s). The type of transgenic depends on what is to be done with the plant. Transient expression allows for rapid expression albeit limited by the capacity of the vector in question to express the protein of interest. Transient expression is particularly useful in the expression of recombinant proteins whereby the plant is used to assemble these proteins for extraction. One particularly useful application of transient transgene expression in plants is the production of biopharmaceuticals such as candidate vaccines such as for example the expression SARS-CoV-2 antigen in *Nicotiana benthamiana* (Ruocco & Strasser, 2022). Here viral vectors are often employed and designed in such a way to facilitate over-production of the recombinant protein and are then later harvested (Regnard, et al., 2010). Transient expression is also often employed for elucidating the subcellular localisation of target proteins. Transient plants do not necessarily have the foreign gene(s) incorporated into their genome and are thus not able to impart the transgenic phenotype onto their offspring. Stable transgenics however, rely on the complete integration of transgene into the plant's genome. These plants are developed for their long term study of transgenes and are often developed with the notion of changing or altering a particular phenotype to suit a particular objective. For instance, stable transgenic food crops are generated to have phenotypes such as disease or drought resistance by stable transgene integration into the genome such as the *Bacillus thuringiensis* Cry proteins which were integrated into *Zea mays* creating the so-called Bt corn. The creation of stable transgenics is a laborious process involving Mendelian genetics and various methods of verifying insertion at each generation. For short lifecycle plants such as *Arabidopsis thaliana*, this is a relatively quick process, theoretically requiring approximately 9 months from infiltration to T3 generation. For longer lifecycle plants, generating stable transgenics can take years. An immediate

deterrent with generating stable transgenics is that the phenotypic characterisation of successful transgene insertion can only be truly evaluated when homozygous transgenics are present. Only then can the hypothesis be tested against its wildtype control. In an effort to streamline the process of generating stable transgenics, the initial screening and testing of suitable vectors transiently is a cost effective route that will assist in evaluating whether the stable line is likely to be feasible.

Successful transgenic plant generation can be divided into the following processes: firstly, selection of most suitable plant transformation vector; selection of suitable agrobacterium strain; and lastly, selection of most appropriate method of vector delivery. These processes will briefly be considered below.

5.1.2 Selection of plant expression vector

As the first processes in generating a transgenic plant (stably or transiently), the selection of the vector that will introduce the transgene into the plant is an important one. Over the past few decades, an array of commercially available vectors has been developed for a variety of applications. The selection of the vector in use is often an arbitrary choice, often limited by the availability within a particular research group, but in general they should be selected on the basis of the following criteria: they should have areas into which the gene(s) of interest can be cloned such as a multiple cloning site; they should be able to be shuttled between an intermediary host such as *E. coli* for initial screening of successful cloning and then into Agrobacteria for plant transformation; they should have selectable markers to identify putative clones; to some extent the target cell or tissue type can also influence the selection; and lastly, they should be selected for the purpose of the experiment, *id est* stable or transient expression.

5.1.2.1 Ti Plasmid

The tumour inducing, Ti, plasmid (Figure 5-1) is undoubtedly the most widely used plant transformation vector as it naturally occurs in the pathogenic plant bacterium *Agrobacterium tumefaciens* which is capable of delivering oncogenic DNA into plant cells causing what is known as Crown Gall disease, characterised as tumour formation within the root-to-shoot transition zone (Peter & Jay, 2014).

LB & RB: T-DNA border regions

Tra: Ti Plasmid processing

Trb: Ti plasmid conjugation

Vir region: virulence region

virA: sensor kinase

virB: translocation

virC: T-DNA processing and recruitment

virD: T-DNA processing/transport

virE: T-DNA movement

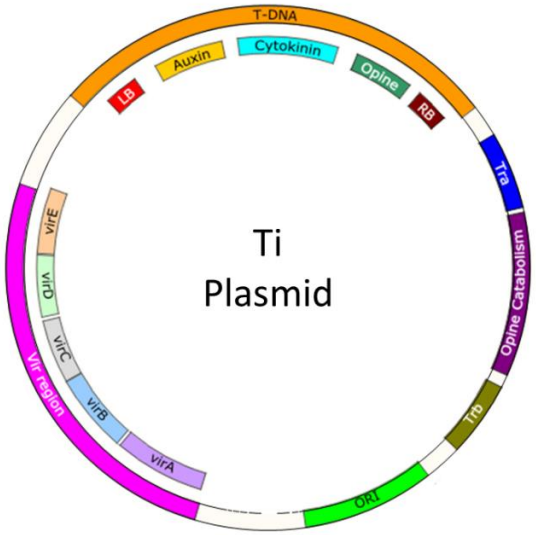


Figure 5-1: Octopine-type Ti Plasmid. Schematic representation of octopine-type Ti plasmid. Figure adapted from (Peter & Jay, 2014)

These plasmids are relatively large spanning 200 to 800 kbp depending on the type of Ti plasmid in question. The Ti plasmids carry with them genes for opine synthesis as well as genes for their catabolism and their classification is dependent on the type of opine synthesis. As shown above in a typified representation of the octopine-type Ti plasmid (Figure 5-1), all Ti plasmids encode for functions related to replicating the plasmid (ORI); conjugated transfer (*Trb*); virulence (*vir* region); opine catabolism; and sensory perception upon wounding (*virA* and *VirG*) (Peter & Jay, 2014; Zhu, et al., 2000; Winans, 1991).

In short, upon plant wounding, signals such as phenolic compounds such as acetosyringone, monosaccharides and an acidic pH induce expression of the virulence region (Winans, 1990; Winans, 1991) via the *VirA*-*VirG* sensor (Lacroix & Citovsky, 2013). This perception then results in a series of events, schematically represented in Figure 5-2 below. VirD2 and VirD1, which are produced upon transcriptional activation of the Vir region, form a nuclease that nicks the left- and right-border region (Figure 5-2) liberating the T-DNA, now a mobile single-stranded T-DNA through strand replacement synthesis (Gelvin, 2003). The T-DNA is then complexed with VirD2 and subsequently transported into the plant cell through a bacterial secretion system (Lacroix & Citovsky, 2013). In addition to the T-DNA complex, several other virulent effectors are also transported via this secretion system. The T-DNA that has entered the host cell can now be transcribed and translated. In nature, the T-DNA region encodes for plant hormones responsible for the promotion of cell growth including auxins and cytokinins (Figure 5-2), which upon entry into the plant cell, results in uncontrolled cell proliferation that finally presents itself as a tumour-like structure (Lacroix & Citovsky, 2013).

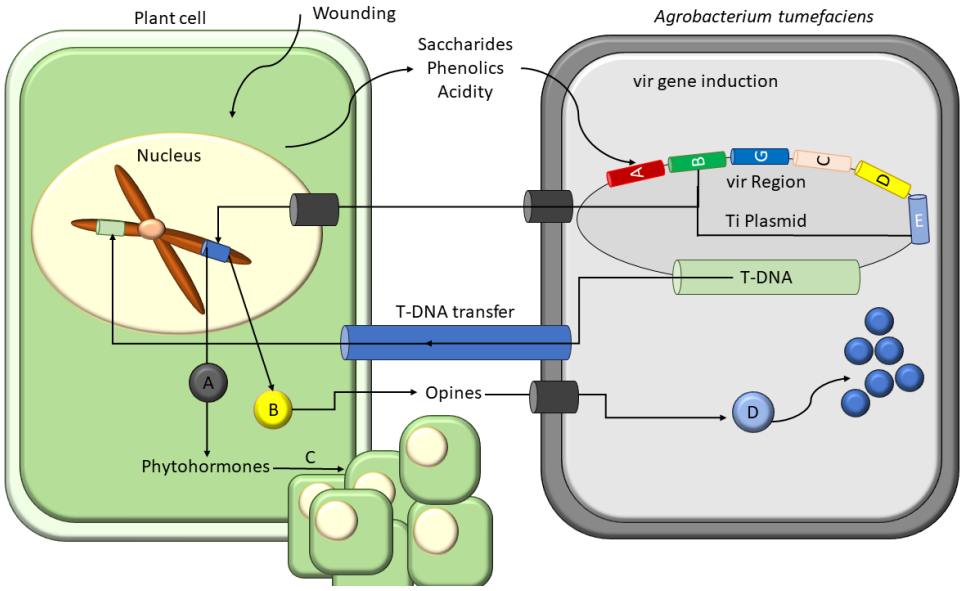


Figure 5-2: Agrobacterium tumefaciens infection. Schematic representation of various stages involved in the transfer of T-DNA from *A. tumefaciens* and a plant cell. A= phytohormone synthases, B= opine synthases; C= cell proliferation and subsequent tumour growth; D= opine hydrolases. Figure adapted from (Zhu, et al., 2000; Peter & Jay, 2014; Heldt & Piechulla, 2010)

In addition to phytohormone-encoding genes, the presence of opine synthesis genes serves the role of providing opines which the agrobacterium, through the activity of opine hydrolases, uses for metabolism. As incredible as it seems in terms of its ability to essentially highjack the plant's protein production machinery, using Ti plasmids alone for plant transformation is cumbersome. Their relatively large size and the fact that they have very low copy numbers in Agrobacteria make them difficult with which to work. In addition, these plasmids do not replicate in *E. coli*.

5.1.2.2 Binary vectors

Although the large size, low-copy number, and non-replication in *E. coli* of the Ti plasmids are cumbersome, if you want to generate a normal plant via transformation then the oncogenic portion of the Ti plasmid is problematic (Lee & Gelvin, 2008). This led to the development of the binary vector. The binary vector is first and foremost much smaller in size and lacks the oncogenic region, thereby allowing for the creation of normal transgenic plants. Most binary vectors contain similar regions performing similar functions. Those being a T-DNA left and right border sequence similar to the Ti plasmid. The gene of interest (GOI) is inserted between those regions (Lee & Gelvin, 2008; Komori, et al., 2007). In Figure 5-3 below, the GOI is inserted using Gateway cloning that results in recombination at the *attR*2 and *attR*1 sites. This insertion allows for bacterial selection using the lethal *ccdB* gene. The *vir* region can reside on a different plasmid, often in the form of a disarmed Ti plasmid that resides in the

Agrobacterium strain. Another important feature is the plant-active selection marker such as a type of antibiotic selection, in the case of pB2WG7 below it is spectinomycin, but can be any other antibiotic such as the aminoglycoside antibiotics kanamycin or hygromycin (Lee & Gelvin, 2008) .

ccdB: bacterial toxin that poisons DNA gyrase – *E. coli* harbouring this gene cannot survive

CmR: Chloramphenicol resistance gene

pVS1 StaA: stability protein from *Pseudomonas* pVS1 plasmid

pVS1 RepA: replication protein from *Pseudomonas* pVS1 plasmid

pVS1 oriV: origin of replication for *Pseudomonas* pVS1 plasmid

bom: basis of mobility region from pBR332

ori: origin of replication in *E. coli*

SmR: spectinomycin and streptomycin resistance gene

BlpR: bialophos and phosphinothricin resistance gene

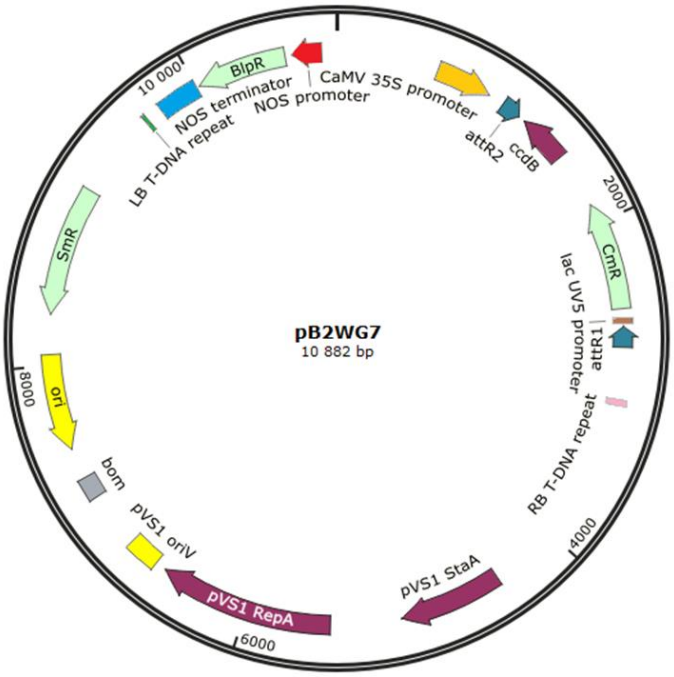


Figure 5-3: A binary vector: vector diagram showing the binary plant expression vector pB2WG7

Though not strictly a requirement, many binary vectors have promoters such as the constitutive CaMV 35S promoter, ubiquitin, or actin (Komori, et al., 2007). A common feature is the presence of rare-cutting or homing endonuclease sites within the T-DNA region where the GOI would be inserted. For pB2WG7 above those are the *attR*2 and *attR*1 sites. Another important factor to consider is the selection of plants once transformed. Here the presence of *BlpR* confers resistance to bialophos or phosphinothricin, allowing one to screen successfully transformed plants by for instance growing them on growth media supplemented with an appropriate herbicide. Other noteworthy sites are the origin of replication for both *E. coli* and *A. tumefaciens* and the presence of an opine promoter such as the nopaline synthase promoter in pB2WG7 above. These vectors are unique as they allow for shuttling between *E. coli* and *A. tumefaciens* for final plant transformation. These vectors can be subjected to either being used for transient or stable transformations and over the past 30 years a suite of highly specialised binary vectors have been developed to meet the needs of the research question; whether they be for protein localisation where a fusion tag such as a fluorescent protein is anchored onto the GOI at the 3' end or for reporter assays where a reporter gene like GUS, or LUC (Komori, et al., 2007) is added to the GOI. The use of the binary vector for plant transformation still exploits the ability of *A. tumefaciens* to infect plant cells and the presence

of a opine synthase on the binary vector allows for the production of opines that would normally have been made if it had just been a regular Ti plasmid.

5.1.2.3 Viral vectors

Viruses are notoriously synonymous with widespread infection. Plants are no exception to the devastation that viruses can cause. A virus in its strictest sense is an obligate parasite, requiring the host's cellular machinery to produce proteins of relevance. There is naturally a plethora of different viruses that infect different plants, for example the tobacco mosaic virus (*Tobamovirus*), the cauliflower mosaic virus (*Caulimovirus*), tomato spotted wilt virus (*Orthotospovirus*), and the potato virus X and Y (*Potexvirus*) to name a few. Although synonymous with negative connotations of disease, these viruses have been engineered to manipulate gene expression in plants through the use virus-induced gene silencing (VIGS) vectors, or coupled with CRISPR/Cas nuclease genome editing to enhance the breeding of economically important crops, or used for their rapid and relatively cheap production of virus-like particles in the development of candidate vaccines (Wang, et al., 2020)..

A viral vector in essence is a vector containing regions associated with a particular virus. In Figure 5-4 below, the authors constructed an autonomously replicating vector based on the bean yellow dwarf geminivirus (BeYDV-m) (Regnard, et al., 2010). The vector also includes T-DNA left and right border regions from the Ti plasmid enabling it to transfect plants. The plasmid has also been constructed in such a way as to allow replication in both *E. coli* and *A. tumefaciens*, thereby ensuring that proper initial screening in *E. coli* can be done before agrobacterium-mediated infiltration.

The BeYDV virus only produces four proteins: a coat, movement, and two replication proteins (Rep and RepA) (Diamos & Mason, 2019). Viral protein production is primarily driven by a promoter in the long intrinsic region (Diamos & Mason, 2019). The Rep protein, which is produced during the early stages of infection, is an important protein responsible for an array of viral functions including initiating rolling circle replication, ligating newly replicated viral DNA to produce a circular viral genome, and is thought to be involved in creating a suitable environment for viral replication (Diamos & Mason, 2019). The pRIC vector below thus incorporates important plant and viral regions that 1) allow for the vector to enter plant cells via *A. tumefacien* transfection (LB and RB regions); 2) viral replication of GOI regions (rep); driven by 3) an enhanced constitutive promoter (P35SS) resulting in much greater expression overall.

LB & RB: T-DNA border regions

SAR: scaffold attachment region of the tobacco *Rb7 gene*

P35SS: Enhanced CaMV35 S promoter

CHS: chalcone synthase 5'-UTR

LIR: BeYDV long intergenic region

SIR: BeYDV short intergenic region

pA35SS: CaMV35 S polyadenylation signal

rep: BeYDV *rep* gene

RK2 ORI: origin of replication for *A. tumefaciens*

bla: ampicillin/ carbenicillin resistance gene

ColE ORI: origin of replication in *E. coli*

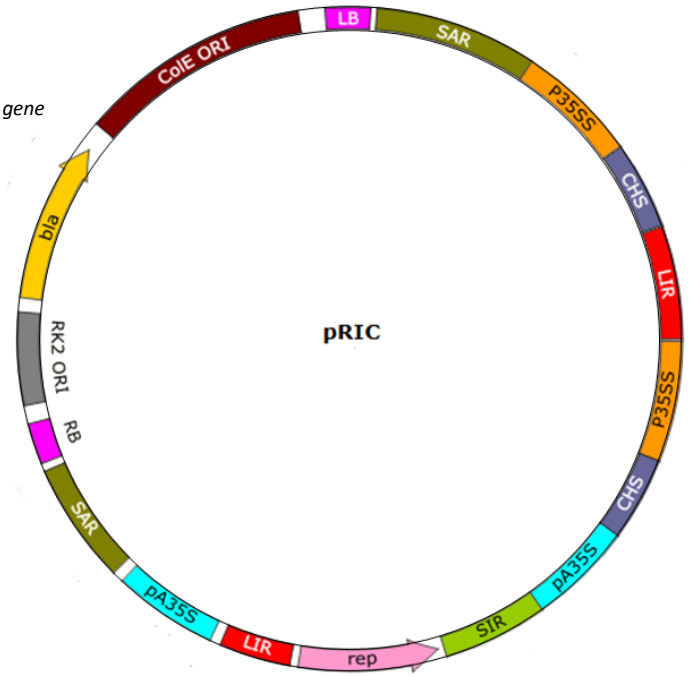


Figure 5-4: Viral vector used in agroinfiltration. Vector map showing points of interest of the viral vector pRIC used in plant transformation. Figure adapted from (Regnard, et al., 2010)

The end product is a plant transformation vector capable of having gene copy numbers as high as 10^9 seven days following Agrobacterium infiltration and producing protein at amounts several fold higher than a non-replicating control (Regnard, et al., 2010). The high gene copy number and greater protein production makes, at least in the case of pRIC, viral vectors prime candidates for scaled up protein production. Another aspect of the viral vectors to consider is the untranslated regions (UTRs) at the 5' and 3' end. For viruses, these UTRs play two important roles; first being viral RNA replication (3'→5' on positive-stranded viral RNA) requiring the synthesis of negative-stranded RNA and the second being the actual translation of viral proteins (5'→3' on positive-stranded RNA) (Peyet, et al., 2019). Though naturally 'good' at what they do, (Peyet, et al., 2019) sought to develop synthetic UTRs to further enhance or improve the current expression of an already widely used (Peyret & Lomonossoff, 2013) viral vector pEAQ-HT. Their work relied on various combinations of 5' and 3' UTRs and they were able to demonstrate that different combinations resulted in varying degrees of protein accumulation (Peyet, et al., 2019) suggesting that not only are aspects of the viral region such as the rep protein important in deciding on a vector, but perhaps the redesign of one or both of the UTRs should be a factor to consider. This naturally lends itself well to optimising the degree of protein production. The viral vectors, such as those routinely used within the Biopharming Research Unit, have a uniquely suited purpose; high production of the protein of interest in a short period of time, often in the range of 2-7 days post-infection. They are

however, limited by the capacity of the vector to maintain self-replication and natural cell division of the host, leading to gradual decrease of protein production. Another factor that impacts the efficacy of transient systems using viral vectors that overproduce is post-translational gene silencing (Chicas & Macino, 2001), often overcome by inclusion of the silencer such as the P19 silencer, and the potential accumulation of exogenous proteins which may become lethal to the plant itself.

5.1.3 Selection of Agrobacterium strain

Agrobacterium is a Gram-negative soil-borne bacterial genus in the rhizobiaceae family. The rhizobiaceae family encompasses an array of genera, all of whom aerobic with some being diazotrophic..

Table 5-1: Commonly used Agrobacterium strains in plant transformation

***Agrobacterium* strain**	**Chromosomal background [a]**	**Ti-plasmid [b]**	**Antibiotic resistance [c]**	**Opine**
A136	C58	Cured Ti	Rif, Nal	Nopaline
AGL-1	C58, RecA	Disarmed pTiBo542ΔT-DNA	Rif, carb	Succinamopine
C58C1	C58	Cured	Rif	Nopaline
C58C1 (pTiB6S3Δ, pCH32)	C58	Disarmed pTiB6S3ΔT Helper pCH32	Rif, Carb, Tet	Octopine
EHA101	C58	Disarmed pTiBo542ΔT-DNA	Rif, Kan	Nopaline
GV3101	C58	Cured	Rif	Nopaline
GV3101 (pMP90)	C58	Disarmed pTiC58ΔT-DNA	Rif, Gent	Nopaline
LBA4404	Ach5	Disarmed pAL4404	Spec, Strep, Rif	Octopine
NTL4 (pKPSF2)	C58	Disarmed pTiChry5ΔT-DNA	Em	Chrysopine

[a] Refers to the genome origination of that particular strain. C58 for example comprises four replicons (two mega-plasmids and one circular and one linear chromosome) (Nester, 2015)

[b] Refers to method used to make strain non-oncogenic

[c] Rif = rifampicin; Nal = nalidixic acid; carb = carbenicillin; Tet = tetracycline; Gent = gentamicin; Kan = kanamycin; Spec = spectinomycin; Strep = streptomycin; Em = erythromycin

Table adapted from Hwang, et al., (2017) and Hellens, et al., (2000) and citations therein.

Many species within this family are non-pathogenic and are found within the rhizobia and are often symbiotic with many plants. *Agrobacterium tumefaciens* and *Rhizobium rhizogene*s however, are pathogenic, the former being responsible for the development of Crown Gall disease. Some commonly used selection antibiotics which form part of the transformation vector include kanamycin, gentamicin, tetracycline, and streptomycin or spectinomycin. Agrobacterial strains, however, have naturally occurring antibiotic resistance, either encoded chromosomally or via their Ti plasmid. For example, GV3101 and EHA101 have chromosomal resistance to rifampicin, whereas EHA101 has additional resistance to kanamycin from its disarmed pTiBo542ΔT-DNA Ti plasmid as shown in Table 5-1 below. LBA4404 has resistance to three antibiotics, two of which are routinely used in plant transformation. Care should thus be taken when selecting vectors and strains so as not to duplicate the resistance (Hellens, et al., 2000).

For general plant transformation such as Arabidopsis floral dipping or tobacco leaf vacuum infiltration, the elimination of Agrobacterium per se is not of great importance. However, for transformation techniques that require tissue culture (see below), the elimination of any residual Agrobacterium following co-cultivation is vital. Some antibiotics such as carbenicillin and amoxicillin are commonly used to curb the further growth of Agrobacteria used during co-cultivation. Some strains, however, have resistance to these penicillin-based antibiotics and the use of other antibiotics such as cefotaxime is required. Additional considerations such as broad host-range replication to allow it to replicate in *E. coli* (in order to enhance plasmid yields) and replicate in Agrobacteria for transformation need to be factored. Taken together, although the choice of the transformation vector may be arbitrary, it has to be compatible with the Agrobacterium strain. There should be no antibiotic duplication between the two, there needs to be compatibility between *E. coli* and Agrobacteria, and lastly, the sterility needs to be maintained by elimination of Agrobacteria using an antibiotic to which it is not resistant.

5.1.4 Selection of transformation method

Once the most suitable plant expression vector has been selected and its compatibility with a suitable agrobacterium strain has been determined, the most appropriate method for plant transformation has to be selected. Naturally, there are numerous methods from which to choose, and traditionally dicotyledonous plants have been much ‘easier’ to transform via agrobacterium-mediated delivery than monocotyledonous plants with a suite of transgenics having been successfully produced (Sood, et al., 2011) to date. The use of agrobacterium-mediated transformation, however, was a much slower process to optimise for monocotyledonous plants. It was not until the early 90s that the first transgenic

monocotyledon, *Oryza sativa*, was made via agrobacterium-mediated gene transfer (Chan, et al., 1992).

Monocotyledons are notoriously difficult to transform, often requiring endless troubleshooting and optimisation. Sood, et al., (2011) propose that the difficulties underlining the effective use of Agrobacteria to mediate the transformation of monocotyledons are related to fundamental differences between the monocotyledons and the dicotyledons. An important aspect of agrobacterium-mediated gene transfer is the ability of the Agrobacteria to adhere to and transfer into the host cell. The chemical composition of the cell walls as well as the ability of underlining meristematic tissue to dedifferentiate between monocotyledons and dicotyledons have been proposed as potential obstacles in the ability of Agrobacteria to infect monocotyledons and transfer T-DNA respectively. The manner in which monocotyledons respond to wounding is another aspect that influences the efficacy of Agrobacterial infection. Wounding, as discussed above, plays a vital role in ensuring Agrobacterial infection. Monocotyledons can differentiate into tissue that is highly lignified or sclerified upon wounding as a mechanism to protect the surrounding tissue from pathogenic attack. It is thus possible to speculate that during wounding and the resulting response can result in only a few cells being competent for transformation and this low number of cells can ultimately lead to insufficient tumour formation (Graves, et al., 1988) (under natural conditions) or ineffective transformation. An additional aspect to consider is the production of *vir*-inducing compounds. As stated above, the induction of the virulence region of Ti-plasmids is accomplished through exudation of phenolics and/or saccharides from the host plant. Traditionally, and within the scope of biotechnology, acetosyringone (a natural phenolic compound often produced in response to wounding and/or plant-pathogen recognition (Baker, et al., 2005)), has been used to induce the *vir* region of Ti and Ti-based vectors in Agrobacteria. It has been shown that monocotyledons exude an array of compounds upon wounding but their effect on agrobacterium recognition and subsequent infection is a contentious one. It can be stated that the interaction between plant cells and agrobacterium during infection is a complex process which is further exasperated by the fact that monocotyledons produce inhibitors in conjunction with *vir* inducers (Sood, et al., 2011). All this taken into consideration, monocotyledons are notoriously difficult to transform, and the successful transformation has to be tailored to the species, especially when that species is not a commonly studied crop such as maize or rice. Despite the difficulties, monocotyledons have been, and still are, being successfully transformed. Some successful transformation processes, shown in Figure 5-5, often rely on the successful establishment of a plant tissue culture process to either isolate single cells such as protoplasts for cell cultures or the induction of callous tissue and subsequent somatic embryogenesis before they are transformed.

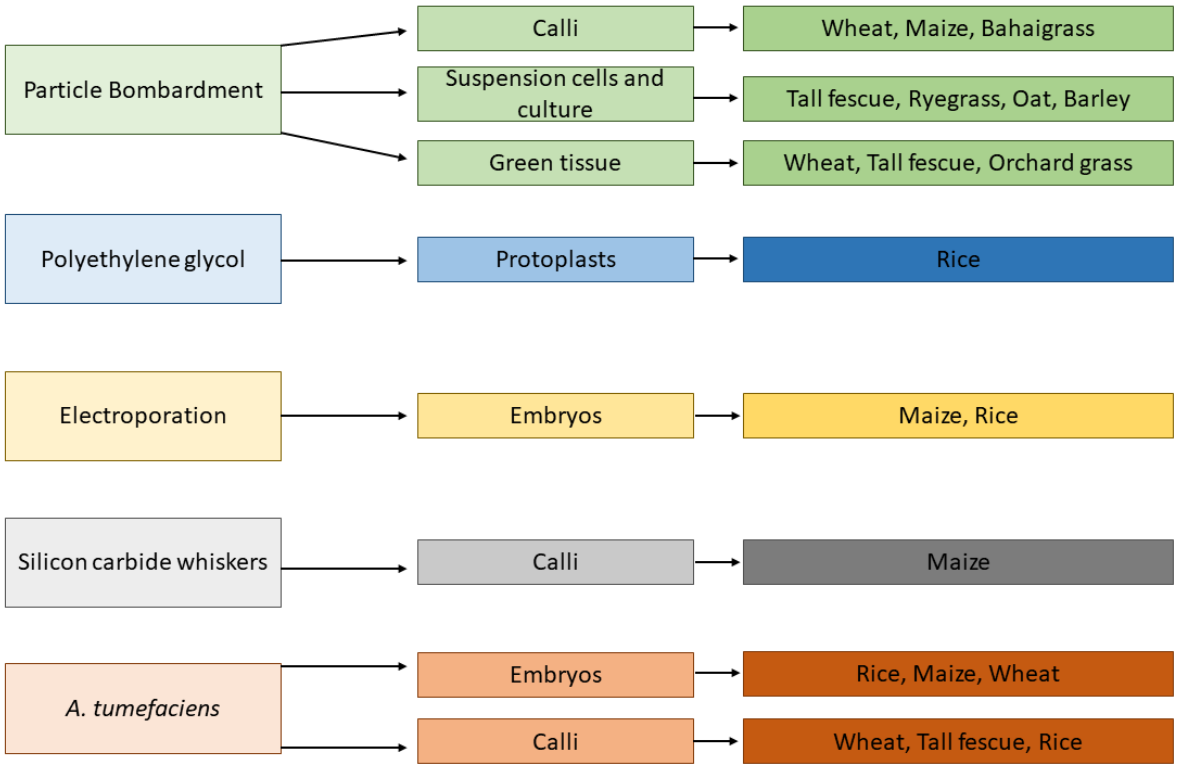


Figure 5-5: Summary of some transformed monocotyledons. Processes used to transform monocotyledons showing source of explant material used for transformation with some examples of successfully transformed monocotyledons. Information from Sood, et al., (2011) and references therein.

5.1.4.1 Callus induction and somatic embryogenesis

Under natural conditions, when some herbaceous dicotyledons are wounded the vascular cambium can form what is known as callus. The callus that forms from wounding is parenchymous in nature and arises from parenchyma cells that undergo division and differentiate into meristematic tissue (Dickison, 2000). This meristematic callous tissue is totipotent and able to regenerate into a whole plant. Calli can be further described as being friable or compact and are often then termed non-embryonic as they do not undergo somatic embryogenesis (Ikeuchii, et al., 2013). Others can be termed embryonic should they show signs of root or shoot development (Ikeuchii, et al., 2013). Somatic embryogenesis refers to the ability of some calli to develop into embryos from parental somatic tissue. This process of callus initiation and subsequent somatic embryogenesis can be exploited through *in vitro* culture of explants on callus induction media (CIM). Typically, this medium is composed of micro-, and macronutrients supplemented with vitamins such as nicotinic acid, pyridoxine HCl, thiamine hydrochloride, and myo-inositol. A carbon source in the form of 2-3% sucrose, maltose, or glucose is added. Antioxidants such as ascorbic acid and citric acid or activated

charcoal are also included to help reduce the accumulation of phenolic compounds. In order to obtain callus from explants, the inclusion of hormones is necessary. Traditionally 2,4-dichlorophenoxy acetic acid (2,4D) is included to initiate the callus induction process. 2,4D is a systemic herbicide commonly used in garden variety herbicides to kill broadleaf weeds. Its mode of action is to mimic the action of auxin which results in uncontrolled cell proliferation. Other hormones commonly used are 6-benzylaminopurine (BAP), a synthetic cytokinin involved in cell division, and α-naphthalene acetic acid (NAA), a synthetic auxin commercially used as a rooting agent. The ratio of cytokinin to auxin in CIM is vital in establishing somatic embryos (Ikeuchii, et al., 2013) and can differ greatly among explant source and species. This particular ratio needs to be determined empirically for each species and can become a laborious process as not all explants respond equally to the same hormone conditions. Explants are first surface sterilized and placed on the sterile CIM and incubated in the dark usually at 25°C. The emergence of callus can take anywhere from a few days to a few weeks depending on explant source. Once calli are obtained, frequent sub-culturing needs to occur to remove any necrotic calli (usually due to phenolic accumulation) and to ensure that the growing calli regularly receive fresh nutrients.

Although laborious and often plagued with fungal and bacterial cross-contamination, the transformation of monocotyledons typically requires the establishment of somatic embryos which are then transformed. The method for transformation is perhaps not as important when compared to the selection of the vector or Agrobacterium strain. In general, the two most widely used methods for transforming calli are via biolistic bombardment, which involves coating a heavy metal with exogenous DNA and firing it into the callus, or via Agrobacterium mediated co-cultivation, which involves cultivating the callus in pre-activated Agrobacterium harbouring the desired construct before culturing on selection media. Other methods used include the isolation of protoplasts from leaf or root tissue and transforming using polyethylene glycol followed by the induction of somatic embryogenesis.

Chapter aims and objectives:

This Chapter is primarily aimed at developing a method for transgenic generation of *E. tef* plants. In order to do this the following will be investigated:

1. Can *E. tef* be used to assess transient expression of eGFP through leaf infiltration using viral vectors?
2. Which explant sources and conditions are required to induce callus formation?
    a. Of these conditions, which yields embryonic calli reliably and reproducibly?
3. Can agrobacterium-mediated transformation be used to transform embryonic calli?
4. What conditions are required for embryonic callus conversion to regenerated plants?

Methods

As this Chapter's main focus was aimed at developing a method for transformation of *E. tef* the following should be noted. First, initial tests of standard leaf infiltration were tested to determine whether this method is suitable for transient expression evaluation. This method yielded inconclusive results but indicated that to some degree it is plausible. Secondly, a method for stable integration was investigated wherein somatic embryos were transfected with agrobacterium first harbouring viral vectors expressing enhanced GFP as a means to evaluate the transformation efficiency of the method. This had yielded inconclusive results but indicated to some degree that the Agrobacterial strain was compatible with calli. Thirdly, the EnHSP70 localisation construct from Chapter 4 was used to transfect embryonic calli for stable integration. Various methods used throughout this process are presented here.

5.2.1 *Eragrostis tef* germination and growth

Eragrostis tef (South African Brown variety) seeds were sown onto a soil mix of 50% peat moss and 50% vermiculite. The surface of the soil was dampened and then compacted before sowing the seeds. Once sown, the seedling trays were covered with a layer of clear plastic wrap and allowed to germinate until the cotyledons were approximately 0.5 cm in height after which the plastic wrap was removed. Plants were then grown at 12 hour day light at 25±2°C/14±2°C day/night. The plants were fertilized two-to-three times during their growth using phostrogen at the recommended dosage. As the plants outgrew their containers, they were potted into larger ones.

5.2.2 Agrobacterium preparation and initial tef leaf infiltration trial

Agrobacterium tumefacien harbouring various expression vectors were obtained from the Biopharming Research Unit (BRU) at the University of Cape Town. All the expression vectors had the enhanced green fluorescent protein cloned within it. One millilitre of a glycerol stock was added to 500 mL of lysogeny broth supplemented with appropriate antibiotics. This was then allowed to incubate at 26-28°C with continuous shaking for three days. On the day of plant infiltration, the culture's OD_{600} was measured and depending on the required OD of the culture, the culture was further diluted using a resuspension buffer (10 mM 2-(N-morpholino)ethanesulfonic acid (MES), 10 mM $MgCl_2$ and 200 µM acetosyringone pH 5.6). This was then left at room temperature for 30 minutes to induce the *vir* genes of the viral vectors. Two-week-old *E. tef* seedlings were vacuum infiltrated with the resuspended Agrobacteria harbouring an expression vector. Infiltration was done by fully submerging the

aerial parts of the seedlings into the Agrobacteria and applying a vacuum for 30 seconds and then releasing the vacuum. The vacuum process was repeated twice. Seedlings were then transferred onto soil and left to resume growth. Total protein was isolated from *E. tef* by first grinding the tissue using a plastic Eppendorf tube pestle with liquid nitrogen until a fine powder was achieved. To this, 200-300 µL of a homogenisation buffer comprising of 50 mM 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES)-KOH pH7.5, 250 mM sucrose , 5% (*v/v*) glycerol, 10 mM EDTA, 0.5% (*w/v*) PVP, 50 mM $Na_4P_2O_7 \cdot 10H_2O$, 1 mM $Na_2MoO_4 \cdot 2H_2O$, 25 mM NaF, 3 mM 1,4-dithiothreitol (DTT) , and 10 $\mu L.mL^{-1}$ phenylmethylsulphonyl fluoride (PMSF) was added, left on ice for 10 minutes, and centrifuged at 4°C for 10 minutes at maximum speed using a benchtop centrifuge. The supernatant was collected as crude isolate. For protein separation, visualisation, and detection, an appropriate quantity (either based on equal total protein amount or equal volume) of the crude protein was heated to 95°C for ten minutes with 5X sample application buffer (50 mM Tris-HCl pH6.8, 5% SDS, 80% glycerol). This was then loaded onto a 10% resolving sodium dodecyl sulphate polyacrylamide gel and run at 100 V in 1x running buffer comprising (0.15% (*w/v*) tris, 0.72% (*w/v*) glycine, and 0.1% (*w/v*) SDS) for approximately one-and-a-half hours. For visualisation purposes, the gel was first rinsed with distilled water, heated in a microwave at maximum heating for 10 seconds before being stained with Coomassie staining solution (40% (*v/v*) methanol, 10% (*v/v*) glacial acetic acid, and 0.1% (*w/v*) bromophenol blue) for one hour at 37°C with continuous shaking. The gel was de-stained by removing the Coomassie stain, adding de-staining solution (10% *(v/v*) methanol and 10% (*v/v*) glacial acetic acid), and heating it for 10 seconds in a microwave at maximum heat and incubating the gel at 37°C for 1-2 hours with a piece of tissue paper added to absorb the bromophenol blue. For detection of proteins of interest, a Western blot was done whereby a non-Coomassie stained gel was sandwiched with a nitrocellulose membrane and allowed to transfer at 50 V for approximately one-and-a-half hours at 4°C in precooled transfer buffer (0.58% (*w/v*) Tris, 0.29 (*w/v*) glycine, 20% (*v/v*) methanol, pH9.2). Prior to blocking, the blot was stained with ponceau (5% (*v/v*) glacial acetic acid and 0.1% (*w/v*) Ponceau S tetrasodium) to ensure complete transfer had occurred. The blot was then blocked with blocking buffer (1x PBS, 5% (*w/v*) skimmed milk power, and 0.1% Tween-20™ (*v/v*)) for 30 minutes with gentle shaking. The blocking buffer was removed and a primary anti-GFP antibody (1:1000) was added and allowed to probe overnight at 4°C with gentle shaking. The primary antibody was then removed, and the blot was then washed four times with blocking buffer with gentle shaking. The secondary antibody, which is conjugated to alkaline phosphatase, was added, and allowed to probe for one hour at 37°C with gentle shaking. The blot was then washed four times with blocking buffer lacking the milk component. Once washed, 3-5 mL of 5-bromo-4-chloro-3-inolyl phosphate/nitro blue tetrazolium (BCIP/NBT)

substrate was added and the blot left in the dark to develop. In addition, relative eGFP fluorescence was measured using a fluorometer (Glomax).

5.2.3 Callus induction

All explant material was surface sterilized by submerging the material in a mixture of 0.3% - 1% sodium hypochlorite and 0.2% Tween20, vortexed briefly and allowed to shake vigorously for 5-10 minutes. For seeds, and additional wash with 0.1% $HgCl_2$ was done. This was removed and 70% (*v/v*) ethanol was added, vortexed and allowed to shake for the same period. The explant material was rinsed five times using filter sterilized autoclaved MiliQ water. All plant tissue culture work was done in a laminar flow hood. All components used for media preparation were filter-sterilized (0.22 µM) except for the agar which was autoclaved. All callus induction and co-cultivation on solid media was done in the dark at 25±2°C unless otherwise stated using 0.8% (*w/*v) agar. Media used for the various stages is summarised in Table 5-2 below with further description provided for the various explants used throughout.

5.2.3.1 Immature embryos (IE)

E. tef was grown until anthesis occurred (approximately 60-70 days post-sowing). Immature embryos were harvested approximately 7-10 days post-anthesis by cutting the peduncle and removing the immature embryos from the spikelets. These were then sterilized and plated onto CIM-IE.

5.2.3.1 Cotyledons and roots

Sterilized seeds of *E. tef* were placed onto SGM and left in the dark for 10 days. Root tissue was cut into two sections with a portion of one half still anchored to the seed. Cotyledons were cut into approximately 0.1 – 0.5 mm strips with a scalpel blade. Excised explants were then plated onto CIM-IE.

5.2.3.1 Mature seeds

Sterilized mature seeds from *E. tef* were initially plated onto CIM-IE. Upon germination, the emerging radicle developed callous tissue which was removed from the seed 1-2 weeks post-callus formation. During this removal, the calli that had formed were further divided by cutting the calli in two and placing the cut surface onto CIM-IE or CIM-MS.

Table 5-2: Media components used during plant tissue culture

Media	Components
Callus Induction of Immature embryos (CIM-IE)	4.4 $g.L^{-1}$ Murashige and Skoog basal salt mixture (MS) with Gamborg's vitamins (+GV), 2.0 $mg.L^{-1}$ 2,4 dichlorophenoxy acetic acid (2,4D), 1 $mg.mL^{-1}$ 6-benzylamino purine (BAP), 10 $mL.L^{-1}$ antibiotic-antimycotic solution (AA), 3% (*w/v*) maltose, 2% (*w/v*) sucrose, 150 $mg.L^{-1}$ ascorbic acid (VC), 150 $mg.L^{-1}$ citric acid (CA), 300 $mg.L^{-1}$ casein hydrolysate (CH), 10 mM MES pH 5.0
Callus induction of mature seeds (CIM-MS)	For induction: 4.4 $g.L^{-1}$ MS -GV, 1 $g.L^{-1}$ CH, 1 $mg.L^{-1}$ thiamine-HCl, 250 $mg.L^{-1}$ myoinositol, 30 $g.L^{-1}$ sucrose, 2 $mg.L^{-1}$ 2,4D. For propagation, 0.5 $mg.L^{-1}$ 2,4D, 0.5 $mg.L^{-1}$ BAP
Regeneration of mature seeds (RMS)	For somatic conversion: 4.4 $g.L^{-1}$ MS -GV, 1 $mg.L^{-1}$ gibberellic acid. For regeneration: 2.2 $g.L^{-1}$ MS -GV, 10 $g.L^{-1}$ sucrose, 4 $g.L^{-1}$ activated charcoal
Liquid co-cultivation callus induction (LCC-CIM)	4.4 $g.L^{-1}$ MS -GV, 2.0 $mg.L^{-1}$ 2,4D, 3% (*w/v*) maltose, 2% (*w/v*) sucrose, 150 $mg.L^{-1}$ VC, 150 $mg.L^{-1}$ CA, 300 $mg.L^{-1}$ CH, 200 µM acetosyringone, 10 mM $MgCl_2$, 10 mM MES pH 5.0
Solid co-cultivation callus induction (SCC-CIM)	4.4 $g.L^{-1}$ MS +GV, 2.0 $mg.L^{-1}$ 2,4D, 1 $mg.mL^{-1}$ BAP, 10 $mL.L^{-1}$ AA, 3% (*w/v*) maltose, 150 $mg.L^{-1}$ VC, 150 $mg.L^{-1}$ CA, 300 $mg.L^{-1}$ CH, 100 µM acetosyringone, 10 mM MES pH 5.0
Co-cultivation selection callus induction (CC-SCIM)	4.4 $g.L^{-1}$ MS +GV, 2.0 $mg.L^{-1}$ 2,4D, 1 $mg.mL^{-1}$ BAP, 10 $mL.L^{-1}$ AA, 3% (*w/v*) maltose, 150 $mg.L^{-1}$ VC, 150 $mg.L^{-1}$ CA, 300 $mg.L^{-1}$ CH, 10 mM MES pH 5.0, 1.5 $mg.L^{-1}$ BASTA
Plant regeneration rooting (PRM)	4.4 $g.L^{-1}$ MS +GV, 1 $mg.L^{-1}$ α-naphthalene acetic acid (NAA), 1 $mg.L^{-1}$ BAP, 10 $mL.L^{-1}$ AA, 3% (*w/v*) maltose, 2% (*w/v*) sucrose, 150 $mg.L^{-1}$ VC, 150 $mg.L^{-1}$ CA, 300 $mg.L^{-1}$ CH, 10 mM MES pH 5.0
Callus cell suspension liquid culture (CSLC)	4.4 $g.L^{-1}$ MS +GV, 2.0 $mg.L^{-1}$ 2,4D, 10 $mL.L^{-1}$ AA, 2% (*w/v*) sucrose, 2% (*w/v*) sorbitol, 150 $mg.L^{-1}$ VC, 150 $mg.L^{-1}$ CA, 10 mM MES pH 5.0
Seed germination (SGM)	4.4 $g.L^{-1}$ MS -GV, 2% (*w/v*) sucrose, 10 $mL.L^{-1}$ AA

Subsequent optimisation of mature seed callus induction was done on CIM-MS as per the methods from Kebebew, et al., (1998) with a few modifications. Seeds were plated on CIM-MS and maintained in the dark at 24°C for 4 weeks upon which developing calli were transferred onto propagation CIM-MS and cultured for a further 4 weeks. A subset of the

propagating calli were transferred to light conditions as described above. Following the 4-week propagation, calli were moved onto RMS containing gibberellic acid to convert somatic emrbyos into plants and maintained on this media until emergence of plantlets occurred upon which they were cultured on media without gibberrillic acid to initiate shoot and root formation.

5.2.4 Co-cultivation of calli and Agrobacteria with viral vectors

Calli which had successfully formed were initially co-cultivated with agrobacterium harbouring a viral vector expressing eGFP. The method used below was a modification of that described by Gebre, et al., (2013). To this end, Agrobacteria were grown until the OD_{600} was approximately 1.0. The Agrobacterial culture was diluted to OD_{600} = 0.7 - 0.8 using LCC-CIM. The *vir* region was induced by shaking the diluted Agrobacteria for 30 minutes at 28°C. Upon induction of the *vir* region, calli were dropped into the diluted Agrobacteria solution and shaken for 4-6 hours to allow for infection. Calli were removed, dapped onto sterile filter paper, rinsed with sterile water once and then rinsed with LCC-CIM before being plated onto SCC-CIM lacking BASTA. The co-cultivated calli were incubated in the dark at 25°C ±2°C for three days before being rinsed once more with sterile water supplemented with 10 $mL.L^{-1}$ antibiotic-antimycotic solution and plated onto SSC-CIM. The use of the viral vectors unfortunately did not impart selection pressure as they lack a selection marker within the T-DNA region of the expression vector, however, the EnHSP70 transgene cloned into pB2WG7 provided selection on the basis of resistance to BASTA. Peak expression of eGFP from the viral vectors is expected to occur between day 5 to day 9 post-infiltration in tobacco, as such calli which had been transfected were sampled on day 5, 7, and 9 post-co-cultivation. Calli were visually assessed for any visible eGFP fluorescence by shining a hand-held UV light over the calli. Total protein was isolated from calli as described above for the tef seedlings. Relative fluorescence was measured on 1:10 diluted total isolated protein. Nickel affinity chromatography purified eGFP, obtained through expressing eGFP which had been cloned into pPROEX in BL21 *E. coli* kindly provided by Shakiera Sattar, served as a positive control and non-infected callus as a negative control. A total of 5µL of crude protein was dotted onto a nitrocellulose and allowed to air-dry for 30 minutes before being blocked by 1X TBST blocking buffer (5% skimmed milk, 20 mM TRIS-HCl, 150 mM NaCl, 0.05% (*v/v*) Tween 20™) for one hour. Once blocked, the blocking TBS buffer was removed and the membrane was washed three times with 1X TBST (20 mM TRIS-HCl, 150 mM NaCl, 0.05% (*v/v*) Tween 20™) for four minutes each then three more washes with 1X TBS (20 mM TRIS-HCl, 150 mM NaCl) for four minutes each. An anti-GFP primary antibody (1:1000) was added, and the blot was incubated at room temperature for 2 hours with gentle shaking. The primary antibody was removed, and the blot washed three times with 1X TBST for four minutes each then washed

three times with 1X TBS for four minutes each. The secondary antibody (Alkaline phosphatase conjugated antibody) was added and incubated at room temperature for one hour with gentle shaking. This antibody was removed, and the blot washed three times with 1X TBST for four minutes each then washed three times with 1X TBS for four minutes each. Once washed, 3-5 mL of 5-bromo-4-chloro-3-inolyl phosphate/nitro blue tetrazolium (BCIP/NBT) substrate was added and the blot left in the dark to develop for 30 minutes. The dots that showed detection of eGFP on the dot blots were further analysed by performing SDS-PAGE and Western Blotting as described above.

5.2.5 Co-cultivation of calli and Agrobacteria with EnHSP70

Co-cultivation of embryonic calli with Agrobacteria harbouring the overexpressing mCherry::EnHSP70 in pB2WG7 or empty vector pB2WG7 was done similar to above with a few modifications. First, 40 mL of the second Agrobacterial overnight culture was centrifuged at 3 500 x g and resuspended in LCC-CIM. A total of 65 calli were added into each Agrobacterial resuspension solution and gently shaken for 6 hours in the dark at room temperature. Calli were removed and blotted dry on sterilized Whatman filter paper and placed on SCC-CIM. Calli were then incubated at 25±2 °C in the dark for one week before being moved onto SCC-SCIM supplemented with 1.5 $mg.L^{-1}$ BASTA (Beyer). These were then allowed to grow for one week before being briefly rinsed in filter-sterilized water and moved onto PRM supplemented with 50 $\mu g.mL^{-1}$ cefotaxime. Calli were allowed to initiate regeneration until formation of roots and/or shoots emerged.

5.2.6 Plant regeneration

Callus regeneration was done in a growth room at 25°C, 16 h light, 8 h dark, and 55% relative humidity. Regenerating callus was sub-cultured every 2-3 weeks until the emergence of roots at which point they were cultured in 50 mL glass tissue culture jars to develop further. Various hormone combinations were tested to establish optimal ratio of auxin to cytokinin.

Results

5.3.1 Initial tef infiltration

The initial pilot infiltration test (data not shown) revealed that an OD_{600} of 0.4 of the Agrobacteria solution yielded highest relative eGFP fluorescence. As such, and for comparative purposes, OD_{600} of 0.3, 0.4, and 0.5 was selected for the second pilot trial. The relative fluorescence values obtained from the partially dehydrated and fully hydrated plants infiltrated with agrobacterium harbouring pRIC3.0_eGFP revealed that eGFP fluorescence was detected in all the treatments (Figure 5-6). The highest fluorescence was detected on day 5 post-infiltration for infiltration at an OD_{600} of 0.4. Across treatments, the hydrated plants displayed higher RFU values than the partially dehydrated or the non-infiltrated plants. A prior pilot study (data not shown) had revealed that beyond day 6 post-infiltration very little eGFP was detectable. This led to focusing purely on day 4,5, and 6. Low levels of autofluorescence was also detected in the non-infiltrated tissue.

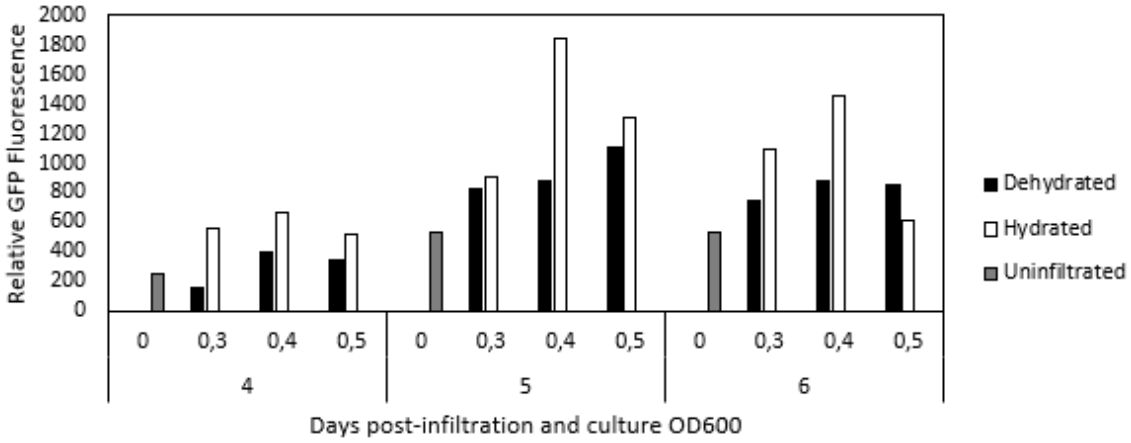


Figure 5-6: eGFP fluorescence in vacuum infiltrated *E. tef* leaves: relative eGFP fluorescence of partially dehydrated, hydrated, and non-infiltrated 2-week old *E. tef* seedlings at 4,5, and 7 days post agrobacterium mediated infiltration at OD600 of 0.3, 0.4, and 0.5. N=3

A comparison was done in parallel with the infiltrated *E. tef* with purified eGFP isolated from BL21 *E. coli* cells and Agrobacterium-infiltrated *Nicotiana benthamiana* from BRU to assess the degree of fluorescence observed from the infiltrated *E. tef*. By comparison, the highest eGFP fluorescence at approximately 1800 RFU was several orders of magnitude lower than either the eGFP sourced from *N. benthamiana* which measured greater than 80 000 or BL21 *E. coli* measuring approximately 40 000. Total protein was isolated from day 5 and day 6 post-infiltrated *E. tef* as well as non-infiltrated *E. tef* and separated on an SDS-PAGE and

subsequently probed for the presence of eGFP. eGFP isolated from 5 days post-infiltrated *N. benthamiana* was included as positive control. As can be seen in Figure 5-7, there is a clear absence of the 26.8 kDa eGFP on the Coomassie stained SDS-PAGE (Figure 5-7 A) and not detectable on the Western blot (Figure 5-7 B) either. During transfer of the blot some protein from the *N. benthamiana* sample contaminated the uninfected lane which does not have a corresponding band in the Coomassie stained gel.

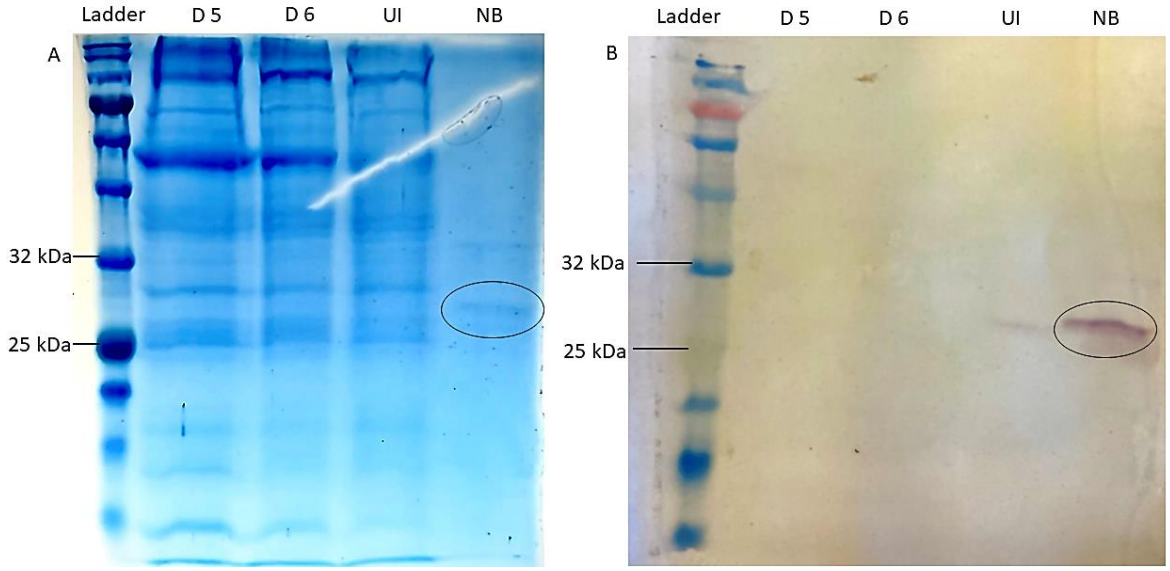


Figure 5-7: eGFP detection of vacuum infiltrated *E. tef*: a Coomassie stained SDS-PAGE (A) and corresponding Western blot (B) for day 5 (D5), day 6 (D6), and non-infiltrated 2-week old E. tef seedling with a *N. benthamiana* positive control (NB)

5.3.2 Callus induction and plant regeneration

Following the above and prompted to find alternative solutions, embryonic calli were successfully induced from *E. tef* immature embryos and mature seeds (Figure 5-8), however, generating embryonic calli proved difficult to achieve and various other explant sources were subsequently tested. This variety of tef started producing immature embryos approximately 80-90 days post-sowing which negatively impacted the time available to test all the conditions and permutations. In addition, tef is known to stagger its flowering time and on the peduncles themselves, embryos towards to the apex of the peduncle were more mature than those towards the base. This too hindered the amount of harvestable immature embryos at any given time.

Time of emergence of potential calli from different explant sources are summarised in Table 5-3 below. Of the explant material tested, mature seeds and immature embryos from *E. tef* were the most successful explant material to use for callus induction. Of the two, mature seeds

proved the least time consuming process yielding potential calli three days post sowing onto callus induction media. However, many of the calli that formed from mature seeds were non-embryonic when plated on the CIM-IE media, however, when plated on the modified CIM-MS and subsequent propagation media, the majority of the developing calli were embryonic and nearly twice as large as those obtained from immature embryos. Following the 4-week induction period and subsequent 4-week proliferation on the CIM-MS calli had a mean weight of 68.09 ± 12.43 mg.

Table 5-3: Summary of callus induction explants showing approximate time for emergence of potential calli post-sowing

Species	Explant	Time for callus emergence post-sowing
E. tef	Leaf	Failed to emerge
	Root attached to seed	± 24 days
	Mature seed	± 3 days
	Immature embryo	± 90 days

The immature embryo calli by comparison to the mature seeds, took longer to differentiate into calli and once calli had formed, their growth was also much slower. Optimisation of the individual components, including but not limited to carbon source, concentration of 2,4D, addition of benzylamino purine, and addition of glutamine to callus induction media revealed that the final selection of components (Table 5-2) proved most reliable at initiating callus formation from immature embryos and mature seeds. Leaf and root tissue failed to develop any appreciable calli except when root tissue was still attached to the seed. However, cotyledons that developed from germinating seeds initiated callus formation when removed from seed derived callus and cultured onto callus induction media. In terms of embryogenesis, it was observed that the absence of BAP resulted in calli being predominantly non-embryonic when using immature embryos.

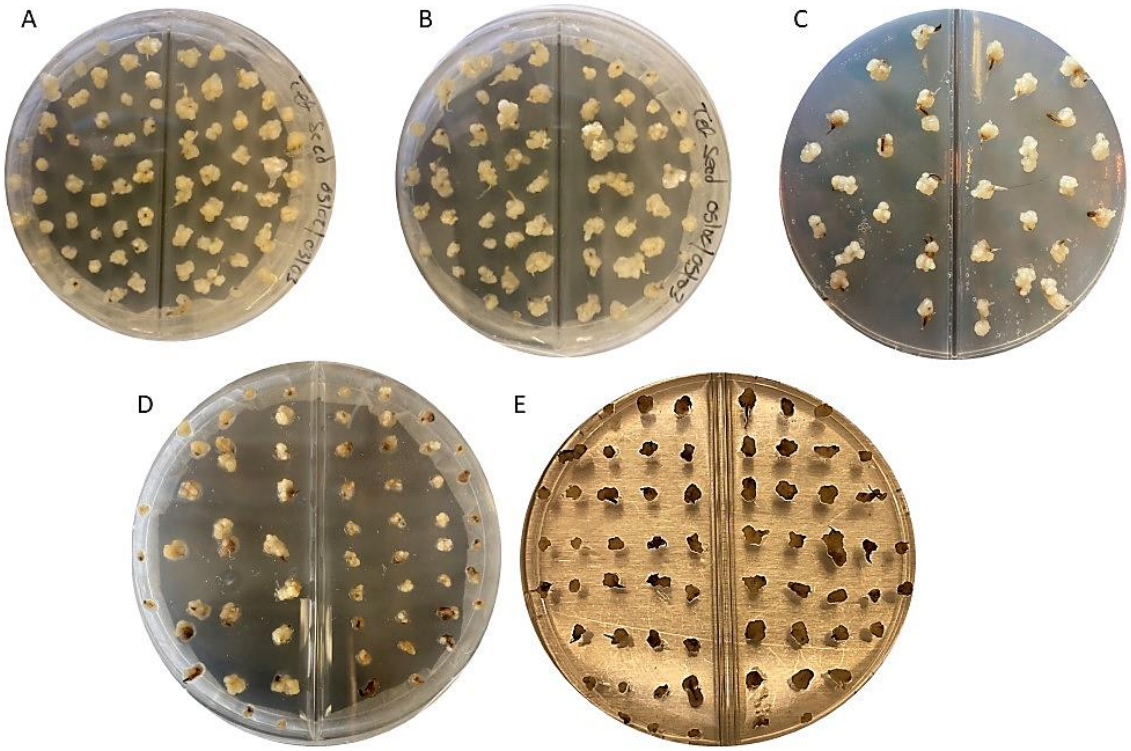


Figure 5-8: Calli from *E. tef* immature embryos and seeds. Callus induction from *E. tef* mature seeds following 4 weeks of culture on modified CIM-MS and then 4 weeks on proliferation CIM-MS (A and B). Callus induction of *E. tef* immature embryos on CIM-IE after approximately 4 weeks (C). Calli after agrobacterium-mediated transformation (D) and calli on selection media showing signs of necrosis (E) after approximately a week in culture.

Plant regeneration from initial immature embryo and mature seeds proved difficult to establish and to this end no successful adult plants have been regenerated. Embryonic calli did not appear to be responding to the presence of BAP alone and only initiated root formation when NAA was added to the media. The inclusion of activated charcoal also appeared to have made a difference in the response of the calli however, after several iterations and approximately 3 months of continuous sub-culturing, the regenerating calli failed to develop any pigmentation despite producing what appeared to be young shoot- and root-like structures as can be seen in Figure 5-9. The calli would initiate shoot- or root-like structures when cultured on PRM and would show signs of somatic conversion but would simply cease to grow beyond a certain point. The resurgence of Agrobacteria from co-cultivation (shown in panel E in Figure 5-9) also negatively impacted the regenerative capacity of the calli as the Agrobacterial growth overwhelmed the calli. Attempts were made to perturb the Agrobacterial contamination post-cultivation by including $AgNO_3$ and cefotaxime, but regular sub-culture was required with limited stock supply of cefotaxime. In most cases, the Agrobacteria became overwhelming for the calli and routine decontamination in an effort to remove the Agrobacteria is a likely agent in preventing proper somatic conversion.

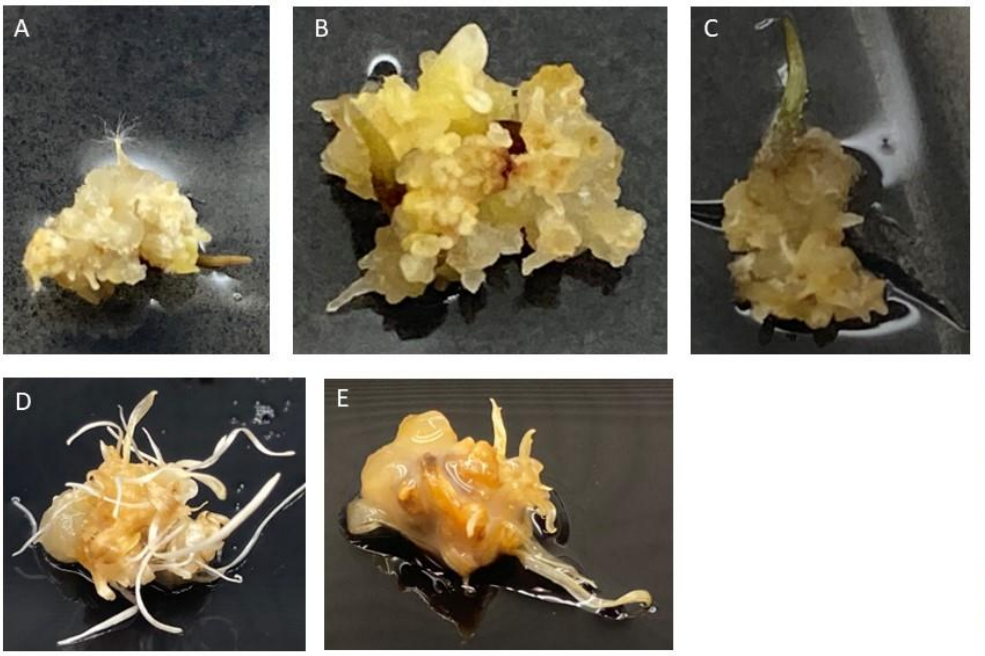


Figure 5-9: Plant regeneration from somatic embryos: Selection of initial regeneration from immature embryos (A-C) and mature seeds (D) of *E. tef*. Panel E is an after image of panel D following agrobacterium-mediated transformation with resurgence of Agrobacteria post-sterilization.

5.3.3 Callus co-cultivation

Following a callus transformation experiment using pRIC4 and pTRAC harbouring eGFP, total protein was isolated from *E. tef* and fluorescence was detected in all treatments (Figure 5-10) with pRIC4 in the *E. tef* yielding the highest amount. Fluorescence was also detected in the uninfected calli.

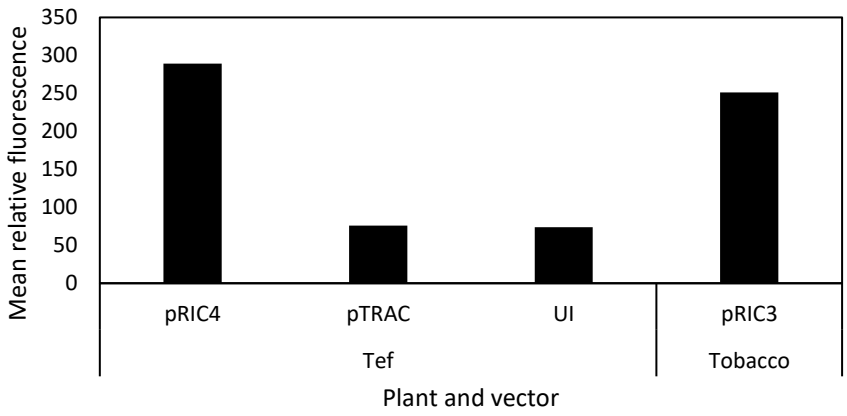


Figure 5-10: Relative eGFP fluorescence of infected calli: Relative fluorescence of pRIC4, pTRAC, and uninfected (UI) *E. tef* calli following co-cultivation with Agrobacteria with the respective vectors and subsequent protein isolation seven days post-infection. pRIC3 tobacco was used as a positive control.

The total protein was then also dotted on to a nitrocellulose membrane and a dot blot was performed using an anti-GFP primary and an anti-mouse horseradish peroxidase conjugated secondary antibody. The ECL detection system was used for the dot blots. The results indicate a strong signal observed in most of the samples and treatments. Between day seven and day nine a decrease in signal is noted. Presence of GFP was also detected in the uninfected sample. The samples were subsequently run on an SDS-PAGE for western detection however this failed to detect any GFP upon separation. The co-cultivation test was repeated twice with the same result implying that there is a signal in the uninfected calli which the primary antibody was able to detect. This suggested that a better method for transformation detection was required.

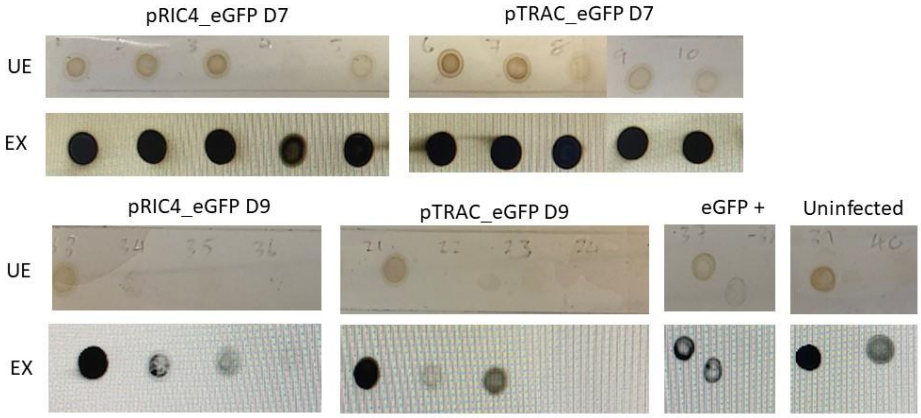


Figure 5-11: Dot blot for eGFP detection: Dot blot results following extraction of total protein from *E. tef* seven and nine days post-infection with pRIC4 and pTRAC harbouring eGFP. Anti-GFP primary antibody and an anti-mouse horseradish peroxidase conjugated secondary antibody was used for detection following ECL treatment. UE= unexposed, EX= exposed.

Using the overexpression construct made in Chapter 4 calli were transformed using the agrobacterium-mediated transformation method used for the viral vectors above. Immature-embryo-derived calli which had been transformed were sterilized by rinsing off the Agrobacteria before being plated on callus induction media for 2 days before being moved selection media supplemented with BASTA. The first attempt resulted in total browning of all calli when moved onto selection media (panel E, Figure 5-8) which suggested that the transformation had not been very successful. A second attempt was made this time opting to do selection during the regeneration stages to allow calli to establish themselves once more after agrobacterium-mediated transformation as it was not clear whether the browning was due to the selection or residual stress from Agrobacterial infection. This too resulted in browning of calli and despite attempts to perturb the resurgence of Agrobacteria, slimy growth quickly consumed the calli rendering the process and evaluation of transgenic integration difficult to establish. A final attempt was made using mature seed-derived calli which as noted above were on average much larger than immature embryo-derived calli. The localisation construct was used for this attempt on the basis that one could identify potential transgenics owing to mCherry fluorescence and not necessarily rely on herbicide selection. Following agrobacterium-mediated transformation, calli were sterilized as before and plated on CIM-SM media used for proliferation for two days before being moved onto 1 $mg.L^{-1}$ gibberellic acid supplemented proliferation media to induce somatic conversion. Calli were routinely sub-cultured and sterilized once during the first sub-culture. Calli were then kept on this media until evidence of somatic conversion were noted. Half of the calli were moved into the growth room (22°C, 16 h light, 8 h dark) while the remainder were kept in the dark. Approximately 3 weeks after being moved onto the gibberellic acid media no noticeable conversion was evident nor was there any increase in overall size when compared to pre-infection stage. Protein was isolated from pooled calli, and relative fluorescence determined as before with the same result. Empty vector calli had the same baseline fluorescence as the mCherry::EnHSP70-infected calli indicating that once more that determining successful transformation during callus stage was not feasible. Fluorescent microscopy on the eGFP calli from above and the mCherry::EnHSP70 indicated the same fluorescent signal between control and transgenic expression calli. Furthermore, the resurgence of Agrobacteria on several of the plates decimated the calli impacting the evaluation of transformation assessment.

Discussion

The primary goal of this chapter was to develop the relevant processes required to proceed with generating transgenic *E. tef*. It involved testing various method for transformation and ultimately led to the development of a tissue culture pipeline to generate embryonic callus

which is then meant to be co-cultivated with Agrobacteria harbouring the transgene and subsequently regenerated into plantlets.

5.4.1 Transient expression in *E. tef* leaves is not a feasible method.

The initial pilot study was aimed at determining whether *E. tef* could be infiltrated with agrobacterium using vacuum infiltration as routinely used within the Biopharming Research Unit. The goal was to determine whether this may pave the way for rapid transient transgene assessment before the implementation of stable transformation processes. While simple in its design and method, what works for one plant does not translate into success for another plant. Being a grass, *E. tef* does not produce large surface area leaves, compared to for example *Nicotiana benthamiana* (a species routinely used in BRU). This inherent decrease in surface area makes it difficult to increase potential infiltration areas. The second problem is the trade-off between age of plant and susceptibility to infection with that of biomass. Younger plants are generally 'easier' to infiltrate than older plants as they have not yet reached full maturity. While grasses lack the ability to initiate secondary meristematic growth, the older they become the tougher the leaves become, with many grasses developing hairy trichomes which can negatively impact leaf infiltration. In addition, gram for gram, more biomass is required to equate to that of *N. benthamiana* for example. This presented a problem as infiltration had to occur when the plants were relatively young but in so doing biomass is lost. Agrobacterium is known to naturally infect a suite of plants, hence its use for plant transformation, however monocotyledonous plants do not display the same susceptibility to Agrobacterial pathogenesis as outlined earlier. The results from the initial infiltration further corroborate this statement. While eGFP fluorescence was detected at day 5 post-infiltration (Figure 5-6) the lack of specific detection (Figure 5-7) brings into question the reliability of the measurements from the fluorometer. Given that no actual recombinant protein was detected it is safe to conclude that this particular method for transfecting *E. tef* did not work. It was the conflicting data that prompted the use of a more suited, well-established method for transforming monocotyledons.

5.4.2 Somatic embryogenesis from *E. tef* required exhaustive optimisation to establish.

One of the biggest problems of plant tissue culture, especially callus induction, is that there is no 'one-size-fits-all' method or recipe for inducing calli. The process requires a lot of optimisation of the various components and can rapidly escalate into a multi-permeated experiment. While there are a few reports of successful callus induction in *E. tef* (Mekbib, et al., 1996; Gebre, et al., 2013; Gugsa & Kumlehn, 2011) and close relative *E. curvula* (Viviana, et al., 2001; Ncanana, et al., 2005) from different explants, the particular variety used here

required extensive optimisation. Something as simple as the carbon source is for instance not consistent among the aforementioned studies. For this reason, a considerable amount of time was spent determining which combination of components yielded calli in the shortest time possible and was applicable to an array of explants. From this, it was observed that 2% sucrose resulted in accelerated browning of the calli owing to phenolic accumulation. For this reason, the media was further supplemented with 3% maltose as there is evidence to suggest that maltose is less likely to cause phenolic accumulation (Kumar, et al., 2015). Furthermore, since the addition of the ascorbic acid and citric acid was shown to markedly reduce the accumulation of phenolics in litchi (Nishchal, et al., 2018) these were thus also included. By changing the carbon source and including 150 $mg.L^{-1}$ ascorbic and 150 $mg.L^{-1}$ citric acid, the prevalence of browning owing to phenolic accumulation was noticeably reduced. Glutamine was an additional amino acid used by Gugsa & Kumlehn, (2011) who stated that 'optimal results' for callus induction from their two varieties of *E. tef* immature embryos included 6 mM glutamine. From what was seen with the callus induction from the SA Brown *E. tef* immature embryo used herein was that glutamine hindered the development of the calli and often resulted in white precipitation in the media. The addition of glutamine is not without its merits. When included in callus induction media, it can serve as a nitrogen source as well as an additional carbon source and plays an important role in a number of biochemical processes such as electron transport. The problem with using glutamine is that it can also readily become toxic through the production of ammonia. For these reasons, and based on observations, glutamine was omitted from the callus induction media. Amino acids were however, supplied with the addition of 300 $mg.L^{-1}$ casein hydrolysate. The inclusion of $1mg.L^{-1}$ BAP was based on the method described by Ncanana, et al., (2005). Upon further observation of callus induction and growth, it became apparent that including BAP resulted in delayed induction and subsequent establishment and growth of calli. However, failure to include BAP in the media resulted in a much higher proportion of calli being non-embryonic. A large setback during the establishment of the tissue culture pipeline was the sporadic instances of contamination. Most common of this was the emergence of furry black mould or bacterial contamination from within the media itself. Setbacks such as this during optimisation often resulted in not being able to fully test certain permutations and having to restart the entire process. While it is difficult to say with absolute certainty whence the sporadic contamination came, it has been speculated that latent fungal spores were dislodged during adjacent building works which caused vibrations in the building in which the work was done. Once building works were concluded and an improved decontamination protocol was established, occurrence and resurgence of contamination was significantly reduced.

The overall process of establishing a callus culture is lengthy and while trying to optimise the conditions while maintaining sterility it becomes difficult to determine with certainty that a particular permutation works or does not. Unlike model organisms such as tobacco or maize with well-established methods, working with non-model organisms inherently presents multiple problems that require optimisation. For instance, the use of mature tef seeds quickly yielded calli, however, they readily developed into non-embryonic calli (data not shown). For the sake of transformation, the calli must be embryonic since non-embryonic calli lack the capacity to regenerate into plantlets. While the use of mature seeds seemed most attractive, the only explant that readily developed into embryonic calli was immature embryos until the media composition was optimised for mature seeds to yield a greater proportion of embryonic calli. Immature embryos too were not without their problems. The first setback encountered was the fact that this variety took nearly 80 days to initiate flowering for immature embryo harvesting, though luckily, once anthesis was initiated it would continue to produce inflorescence for several weeks to months. Secondly, the rate of callus induction and proliferation was exceptionally low for the immature embryos. For instance, out of 200 immature embryos plated onto the optimised CIM-IE, approximately 20-30% would develop into callus, of which 60-80% would be embryonic. The modified mature seed method described above worked exceptionally well. Callus initiation was rapid and once moved onto callus proliferation media the calli doubled in size and were predominantly embryonic. Sub-culture onto gibberellic-supplemented media was proposed by Kebebew, et al., (1998) to convert somatic embryos into developing plantlets. The authors however failed to disclose the final media composition of this sub-culture step as well as the length of time required for calli to be maintained on this media. Seed-derived calli from this method were transfected with the mCherry::EnHSP70 construct generated in Chapter 4 and were moved onto gibberellic-supplemented media. They were initially kept in the dark for approximately two weeks and sub-cultured routinely to reduce the resurgence of Agrobacteria. Of the 65 calli each which were infected with either the mCherry::EnHSP70 construct or the empty vector pB2WG7, only 1 and 2 respectively showed signs of conversion after approximately three weeks of culture. These were then moved onto the recommended plant regeneration media from Kebebew, et al., (1998) but failed to develop any further.

The regeneration of embryonic calli required extensive optimisation mostly due to the fact that the process leading up to regeneration required extensive optimisation. Upon moving calli onto regeneration media, it took approximately 2 months for the calli to initiate root formation. Interestingly, under these conditions, the calli remained unpigmented during this culture period. Owing to limited availability of suitable plant tissue culture environments such as controlled growth chambers, the regeneration had to occur in a regular growth room used for

the initial growth of tef seedlings. It is likely that by not establishing a controlled environment, the regeneration process took much longer than it ought to have. As noted in the results, the regenerated calli obtained from immature embryos simply did not develop into plants which naturally impacts the establishment of a transformation and regeneration protocol for tef and maintaining calli on media for too long inherently increases the risks of contamination occurring. For example, the attempt made to establish a liquid cell suspension culture took approximately 8 weeks, requiring weekly media replenishment. While this appeared to be working in that smaller clusters of embryonic calli were developing at about 7 weeks of culture, the media were contaminated by fungus and the experiment had to be abandoned. The use of other explant sources such as immature cotyledons and roots appeared at first to be suitable candidates to shorten the time required from immature embryos but the capacity of these explant sources to initiate and maintain callous tissue could not be established. Root material still anchored to the original seed seemed most susceptible to callus induction and produced callus but sub-culture of these resulted in failure to propagate sufficiently. Mature seeds once plated on CIM-MS readily formed callus though at first glance these too appeared to be predominantly non-embryonic as exemplified by appearance of a milky-white glossy appearance. Some of them however developed calli which were more compact in appearance. Once these were moved to propagation media which contained a lower amount of 2,4D and included BAP, the calli further differentiated into embryonic calli.

While the viral vectors routinely used by BRU are well suited for tobacco, their use in callus transformation became a contentious one. The primary aim was to use these vectors to evaluate the transformation efficiency of Agrobacteria-mediated co-cultivation and determine whether their use in transient recombinant protein production in tobacco could be used to assess the feasibility of transgene expression for improved abiotic tolerance in calli. Several problems were encountered which ultimately led to exclusion of these vectors for subsequent study. One such problem was the transient nature of the vectors having peak expression sometime between 5-9 days post-infection. This is of course optimal for the production of recombinant protein in tobacco but not optimal for callus transformation since the culture time required to firstly ensure complete removal of Agrobacteria and secondly selection is much greater than this peak expression period. Secondly, the issue with biomass became very apparent when trying to isolate protein from infected calli. On average, one callus weighed 3 – 5 mg yielding total protein in the nanogram range. The volume of extraction buffer had to be significantly reduced so as to obtain a crude sample with a sufficient amount of protein. The detection of the eGFP visually was suggestive of successful transformation, however, it became difficult to validate the presence of eGFP. The relative fluorescence readings across treatments revealed that while there was more fluorescence in the infected calli some

fluorescence was still being detected in the uninfected calli. Upon closer inspection by viewing the infected and non-infected calli under a fluorescent microscope, the degree of auto-fluorescence from non-infected calli made it impossible to differentiate any true eGFP fluorescence from the infected calli from that of auto-fluorescence. The most conclusive result was that when probed with an anti-GFP primary antibody, no detection was observed in the infected calli. It was also discovered much later that pEAQ-HT, selected on the basis of it being suitable for stable and transient transformation, had a P-19 silencer which could hinder the regeneration of callus. While it would be erroneous to state that the use of the viral vectors did not perform as intended, it is valid to conclude that attempting to do any molecular analysis on undifferentiated callus is not recommended, especially when one is limited by the number of receptive callus to have a sufficient sample size.

Given that the culture conditions for callus induction was to some degree established and the use of the viral vectors had been determined to not be suitable, the next step was to transform the calli with a more suitable expression vectors. The overexpressing EnHSP70 construct and empty pB2WG7 generated in Chapter 4 was used to infect embryonic calli derived from *E. tef* immature embryos. The calli were co-cultured with Agrobacteria and then plated on co-cultivation media. The calli were incubated for a week on this media before being moved onto selection media supplemented with BASTA. They were cultured on selection media for one week before being moved onto regeneration media. Initial observations suggested a very low transformation efficiency as most of the calli started browning on BASTA supplemented media. A second attempt at this yielded greater promise as it was decided to only include BASTA during the regeneration stage, but as noted earlier, the resurgence of Agrobacteria and lack of sufficient regeneration during plant regeneration rendered this experiment inconclusive. A final attempt was made using calli derived from the modified mature seed media. This too failed to provide any conclusive evidence of successful transformation. Calli infected with these had significantly reduced overall growth when moved onto the gibberellic media. The 3 calli as noted above which had shown signs of regeneration failed too to accumulate any pigmentation and did not resume growth.

Conclusion

The transformation of monocotyledonous plants requires the establishment of a callus induction and regeneration pipeline. Orphan cereals such as tef are less studied and methods for their genetic transformation are not that well characterised or developed. Initial transient evaluation of transgene expression in leaves indicated that some expression could be

detected but it is not a feasible method for evaluating whether transgene expression could impart any phenotypic advantage. The establishment of a reproducible method for callus induction and proliferation presented herein firmly lays the foundation for continued investigation as to the conditions required for proper propagation of transformed embryonic calli for enhanced abiotic stress tolerance. While it should be noted that although within the time frame allotted plants could not be generated based on previously reported methods from embryonic calli as for instance reported by Kebebew, et al., (1998) and Gebre, et al., (2013), considerable progress has been made towards the establishment of a method for the transformation of *E. tef*.

Chapter 6: Concluding thoughts and future work.

Vegetative desiccation tolerance is a fascinating state and the implications from understanding the molecular mechanisms that underpin it are even more fascinating. Over the past few decades and with growing precision in mass-spectrometry and DNA sequencing, researchers from across the globe working on a myriad of resurrection plants have all added to our fundamental understanding of the drivers of vegetative desiccation tolerance, yet we are still nowhere closer to having the full picture (Tebele, et al., 2021). The most widely accepted paradigm of vegetative desiccation tolerance is its resemblance to orthodox seed maturation in that resurrection plants, *inter alia,* accumulate carbohydrates and thermo-tolerant proteins such as LEAs and heat shock proteins (Costa, et al., 2017a,b). This is further exemplified by the lack of these responses in sensitive plants and implicates a switch in vegetative tissues to adopt seed maturation processes which are simply not switched on in sensitive plants. Dehydration stress imposes great strain on the cellular environment where for instance photosynthesis becomes lethal through the generation of reactive oxygen species which attack lipids, proteins, and genetic material. Methods used by resurrection plants to combat this is to either shut down photosynthesis completely which is a strategy used by poikilochlorophyllous resurrection plants, or the accumulation of massive chloroplastic protection (Charuvi, et al., 2015; Zia, et al., 2016) and anthocyanins as seen in the homoiochlorophyllous resurrection plants. Accumulation of an antioxidant response to sequester and dampen ROS-mediated oxidation is highly upregulated in all resurrection plants and several folds greater when compared to sensitive plants. The accumulation of intrinsically disordered proteins such as LEAs and thermostable proteins such as HSPs, again several folds greater compared to sensitive plants, is yet another method commonly used.

The complemented proteomic study done herein demonstrated firstly that while transcriptomic studies are a great source of indispensable knowledge, conclusions drawn from them must be handled carefully. The data presented here clearly demonstrated that while there are tens of thousands of genes that are perceived as differentially expressed from the transcriptome, only a small subset of those are eventually differentially translated. While it should be noted that owing to differences in annotations between the two data sets inherently limits the comparative capacity, the subset of proteomic-matching transcripts illustrates that in some instances transcripts are immediately translated while others are delayed. Generally, most of the transcripts were translated at the same RWC suggesting for the most part that there is very little translational delay which is in contrast to that observed in *C. plantagineum* (Xu, et al., 2021). While the observation herein of little delayed translation does not invalidate those by Xu et al., (2021), it might indicate a quicker turnover of mRNA to necessary proteins in *E.*

nindensis as compared to *C. plantagineum* which could possibly be attributed to the morphological differences between the two species. However, the fact that fewer proteins showed differential abundance points to greater translational control over transcripts and that perhaps the production of a large number of transcripts is a means for *E. nindensis* to decide which are more favourable to translate. There is evidence to suggest that it might lie in differential methylation of nascent mRNA where some transcripts are tagged for repression while others are tagged for translation. The same can be said for the protein pool within the cell where ubiquitinated-mediated proteolysis implies selective degradation of proteins. A reasonable assumption to make is that these are likely photosynthetic proteins since *E. nindensis* dismantles its thylakoids upon dehydration stress. In addition, these proteins could also be from damaged or misfolded proteins as a result of oxidative stress. Shifts in metabolism indicate that *E. nindensis* favours degradation processes as means to recycle and remobilise those constituents. While a much more detailed study of all the differentially abundant proteins, which could not be done here owing to space limitations, is required to fully appreciate the proteomic nature of *E. nindensis*, the study demonstrated that there were clear differences between the non-senescent desiccation-tolerant tissue when compared to the senescent desiccation-sensitive tissue. A hypothesis presented here is that small and subtle changes of proteins, for instance raffinose synthesis in the NST exclusively, are what ultimately leads to desiccation survival. Another hypothesis is that all ST were desiccation tolerant at the onset of their growth but lose this ability as they mature. It is no secret that many of the vegetative desiccation tolerant plants are considerably small, with the exception of *Myrothamnus flabellifolia* which is the only woody resurrection plant. This small size has one large advantage in that it is easier regulate a few thousand cells compared to a few billion and this argument can be extended when considering *E. nindensis* in a natural environment. Owing to similar protective responses between the NST and ST, a hypothesis is that *E. nindensis* simply sacrifices the ST upon rehydration as a means to turn it into a source of nutrients for the resurrecting NST. Thus, it would be worthwhile to investigate the source-to-sink mechanics between these two tissues perhaps though isotope or fluorescent labelling. In addition, there is evidence to support not only ABA-mediated regulation of desiccation response but also jasmonate, salicylic acid, and ethylene and it would be worthwhile to investigate the hormonal signatures of *E. nindensis*, something which has never been done in any resurrection plant to date.

The subset of selected senescence-related protein in Chapter 3 illustrated that there was much greater suppression of senescence in the NST than in the ST. Some of these proteins were functionally assigned to belonging to chloroplast development while others were involved in removal of post-translational protein modifications in the NST. This selection of proteins and

transcription factors narrowed the range of potential candidates for bioengineering. These candidate genes can be induced or silenced in economically important crops such as maize and orphan crops such as teff to supress the early onset of senescence.

During dehydration there is great crowding occurring where biomolecules are forced to interact with one another when under normal conditions they would not. This forced interaction can lead to the formation of spontaneous de-mixing via the formation of membraneless compartments through liquid-liquid phase separation (LLPS). Chapter 4 illustrated that a selected HSP70 from *E. nindensis* was capable of undergoing LLPS, something which has been well-characterised in the other seed-related thermotolerant chaperone, LEA. The ability of it to undergo LLPS is highly suggestive of it functioning as means to sequester biomolecules and could facilitate biochemical activity within these condensates. Further characterisation of the biomolecules with which EnHSP70 can interact is needed to further expand on this hypothesis. In addition, only one of the many of HSP70s from *E. nindensis* was studied, and as such, it would be worthwhile to investigate whether all of the isoforms of HSP70 from *E. nindensis* can undergo LLPS. The studied EnHSP70 was shown to localise to the chloroplast and transcriptomic data indicated that it was not significantly upregulated. In fact, this gene was completely absent in the transcriptome which only became evident after much of the work had been completed. Bioinformatic analysis indicated that it belonged to the subgroup of isoform 4. Two of the greatest upregulated genes from the transcriptome were denoted as isoform 4 and sequence similarity indicated that the *EnHSP70* used herein was closely related. Its localisation to chloroplast suggests that it is mainly involved in the chaperoning of chloroplastic proteins destined for degradation. HSP70 is a universal stress protein and sequence comparison between *E. nindensis*' *HSP70* and other sensitive plant species indicated conserved areas but also small differences in amino acid composition, suggesting that perhaps isoforms of HSP70, while evolutionary conserved, might perform better than those from sensitive plants in stress response. The protein however appeared to have what can be seen as negative effects when overexpressed in *A. thaliana* and BL21 *E. coli*. For Arabidopsis it was noted that germination was negatively affected to the extent that seeds would take on average 1-2 weeks longer to germinate than the empty vector control seeds. While it was noted that during the drying of the T_2 overexpressing plants their leaves remained greener and did not succumb to rapid wilting when compared to the empty vector plants, the expression observed in these plants suggested that it was not fully over-expressed. In light of this and the observation that EnHSP70 did not enhance salt or mannitol stress in BL21 *E. coli* cells nor imparted improved thermotolerance to lactate dehydrogenase it was decided that pursuing the evaluation of *EnHSP70* overexpression in Arabidopsis was not feasible. Attempts at present are being made to isolate T_3 lines to explore the latent seed germination phenotype

in more detail. Thus, based on its localisation and its possible role in *E. nindensis*, we drew the conclusion that it is firstly not stress-inducible and secondly that its primary role is most likely involved with maintaining proteostasis during the early stages of thylakoid disassembly in *E. nindensis* where it chaperones proteins from this disassembly towards proteolysis sites.

A primary goal of our research group is to take information from the study of resurrection plants and apply them to the improvement of commercially important crops and underutilised but regionally important orphan crops. Great progress in the understanding of the key drivers of desiccation tolerance has been made and attempts were made herein to genetically transform the Ethiopian orphan crop *E. tef*, which is an already drought tolerant close relative to *E. nindensis*. The process of *E. tef* transformation was severely impacted by the lack of suitable plant tissue culture facilities within the department, however, a reproducible method for generating somatic embryos from immature embryos and mature seeds was developed. Further experimentation is required to find the most optimal culture and media conditions to convert somatic embryos into plantlets. The inability to establish an appropriate culture condition for the regeneration of embryonic calli negatively impacted the progress of *E. tef* transformation and the efficiency of agrobacterium-mediated transformation of calli could not be evaluated. Nevertheless, the methods and procedures developed herein lays the foundation for future expansion and testing.

Taken together, this thesis sought to understand what the proteomic drivers of desiccation tolerance in *E. nindensis* were and whether it corroborates and validates observations seen in the transcriptome. It also sought to functionally characterise a highly conserved protein to investigate its potential application to crop improvement. Lastly, this thesis sought to develop a method for transformation of *E. tef* to bridge the gap between discovery omics and application. To this end, the primary goals set out were accomplished to varying success and ultimately paves the way for future discovery omics and direct application.

References

Abobatta, W F. 2018. “Drought adaptive mechanisms of plants – a review.” Advances in Agriculture and Environmental Science 2 (1): 42-45. doi: 10.30881/aaeoa.00021.

An, J U, I G Lee, Y J Ko, and D K Oh. 2019. “Microbial Synthesis of Linoleate 9S-Lipoxygenase Derived Plant C18 Oxylipins from C18 Polyunsaturated Fatty Acids.” American Chemical Society 67 (11): 3209-3219. doi: https://doi.org/10.1021/acs.jafc.8b05857.

Anum, J, C O'Shea, M Z Hyder, S Farrukh, K Skriver, S I Malik, and T Yasmin. 2022. “Germin like protein genes exhibit modular expression during salt and drought stress in elite rice cultivars.” Molecular Biology Reports 49 (1): 293-302. doi:10.1007/s11033-021-06871-3 .

Arent, S, V E Pye, and A Henrisksen. 2008. “Structure and function of plant acyl-CoA oxidases.” Plant Physiology and Biochemistry 46 (3): 292-301. doi:https://doi.org/10.1016/j.plaphy.2007.12.014.

Arias, C L, T Pavlovic, G Torcolese, M B Badia, M Gismondi, V G Maurino, C S Andreo, M F Drincovich, M C Gerrard Wheeler, and M Saigo. 2018. “NADP-Dependent Malic Enzyme 1 Participates in the Abscisic Acid Response in *Arabidopsis thaliana*.” Frontiers in Plant Science 9. doi:https://doi.org/10.3389/fpls.2018.01637.

Armenteros, J JA, C K Sønderby , S K Sønderby , H Nielsen, and O Winther. 2017. “DeepLoc: prediction of protein subcellular localization using deep learning.” Bioinformatics 33 (21): 3387-3395. doi:https://doi.org/10.1093/bioinformatics/btx431.

Armenteros, J JA, M Salvatore, O Emanuelsson, O Winther, G von Heijne, A Eloffson, and H Nielsen. 2019. “Detecting sequence signals in targeting peptides using deep learning.” Life Science Alliance 30 (2). doi:10.26508/lsa.201900429.

Arce, D P, A V Godoy, K Tsuda, KI Yamazaki, E M Valle, M J Iglesias, M F Di Mauro, and C A Casalongue. 2010. "The analysis of an Arabidopsis triple knock-down mutant reveals functions for MBF1 genes under oxidative stress conditions." Journal of Plant Physiology 167 (3): 194-200. doi:https://doi.org/10.1016/j.jplph.2009.09.003.

Artur, M A, M C Costa, J M Farrant, and H WM Hilhorst. 2019a. “Genome-level responses to the environment: plant desiccation tolerance.” Emerging Topics in Life Sciences 3: 153-163. doi:https://doi.org/10.1042/ETLS20180139.

Artur, M A.S, J Rienstra, T J Dennis, J M Farrant, W Ligterink, and H Hilhorst. 2019. "Structural Plasticity of Intrinsically Disordered LEA Proteins from Xerophyta schlechteri Provides Protection In Vitro and In Vivo." Frontiers in Plant Science 10. doi:https://doi.org/10.3389/fpls.2019.01272.

Ashraf, M, and S Mehmood. 1990. “Response of four Brassica species to drought stress.” Environmental and Experimental Botany 30 (1): 93-100. doi:ttps://doi.org/10.1016/0098-8472(90)90013-T.

Ashraf, M, Q Mao, J Hong, L Shi, X Ran, F Liaquat, M Uzair, W Liang, A R Fernie, and J Shi. 2021. “HSP70-16 and VDAC3 jointly inhibit seed germination under cold stress in Arabidopsis.” Plant, Cell & Environment 44 (11): 3616-3627. doi:https://doi.org/10.1111/pce.14138.

Assefa, K, J K Yu, G Belay, H Tefera, and M E Sorrells. 2011. “Breeding tef [*Eragrostis tef* (Zucc.) trotter]: conventional and molecular approaches.” Plant Breeding 130 (1). doi:https://doi.org/10.1111/j.1439-0523.2010.01782.x.

Augustine, S M, J A Narayan, D S Syamaladevi, C Appunu, M Chakravarthi, V Ravichandran, and N Subramonian. 2015. “*Erianthus arundinaceus* HSP70 (EaHSP70) overexpression increases drought and salinity tolerance in sugarcane (Saccharum spp. hybrid).” Plant Science 232: 23-34. doi:https://doi.org/10.1016/j.plantsci.2014.12.012.

Baker, J C, N M Mock, B D Whitaker, D P Roberts, C P Rice, K L Deahl, and A Aver'Yanov. 2005. “Involvement of acetosyringone in plant-pathogen recognition.” Biochemical and Biophysical Research Communications 328 (1): 130-136. doi:10.1016/j.bbrc.2004.12.153.

Balsamo, R A, C Vander Willigen, A M Bauher, and J Farrant. 2006. “Drought Tolerance of selected Eragrostis species correlates with leaf tensile properties.” Annals of Botany 985-991. doi:10.1093/aob/mcl068.

Bandurska, H, and A Stroinski. 2005. “The effect of salicylic acid on barley response to water deficit.” Acta Physiologiae Plantarum volume 27: 379-386. doi:https://doi.org/10.1007/s11738-005-0015-5.

Bardou, P, J Marieete, F Escudié, C Djemiel, and C Klopp. 2014. “jvenn: an interactive Venn diagram viewer.” BMC Bioinformatics. doi:https://doi.org/10.1186/1471-2105-15-293.

Barrat, D H.P, P Derbyshire, P Findlay, K Pike, M Wellner, N Lunn, R Feil, C Simpson, A J Maule, and A M Smith. 2009. “Normal growth of Arabidopsis requires cytosolic invertase but not sucrose synthase.” Proceedings of the National Academy of Sciences 106 (31): 13124-13129. doi:https://doi.org/10.1073/pnas.0900689106.

Berka, M, R Kopecká, V Berková, B Brzobohatý, and M Černý. 2022. “Regulation of heat shock proteins 70 and their role in plant immunity.” Journal of Experimental Botany 73 (7): 1894-1909. doi:https://doi.org/10.1093/jxb/erab549.

Bettencourt, B R, M E Feder, and S Cavicchi. 1999. “Experimental evolution of HSP70 expression and thermotolerance in Drosophila melanogaster.” Evolution 53 (2): 484-492. doi:https://doi.org/10.1111/j.1558-5646.1999.tb03783.x.

Bharath, P, S Gahir, and A S Raghavendra. 2021. “Abscisic Acid-Induced Stomatal Closure: An Important Component of Plant Defense Against Abiotic and Biotic Stress.” Frontiers in Plant Science 12. doi: https://doi.org/10.3389/fpls.2021.615114.

Bhaskara, G B, T T Nguyen, and P E Verslues. 2012. “Unique Drought Resistance Functions of the Highly ABA-Induced Clade A Protein Phosphatase 2Cs.” Plant Physiology 160 (1): 379-395. doi:https://doi.org/10.1104/pp.112.202408.

Bieniawska, Z, D H.P Barrat, A P Garlick, V Thole, N J Kruger, C Martin, R Zrenner, and A M Smith. 2007. “Analysis of the sucrose synthase gene family in Arabidopsis.” The Plant Journal 49 (5): 810-828. doi:https://doi.org/10.1111/j.1365-313X.2006.03011.x.

Bilova, T, G Paudel, N Shilyaev, C Birkemeyer, L A Wessjohann, and A Frolov. 2017. "Global proteomic analysis of advanced glycation end products in the Arabidopsis proteome provides evidence for age-related glycation hot spots." Journal of Biological Chemistry 292 (38): 15758-15776. doi:https://doi.org/10.1074/jbc.M117.794537.

Blomstedt, C K, C A Griffiths, D P Frederick, J D Hamill, D F Gaff, and A D Neale. 2010. “The resurrection plant *Sporobolus stapfianus*: An unlikely model for engineering enhanced plant biomass?” Plant Growth Regulation 62 (3): 217-231. doi:https://doi.org/10.1007/s10725-010-9485-6.

Bonnot, T, M B Gillard, and D H Nagel. 2019. “A Simple Protocol for Informative Visualization of Enriched Gene Ontology Terms.” Bio-101. doi:10.21769/BioProtoc.3429.

Boorstein, W R, T Zeigelhoffer, and E A Craig. 1994. “Molecular evolution of the HSP70 multigene family.” Journal of Molecular Evolution 38 (1): 1-17. doi: 10.1007/BF00175490..

Boyd, R S, and S N Martens. 1998. “The significance of metal hyperaccumulation for biotic interactions.” Chemoecology 8 (1): 1-7. doi:10.1007/s000490050002.

Bresson, J, S Bieker, L Riester, J Doll, and U Zentgraf. 2018. “A guideline for leaf senescence analyses: From quantification to physiological and molecular investigations.” Journal of Experimental Botany 69 (4): 769-786. doi:https://doi.org/10.1093/jxb/erx246.

Brito, V C, C P de Almeida, R R Barbosa, M G.A Carasio, A G Ferreira, L G Fernandez, R D de Castro, H Hilhorst, W Ligterink, and P R Ribeiro. 2020. “Overexpression of *Ricinus communis* L. malate synthase enhances seed tolerance to abiotic stress during germination.” Industrial Crops and Products 148. doi:https://doi.org/10.1016/j.indcrop.2020.112110.

Buchaupt, M, F Kähne, M M.W Etchmann, and J Schrader. 2014. “Biotechnological Production of Fatty Aldehydes.” In Flavour Science, edited by V Ferreira and R Lopez, 195-199. Academic Press. doi:doi.org/10.1016/B978-0-12-398549-1.00037-4.

Bürglin, T. 2001. “Homeobox.” In Encyclopaedia of Genetics, edited by S Brenner and J H Miller, 958-962. Academic Press. doi:https://doi.org/10.1006/rwgn.2001.0625.

Campell, M K, and S O Farrell. 2012. “Photosynthesis.” In Biochemistry. Brooks/Cole Cengage Learning.

Chan, M T, T M Lee, and H H Chang. 1992. “Transformation of Indica rice (*Oryza sative* L.) mediated by agrobacterium tumefaciens.” Plant and Cell Physiology 33 (5): 577-583.

Charuvi, D, R Nevo, E Shimoni, L Narvey, A Zia, J M Farrant, H Kirchoff, and Z Reich. 2015. “Photoprotection conferred by changes in photosynthetic protein levels and organization of a homoiochlorophyllous resurrection plant.” Plant Physiology 167: 1554-1565. doi:10.1104/pp:11425594.

Chen, J C.F, C C.Y Tsai, and J T.C Tzen. 1999. “Cloning and Secondary Structure Analysis of Caleosin, a Unique Calcium-Binding Protein in Oil Bodies of Plant Seeds.” Plant and Cell Physiology 40 (10): 1079-1086. doi:https://doi.org/10.1093/oxfordjournals.pcp.a029490.

Chen, J, W-Q Li, and Y-X Jia. 2020. “The Serine Carboxypeptidase-Like Gene SCPL41 Negatively Regulates Membrane Lipid Metabolism in *Arabidopsis thaliana*.” Plants 9 (6). doi:doi.org/10.3390/plants9060696.

Chen, L, R Gómez, and L C Weiss. 2021. “Distinct Gene Expression Patterns of Two Heat Shock Protein 70 Members During Development, Diapause, and Temperature Stress in the Freshwater Crustacean Daphnia magna.” Frontiers in Cell Developmental Biology 9. doi:https://doi.org/10.3389/fcell.2021.692517.

Chen, L, Z Wu, and S Hou. 2020. “SPEECHLESS Speaks Loudly in Stomatal Development.” Frontiers in Plant Science 11. doi:https://doi.org/10.3389/fpls.2020.00114.

Chen, W, J Huang, S Chen, L Zhang, J-D Rochaix, L Peng, and Q Xin. 2022. “Stromal Protein Chloroplast Development and Biogenesis1 Is Essential for Chloroplast Development

and Biogenesis in Arabidopsis thaliana.” Frontiers in Plant Science 13. doi:https://doi.org/10.3389/fpls.2022.815859.

Chen, Y, X Lei, Z Jiang, and K Fitzgerald. 2021. “Cellular nucleic acid–binding protein is essential for type I interferon–mediated immunity to RNA virus infection.” PNAS 118 (26). doi:https://doi.org/10.1073/pnas.210038311.

Cheng, X, X Xiao, and K-C Chou. 2017. “pLoc-mPlant: predict subcellular localization of multi-location plant proteins by incorporating the optimal GO information into general PseAAC.” Molecular Biosystems 13 (9): 1722-1727. doi:10.1039/c7mb00267j .

Chia, T Y, M J Pike, and S Rawsthorne. 2005. “Storage oil breakdown during embryo development of *Brassica napus* (L.).” Journal of Experimental Botany 56 (415): 1285-1596. doi:https://doi.org/10.1093/jxb/eri129.

Chicas, A, and G Macino. 2001. “Characteristics of post-transcriptional gene silencing.” EMBO Reports 2: 992-996. doi: 10.1093/embo-reports/kve231.

Cho, E K, and C B Hong. 2005. “Over-expression of tobacco NtHSP70-1 contributes to drought-stress tolerance in plants.” Physiology and Biochemistry 25: 349-358. doi:https://doi.org/10.1007/s00299-005-0093-2.

Choudhury, R, and S Pandey. 2016. “The role of PLDα1 in providing specificity to signal-response coupling by heterotrimeric G-protein components in Arabidopsis.” Plant Journal 86: 50-61. doi:10.1111/tpj.13151 .

Choudhury, S, P Panda, L Sahoo, and S K Panda. 2013. “Reactive oxygen species signalling in plants under abiotic stress.” Plant Signalling & Behaviour 8. doi:http://dx.doi.org/10.4161/psb.23681.

Christ, B, A Egbert, I Süssenbacher, B Kräutler, D Bartels, S Peters, and S Hörtensteiner. 2014. “Water deficit induces chlorophyll degradation via the “PAO/phyllobilin” pathway in leaves of homoio- (*Craterostigma pumilum*) and poikilochlorophyllous (*Xerophyta viscosa*) resurrection plants.” Plant, Cell and Environment 37 (11): 2521-2531. doi:https://doi.org/10.1111/pce.12308.

Collett, H, A Shen, M Gardner, J M Farrant, K J Denby, and N Illing. 2004. “Towards transcript profiling of desiccation tolerance in *Xerophyta humilis*: Construction of a normalized 11 k X. humilis cDNA set and microarray expression analysis of 424 cDNAs in response to dehydration.” Physiologia Plantarum 122 (1): 39-53. doi:https://doi.org/10.1111/j.1399-3054.2004.00381.x.

Corpas, F J, and J B Barroso. 2014. “NADPH-generating dehydrogenases: their role in the mechanism of protection against nitro-oxidative stress induced by adverse environmental conditions.” Frontiers in Environmental Science 2. doi: https://doi.org/10.3389/fenvs.2014.00055.

Costa, M C D, K Cooper, H W M Hilhorst, and J M Farrant. 2017a. “Orthodox seeds and resurrection plants: two of a kind?” Plant Physiology. doi:https://doi.org/10.1104/pp.17.00760.

Costa, M C D, M A S Arthur, J Maia, E Jonkheer, M F L L Derks, H Nijveen , and H W Hilhorst. 2017b. “A footprint of desiccation tolerance in the genome of *Xerophyta viscosa*.” Nature Plants. doi:https://doi.org/10.1038/nplants.2017.38.

Costa, M C, J M Farrant, M J Oliver, W Ligterink, J Buitink, and H M.W Hilhorst. 2016. “Key genes involved in desiccation tolerance and dormancy across life forms.” Plant Science 251: 162-168. doi:https://doi.org/10.1016/j.plantsci.2016.02.001.

Culver, J A, and M Mariappan. 2020. “Membrane Protein Biogenesis: PAT Complex Pats Membrane Proteins into Shape.” Current Biology 30 (22): 1387-1389. doi:doi.org/10.1016/j.cub.2020.09.072.

da Silva Georsch, M C, L Schafer, M Tonial, V R de Oliveira, A de Borros Falcao Ferraz, J Fachini, J B da Silva, et al. 2019. “Nutritional composition of *Eragrostis teff* and its association with the observed antimutagenic effects.” Royal Society of Chemistry 9: 3764-3776.

Dace, H JW, A E Adetunji, J P Moore, J M Farrant, and H WM Hilhorst. 2023. “A review of the role of metabolites in vegetative desiccation tolerance of Angiosperms.” Current opinion in Plant Biology. doi:In press.

Darlinger, D L, J P Rinehart, and G D Yocum. 2001. “tress proteins: a role in insect diapause?” Insect Timing: Circadian Rhythmicity to Seasonality. 155-171. doi:https://doi.org/10.1016/B978-044450608-5/50045-3.

de Vetten, N C, and R J Ferl. 1994. “Transcriptional regulation of environmentally inducible genes in plants by an evolutionary conserved family of G-box binding factors.” International Journal of Biochemistry 26 (9): 1055-1068. doi:10.1016/0020-711x(94)90128-7.

Deeba, F, V Pandey, U Pathre, and S Kanojiya. 2009. “Proteome Analysis of Detached Fronds from a Resurrection Plant *Selaginella bryopteris* - Response to dehydration and rehydration.” Journal of Proteomics and Bioinformatics 2 (2): 108-116.

Dempewolf, H, G Baute, J Anderson, B Killian , C Smith, and L Gaurino. 2017. “Past and future use of wild relatives in crop breeding.” Crop Science 57 (3): 1070-1082. doi:https://doi.org/10.2135/cropsci2016.10.0885 .

Diamos, A G, and H S Mason. 2019. “Modifying the Replication of Geminiviral Vectors Reduces Cell Death and Enhances Expression of Biopharmaceutical Proteins in Nicotiana benthamiana Leaves.” Frontiers in Plant Science 9. doi:https://doi.org/10.3389/fpls.2018.01974.

Dickison, W C. 2000. “Defense mechanisms and structural responses of plants.” In Integrative Plant Anatomy, 364-367. Elsevier.

Dietz, KJ. 2011. “Peroxiredoxins in Plants and Cyanobacteria.” Antioxidants and REDOX signalling 15 (4): 1129-1159. doi:https://doi.org/10.1089%2Fars.2010.3657.

Dinakar, C, and D Bartels. 2013. “Desiccation tolerance in resurrection plants: new insights from transcriptome, proteome, and metabolome analysis.” Frontiers in Plant Science 4. doi:https://doi.org/10.3389/fpls.2013.00482.

Disfani, F M, W-L Hsu, M J Mizainty, C J Oldfield, B Xue, A K Dunker, V N Uversky, and L Kurgan. 2012. “MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins.” Bioinformatics 28 (12). doi:https://doi.org/10.1093/bioinformatics/bts209.

Distéfano, A M, G A López, N Setzes, F Marchetti, M Cainzos, M Cascallares, E Zabaleta, and G C Pagnussat. 2021. “Ferroptosis in plants: triggers, proposed mechanisms, and the role of iron in modulating cell death.” Journal of Experimental Botany 72 (6): 2125-2135. doi:https://doi.org/10.1093/jxb/eraa425.

Distéfano, A M, M V Martin, J P Córdoba, A M Bellido, S D’Ippólito, S L Colman, D Soto, et al. 2017. “Heat stress induces ferroptosis-like cell death in plants.” Journal of Cell Biology 216 (2): 463-476. doi:https://doi.org/10.1083%2Fjcb.201605110.

Dixon, S J, K M Lemberg, M R Lamprecht, W S Yang, B Morrison III, and B R Stockwell. 2012. “Ferroptosis: An Iron-Dependent Form of Nonapoptotic Cell Death.” Cell 149 (5): 1060-1072. doi:https://doi.org/10.1016/j.cell.2012.03.042.

Do, T TH, E Martinoia, Y Lee, and J-U Hwang. 2021. “2021 update on ATP-binding cassette (ABC) transporters: how they meet the needs of plants.” Plant Physiology 187 (4): 1876-1892. doi:10.1093/plphys/kiab193.

Dock-Bregeon, A-C, K A Lewis, and M R Conte. 2021. “The La-related proteins: structures and interactions of a versatile superfamily of RNA-binding proteins.” RNA Biology 18 (2): 178-193. doi:https://doi.org/10.1080%2F15476286.2019.1695712.

Dolgikh, V A, E M Pukhovaya, and E V Zemlyanskaya. 2019. “Shaping Ethylene Response: The Role of EIN3/EIL1 Transcription Factors.” Frontiers in Plant Science 10. doi:https://doi.org/10.3389%2Ffpls.2019.01030.

Doroodian, P, and Z Hua. 2021. “The Ubiquitin Switch in Plant Stress Response.” Plants 10 (2). doi:10.3390/plants10020246.

Du Toit, S F, J Betley, and J M Farrant. 2021. “NaDES formation in vegetative desiccation tolerance: Prospects and challenges.” Advances in Botanical Research 97: 225-252. doi:https://doi.org/10.1016/bs.abr.2020.09.007.

Du, Z, H Su, W Wang, L Ye, H Wei, Z Peng, I Anishchenko, D Baker, and J Yang. 2021. “The trRosetta server for fast and accurate protein structure prediction.” Nature Protocols 15: 5634-5651. doi:https://doi.org/10.1038/s41596-021-00628-9.

Duagaard, M, M Rohde, and M Jäättelä. 2007. “The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions.” FEBS Letters 581 (19): 3702-3710. doi:https://doi.org/10.1016/j.febslet.2007.05.039.

Düster, R, I H Kaltheuner, M Schmitz, and M Geyer. 2021. “1,6-Hexanediol, commonly used to dissolve liquid–liquid phase separated condensates, directly impairs kinase and phosphatase activities.” Journal of Biological Chemistry 269. doi:https://doi.org/10.1016/j.jbc.2021.100260.

Dutilleul, C, M Garmier, G Noctor, C Mathieu, P Chétrit, C H Foyer, and R de Paepe. 2003. “Leaf Mitochondria Modulate Whole Cell Redox Homeostasis, Set Antioxidant Capacity, and Determine Stress Resistance through Altered Signalling and Diurnal Regulation.” The Plant Cell 15 (5): 1212-1226. doi:https://doi.org/10.1105/tpc.009464.

Ellison, C T, F Vandenbussche, D Van der Straeten, and S L Harmer. 2011. "XAP5 CIRCADIAN TIMEKEEPER Regulates Ethylene Responses in Aerial Tissues of Arabidopsis." Plant Physiology 155 (2): 988-999. doi:https://doi.org/10.1104/pp.110.164277.

El-Maarouf-Bouteau, H, and C Bailly. 2008. “Oxidative signalling in seed germination and dormancy.” Plant Signalling & Behaviour 3 (3): 175-182. doi:https://doi.org/10.4161/psb.3.3.5539.

Fait, A, H Fromm, D Walter, G Galili, and A R Fernie. 2008. “Highway or byway: the metabolic role of the GABA shunt in plants.” Trends in Plant Science 13 (1): 14-9. doi:10.1016/j.tplants.2007.10.005.

Farmer, E E, and C Davoine. 2007. “Reactive electrophile species.” Current Opinion Plant Biology 10 (4): 380-386. doi:10.1016/j.pbi.2007.04.019.

Farooq, M, A Wahid, N Kobayashi, D Fujita, and S M.A Basra. 2009. “Plant drought stress: effects, mechanisms and management.” Agronomy for Sustainable Development 185-212. doi:10.1051/agro:2008021.

Farrant, J M. 2000. “A comparison of mechanisms of desiccation tolerance among three angiosperm resurrection plant species.” Plant Ecology 151: 29-39. doi:https://doi.org/10.1023/A:1026534305831.

Farrant, J M, and H WM Hilhorst. 2022. “Crops for dry environments.” Current Opinion in Biotechnology 74: 84-91. doi:https://doi.org/10.1016/j.copbio.2021.10.026.

Farrant, J M, and H WM Hilhorst. 2021. “What is dry? Exploring metabolism and molecular mobility at extremely low water contents.” Journal of Experimental Botany 72 (5): 1507-1510. doi:https://doi.org/10.1093/jxb/eraa579.

Farrant, J M, K Cooper, A Hilgart, K O Abdalla, J Bentley, J A Thomson, and M S Rafudeen. 2015. “A molecular physiological review of vegetative desiccation tolerance in the resurrection plant *Xerophyta viscosa* (Baker).” Planta 242 (2): 407-426. doi:https://doi.org/10.1007/s00425-015-2320-6.

Farrant, J M, K Cooper, and J S Nell. 2012. “Desiccation Tolerance.” Chap. 11 in Plant Stress Physiology, edited by S Shabala, 238-265. Wallingford: CAB International. doi:1845939956.

Farrant, J M, W Brandt, and G G Lindsey. 2007. “An overview of mechanisms of desiccation tolerance in selected angiosperm resurrection plants.” Plant Stress 1 (1): 72-84. doi:https://doi.org/10.1002/9780470376881.ch3.

Feng, S, Q Hu, W Huang, CH Ho, R Li, and Z Tang. 2014. “Projected climate regime shift under future global warming from multi-model, multi-scenario CMIP5 simulations.” Global and Planetary Change 112: 41-52.

Fourie, A M, J F Sambrook, and MJ H Gething. 1994. “Common and divergent peptide binding specificities of hsp70 molecular chaperones.” The Journal of Biological Chemistry 269 (48): 30470-30478. doi:https://doi.org/10.1016/S0021-9258(18)43837-9.

Foyer, C H. 2018. “Reactive Oxygen Species, Oxidative Signalling and the Regulation of Photosynthesis.” Environmental and Experimental Botany 154: 134-142. doi:https://doi.org/10.1016/j.envexpbot.2018.05.003.

Frank, W, T Munnik, K Kerkmann, F Salamini, and D Bartels. 2000. “Water Deficit Triggers Phospholipase D Activity in the Resurrection Plant Craterostigma plantagineum.” The Plant Cell 12 (1): 111-123. doi:https://doi.org/10.1105/tpc.12.1.111.

Fraser, C M, M G Thomson, A M Shirley, J Ralph, J A Schoenherr, T Sinlapadech, M C Hall, and C Chapple. 2007. “Related Arabidopsis Serine Carboxypeptidase-Like Sinapoylglucose Acyltransferases Display Distinct But Overlapping Substrate Specificities.” Plant Physiology 144 (4): 1986-1999. doi:https://doi.org/10.1104/pp.107.098970.

Fu, JY, XC Wang, TF Mao, H Cheng, F Chen, and YJ Yang. 2018. “Identification and functional analysis of germin-like protein Gene family in tea plant (*Camellia sinensis*).” Scientia Horticulturae 234: 166-175. doi:https://doi.org/10.1016/j.scienta.2018.02.024.

Fujii, H, and JK Zhu. 2009. “Arabidopsis mutant deficient in 3 abscisic acid-activated protein kinases reveals critical roles in growth, reproduction, and stress.” PNAS 106 (20): 8380-8385. doi:https://doi.org/10.1073/pnas.0903144106.

Fujii, H, P E Verslues, and JK Zhu. 2007. “Identification of Two Protein Kinases Required for Abscisic Acid Regulation of Seed Germination, Root Growth, and Gene Expression in Arabidopsis.” Plant Cell 19 (2): 485-494. doi:https://doi.org/10.1105%2Ftpc.106.048538.

Gabier, H, D L Tabb, J M Farrant, and M S Rafudeen. 2021. “A Label-Free Proteomic and Complementary Metabolomic Analysis of Leaves of the Resurrection Plant *Xerophyta schlechteri* during Dehydration.” Life 11 (2). doi:https://doi.org/10.3390/life11111242.

Gaff, D F. 1971. "Desiccation-Tolerant Flowering Plants in Southern Africa." Science 174 (4013). doi:https://doi.org/10.1126/science.174.4013.1033.

Gaff, D F, and M J Oliver. 2013. “The evolution of desiccation tolerance in angiosperm plants: A rare yet common phenomenon.” Functional Plant Biology 40 (4): 315-328. doi:https://doi.org/10.1071/FP12321.

Gasteiger, E, C Hoogland, C Gattiker, S Duvaud, M R Wilkins, R D Appel, and A Bairoch. 2005. “Protein Identification and Analysis Tools on the Expasy Serve.” In The Proteomics Protocols Handbook, edited by J M Walker, 571-607. Humana Press.

Gebre, E, L Gugsa, U Schulter, and K Kunert. 2013. “Transformation of tef (*Eragrostis tef*) by Agrobacterium through immature embryo regeneration system for inducing semi-

dwarfism." South African Journal of Botany 87: 9-17. doi:https://doi.org/10.1016/j.sajb.2013.03.004.

Gelhaye, E, N Rouhier, J Gérard, Y Jolivet, J Gualberto, N Navrot, PI Ohlsson, et al. 2004. "A specific form of thioredoxin h occurs in plant mitochondria and regulates the alternative oxidase." PNAS 101 (40): 14545-14550. doi:https://doi.org/10.1073%2Fpnas.0405282101.

Gelhaye, E, N Rouhier, N Navrot, and J P Jacquot. 2005. "The plant thioredoxin system." Cellular and Molecular Life Sciences 62: 24-35. doi:10.1007/s00018-004-4296-4.

Gelvin, S B. 2003. "Agrobacterium-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool." Microbiology and Molecular Biology Reviews 67 (1): 16-37. doi:10.1128/MMBR.67.1.16-37.2003.

Ghasempour, H R, D F Gaff, and R D Gianello. 1998. "Contents of sugars in leaves of drying desiccation tolerant flowering plants, particularly grasses." Plant Growth Regulation 24: 185-191. doi:https://doi.org/10.1023/A:1005927629018.

Ginbot, Z G, and J M Farrant. 2011. "Physiological response of selected Eragrostis species to water-deficit stress." African Journal of Biotechnology 10 (51): 10405-10417. doi:10.5897/AJB11.1124.

Ginsawaeng, O, C Heise, R Sangwan, D Karcher, I E Hernández-Sánchez, A Sampathkumar, and E Zuther. 2021. "Subcellular Localization of Seed-Expressed LEA_4 Proteins Reveals Liquid-Liquid Phase Separation for LEA9 and forLEA48 Homo- and LEA42-LEA48 Heterodimers." Biomolecules 11. doi:doi.org/10.3390/biom11121770.

Giuliani, S, M C Sanguineti, R Tuberosa, M Bellotti, S Salvi, and P Landi. 2005. "Root-ABA1 a major constitutive QTL affects maize root architecture and leaf ABA concentration at different water regimes." Journal of Experimental Botany 56 (422): 3061-3070. doi:https://doi.org/10.1093/jxb/eri303.

Goldberg, T, M Hecht, T Hamp, T Karl, G Yachdav, N Ahmed, U Altermann, et al. 2014. "LocTree3 prediction of localization." Nucleic Acids Research 42: 350-355. doi:10.1093/nar/gku396.

Graham, L E, J M Graham, and L W Wilcox. 2014. "Photosynthesis and Respiration." In Plant Biology. Essex: Pearson Education Limited.

Graves, A E, S L Goldman, S W Banks, and A C Graves. 1988. "Scanning Electron Microscope Studies of Agrobacterium tumefaciens Attachment to *Zea mays*, Gladiolu ssp.,

and *Triticum aestivum*." Journal of Bacteriology 170 (5): 2395-2400. doi:10.1128/jb.170.5.2395-2400.1988.

Gribaldo, S, V Lumia, R Creti, E C de Macario, A Sanangelantoni, and P Cammarano. 1999. "Discontinuous Occurrence of the hsp70(dnaK) Gene among Archaea and Sequence Features of HSP70 Suggest a Novel Outlook on Phylogenies Inferred from This Protein." Journal of Bacteriology 181 (2). doi:https://doi.org/10.1128/JB.181.2.434-443.1999.

Gugsa, L, and J Kumlehn. 2011. "Somatic Embryogenesis and Massive Shoot Regeneration from Immature Embryo Explants of Tef." Biotechnology Research International 2011. doi:doi:10.4061/2011/309731.

Guhr, A, M A Horn, and A R Weig. 2017. "Vitamin B2 (riboflavin) increases drought tolerance of Agaricus bisporus." Mycologia 109 (6): 860-873. doi:10.1080/00275514.2017.1414544.

Guiboileau, A, R Sormani, R Meyer, and C Masclaux-Daubresse. 2010. "Senescence and death of plant organs: Nutrient recycling and developmental regulation." Comptes Rendus - Biologies 333 (4): 382-391. doi:https://doi.org/10.1016/j.crvi.2010.01.016.

Hanson, A D, G A Beaudoin, D R McCarty, and J F Gregory III. 2016. "Does Abiotic Stress Cause Functional B Vitamin Deficiency in Plants?" Plant Physiology 172 (4): 2082-2097. doi:https://doi.org/10.1104%2Fpp.16.01371.

He, J, P Xin, X Ma, J Chu, and G Wang. 2020. "Gibberellin Metabolism in Flowering Plants: An Update and Perspectives." Frontiers in Plant Science 11. doi:https://doi.org/10.3389/fpls.2020.00532.

Hedderson, T A, and R H Zander. 2008. "*Vrolijkheidia circumscissa* (Pottiaceae), a new moss genus and species from the Succulent Karoo of South Africa." Journal of Brylogy 30: 143-146. doi:10.1179/174328208X282157.

Heldt, HW, and B Piechulla. 2010. "Biotechnology alters plants to meet requirements." In Plant Biochemistry, 565. Elsevier.

Hellens, R, P Mullineaux, and H Klee. 2000. "Technical Focus:A guide to Agrobacterium binary Ti vectors." Trends in Plant Science 5 (10): 446-451. doi:https://doi.org/10.1016/S1360-1385(00)01740-4.

Hernández, L A, S Bressendorff, M H Hansen, C Poulson, S Erdmann, and P Brodersen. 2018. "An m6A-YTH Module Controls Developmental Timing and Morphogenesis in Arabidopsis." Plant Cell 30 (5): 952-967. doi:10.1105/tpc.17.00833.

Hobisch, M, D Holtmann, P G de Santos, M Alcalde, F Hollmann, and S Kara . 2021. “Recent developments in the use of peroxygenases – Exploring their high potential in selective oxyfunctionalisations.” Biotechnology advances 51. doi:https://doi.org/10.1016%2Fj.biotechadv.2020.107615.

Hoekstra, F, E Golovina, and J Buitink. 2001. “Mechanisms of plant desiccation tolerance.” Trends in Plant Science 6 (9): 431-438. doi:https://doi.org/10.1016/S1360-1385(01)02052-0.

Huang, K, Y Liu, Y Shi, J Tian, T Shi, H Peng, X Tian, W Zhang, and K Wang. 2022. "Overexpression of TaMBF1c improves thermo-tolerance of perennial ryegrass." Scientia Horticulturae 295 (15). doi:https://doi.org/10.1016/j.scienta.2021.110812.

Huang, A H.C. 2018. “Plant Lipid Droplets and Their Associated Proteins: Potential for Rapid Advances.” Plant Physiology 176 (3): 1894-1918. doi:https://doi.org/10.1104/pp.17.01677.

Hwang, HH, M Yu, and EM Lai. 2017. “Agrobacterium-mediated plant transformation: biology and applications.” American Society of Plant Biologist. doi:10.1199/tab.0186.

Ikeuchii, M, K Sugimoto, and A Iwase. 2013. “Plant Callus: Mechanisms of Induction and Repression.” The Plant Cell 25 (9): 3159-3173. doi:https://doi.org/10.1105/tpc.113.116053.

Illing, N, K J Denby, H Collett, A Shen, and J M Farrant. 2005. “The Signature of Seeds in Resurrection Plants: A Molecular and Physiological Comparison of Desiccation Tolerance in Seeds and Vegetative Tissues.” Integrative and Comparative Biology 45 (5): 771-787. doi:https://doi.org/10.1093/icb/45.5.771.

Imamoglu, R, D Balchin, M Hayer-Hartl, and F U Hartl. 2020. “Bacterial Hsp70 resolves misfolded states and accelerates productive folding of a multi-domain protein.” Nature Communications 365 (11). doi:https://doi.org/10.1038/s41467-019-14245-4.

Ingle, R A, U G Schmidt, J M Farrant, J A Thomson, and S G Mundree. 2007. “Proteomic analysis of leaf proteins during dehydration of the resurrection plant *Xerophyta viscosa*.” Plant, Cell, and Environment 30 (4): 435-446. doi: https://doi.org/10.1111/j.1365-3040.2006.01631.x.

Iqbal, N, N A Khan, A Ferrante, A Trivellini, A Francini, and M I.R Khan. 2017. “Ethylene Role in Plant Growth, Development and Senescence: Interaction with Other Phytohormones.” Frontiers in Plant Science 8 (475). doi:https://doi.org/10.3389%2Ffpls.2017.00475.

Izquierdo, Y, S Kulasekaran, P Benito, B López, R Marcos, and T Cascón. 2018. “Arabidopsis nonresponding to oxylipins locus NOXY7 encodes a yeast GCN1 homolog that

mediates noncanonical translation regulation and stress adaptation.” Plant, Cell & Environment 41: 1438-1452. doi:doi.org/10.1111/pce.13182.

Jiang, G, Z Wang, H Shang, W Yang, Z Hu, J Phillips, and X Deng. 2006. “Proteome analysis of leaves from the resurrection plant *Boea hygrometrica* in response to dehydration and rehydration.” Planta 225. doi:https://doi.org/10.1007/s00425-006-0449-z.

Kamies, R, J M Farrant, Z Tadele, G Cannarozzi, and M S Rafudeen. 2017. “A Proteomic Approach to Investigate the Drought Response in the Orphan Crop *Eragrostis tef*.” Proteomes 5 (32). doi:doi:10.3390/proteomes5040032 .

Kato, Y, and W Sakamoto. 2018. “FtsH Protease in the Thylakoid Membrane: Physiological Functions and the Regulation of Protease Activity.” Frontiers in Plant Science 9. doi:doi.org/10.3389/fpls.2018.00855.

Kebebew, A, M D Gaj, and M Maluszynski. 1998. “Somatic embryogenesis and plant regeneration in callus culture of *tef, Eragrostis* tef (Zucc.) Trotter.” Plant Cell Reports 18: 154-158. doi:https://doi.org/10.1007/s002990050549.

Kedersha, N, P Ivanov, and P Anderson. 2013. “Stress granules and cell signalling: more than just a passing phase?” Trends Biochem Sci 38 (10). doi:10.1016/j.tibs.2013.07.004.

Khanna, R, B Kronmiller, D R Maszle, G Coupland, M Holm, T Mizuno, and S-H Wu. 2009. “The Arabidopsis B-Box Zinc Finger Family.” Plant Cell 21 (11): 3416-3420. doi:10.1105/tpc.109.069088.

Kim, B, G Kim, S Fujioka, S Takatsuto, and S Choe. 2012. “Overexpression of 3β-Hydroxysteroid Dehydrogenases/C-4 Decarboxylases Causes Growth Defects Possibly Due to Abnormal Auxin Transport in Arabidopsis.” Molecules and Cels 34: 77-84. doi:10.1007/s10059-012-0102-6.

Kim, W, Y Lee, J Park, N Lee, and G Choi. 2013. “HONSU, a Protein Phosphatase 2C, Regulates Seed Dormancy by Inhibiting ABA Signalling in Arabidopsis.” Plant Cell Physiology 54 (4): 555-572. doi:doi:10.1093/pcp/pct017.

Kobayashi, Y, N Fukuzawa, A Hyobo, H Kim, S Mashiyama, T Ogihara, H Yoshioka, et al. 2020. “Role of salicylic acid glucosyltransferase in balancing growth and defence for optimum plant fitness.” Molecular Plant Pathology 21 (3): 429-442. doi:https://doi.org/10.1111%2Fmpp.12906.

Kumimoto, R W, C T Ellison, T Y Toruño, A Bak, H Zhang, C L Casteel, G Coaker, and S L Harmer. 2021. "XAP5 CIRCADIAN TIMEKEEPER Affects Both DNA Damage Responses

and Immune Signaling in Arabidopsis." Frontiers in Plant Science 12. doi:https://doi.org/10.3389/fpls.2021.707923.

Komori, T, T Imayama, N Kato, Y Ishida, J Ueki, and T Komari. 2007. “Current Status of Binary Vectors and Superbinary Vectors.” Plant Physiology 145: 1155-1160. doi: 10.1104/pp.107.105734.

Krenek, S, M Schlegel, and T U Berendonk. 2013. “Convergent evolution of heat-inducibility during subfunctionalization of the Hsp70 gene family.” BMC Evolutionary Biology 13 (49). doi:https://doi.org/10.1186/1471-2148-13-49.

Kumar, G P, S Subiramani, S Govindarajan, V Sadasivam, V Manickam, K Mogilicherla, S K Thiruppathi, and J Narayanasamy. 2015. “Evaluation of different carbon sources for high frequency callus culture with reduced phenolic secretion in cotton (*Gossypium hirsutum* L.) cv. SVPR-2.” Biotechnology Reports 7: 72-80. doi:doi.org/10.1016/j.btre.2015.05.005.

Lacroix, B, and V Citovsky. 2013. “The roles of bacterial and host plant factors in Agrobacterium-mediated genetic transformation.” The International Journal of Developmental Biology 57: 467-481. doi:10.1387/ijdb.130199b.

Latrasse, D, M Benhamed, C Bergounioux, C Raynaud, and M Delarue. 2016. "Plant programmed cell death from a chromatin point of view." Journal of Experimental Botany 67 (20): 5887-5990. doi:https://doi.org/10.1093/jxb/erw329.

Lee, LY, and S B Gelvin. 2008. “T-DNA Binary Vectors and Systems.” American Society of Plant Biologist. doi: https://doi.org/10.1104/pp.107.113001.

Lee, S, A Marmagne, J Park, C Fabien, Y Yim, S-J Kim, T-H Kim, P O Lim, C Masclaux-Daubresse, and H G Nam. 2020. “Concurrent activation of OsAMT1;2 and OsGOGAT1 in rice leads to enhanced nitrogen use efficiency under nitrogen limitation.” The Plant Journal 103 (1): 7-20. doi:https://doi.org/10.1111/tpj.14794.

Lee, S, and C Masclaux-Daubresse. 2021. “Current Understanding of Leaf Senescence in Rice.” International Journal of Molecular Science 22 (9). doi:https://doi.org/10.3390/ijms22094515.

Lei, P, Z Liu, J Li, G Jin, L Xu, X Ji, X Zhao, L Tao, and F Meng. 2022. “Integration of the Physiology, Transcriptome and Proteome Reveals the Molecular Mechanism of Drought Tolerance in Cupressus gigantea.” Forests 13 (3): 401. doi:https://doi.org/10.3390/f13030401.

Li, C, C Yan, Q Sun, J Wang, C Yuan, Y Mou, and X Zhao. 2022a. “Proteomic profiling of *Arachis hypogaea* in response to drought stress and overexpression of AhLEA2 improves drought tolerance.” Plant Biology 24 (1): 75-84. doi:https://doi.org/10.1111/plb.13351.

Li, Q, Q Zheng, W Shen, D Cram, D B Fowler, Y Wei, and J Zou. 2015. “Understanding the Biochemical Basis of Temperature-Induced Lipid Pathway Adjustments in Plants.” The Plant Cell 27 (1): 86-103. doi: https://doi.org/10.1105/tpc.114.134338.

Li, R, X Su, R Zhou, Y Zhang, and T Wang. 2022b. “Molecular mechanism of mulberry response to drought stress revealed by complementary transcriptomic and iTRAQ analyses.” BMC Plant Biology 22 (36). doi:https://doi.org/10.1186/s12870-021-03410-x.

Li, S. 2015. “The Arabidopsis thaliana TCP transcription factors: A broadening horizon beyond development.” Plant Signal Behaviour 10 (7). doi:10.1080/15592324.2015.1044192.

Li, S, D Li, P Zhang, R Wang, L Sun, J Wan, and J Xu. 2018. “ABCF3 regulates the expression of aquaporin genes and endoplasmic reticulum stress-related genes in Arabidopsis.” Theoretical and Experimental Plant Physiology 30: 215-222. doi:doi.org/10.1007/s40626-018-0116-3.

Li, Y, J G Gu, C Wang, J Hu, S Zhang, C Liu, S Zhang, Y Fang, and D Li. 2022c. “Hsp70 exhibits a liquid-liquid phase separation ability and chaperones condensed FUS against amyloid aggregation.” Science 25 (6). doi:https://doi.org/10.1016/.

Li, Yipu, and M Xu. 2017. “CCT family genes in cereal crops: A current overview.” The Crop Journal 5 (6): 449-458. doi:https://doi.org/10.1016/j.cj.2017.07.001.

Liao, WB, ML Zhang, GB Haung, and JH Yu. 2012. “Hydrogen peroxide in the vase solution increases vase life and keeping quality of cut Oriental × Trumpet hybrid lily ‘Manissa’.” Scientia Horticulturae 139: 32-38. doi:https://doi.org/10.1016/j.scienta.2012.02.040.

Lim, P O, H J Kim, and H G Nam. 2007. “Leaf Senescence.” Annual Review of Plant Biology 58 (1): 115-136. doi:https://doi.org/10.1146/annurev.arplant.57.032905.105316.

Lin, J, T J Nazarenus, J L Frey, X Liang, M A Wilson, and J M Stone. 2011. “A Plant DJ-1 Homolog Is Essential for *Arabidopsis thaliana* Chloroplast Development.” PLoS ONE 6 (8). doi:https://doi.org/10.1371/journal.pone.0023731.

Liu, D, Z M Wu, and L Hou. 2015. “Loss-of-function mutation in SCY1 triggers chloroplast-to-nucleus retrograde signalling in *Arabidopsis thaliana*.” Biologia Plantarum 59 (3): 469-476. doi:10.1007/s10535-015-0514-1.

Liu, S, Y Cheng, X Zhang, Q Guan, S Nishiuchi, K Hase, and T Takano. 2007. “Expression of an NADP-malic enzyme gene in rice (*Oryza sativa*. L) is induced by environmental stresses; over-expression of the gene in Arabidopsis confers salt and osmotic stress tolerance.” Plant Molecular Biology 64: 49-58. doi:https://doi.org/10.1007/s11103-007-9133-3.

Lorenzo, O. 2019. “bZIP edgetic mutations: at the frontier of plant metabolism, development and stress trade-off.” Journal of Experimental Botany 70 (20): 5517-5520. doi: https://doi.org/10.1093/jxb/erz298.

Lu, J, X Zhang, Y Wu, Y Sheng, W Li, and W Wang. 2021. “Energy landscape remodelling mechanism of Hsp70-chaperone-accelerated protein folding.” Biophsyical Journal 120 (10). doi:https://doi.org/10.1016/j.bpj.2021.03.013.

Lu, Y, T Wu, O Gutman, H Lu, Q Zhou, Y I Henis, and K Luo. 2020. “Phase separation of TAZ compartmentalizes the transcription machinery to promote gene expression.” Nature Cell Biology 22: 453-464. doi:https://doi.org/10.1038/s41556-020-0485-0.

Lyall, R, and T Gechev. 2020. “Multi-omics insights into the evolution of angiosperm resurrection plants.” Annual Plant Reviews 3: 77-110. doi: 10.1002/9781119312994.apr0730.

Ma, C, H Wang, A J Macnish, A C Estrada-Melo, J Lin, Y Chang, and C Z Jiang. 2015. “Transcriptomic analysis reveals numerous diverse protein kinases and transcription factors involved in desiccation tolerance in the resurrection plant *Myrothamnus flabellifolia*.” Horticulture Research. doi:https://doi.org/10.1038/hortres.2015.34.

Macario, A J.L, M Lange, B K Ahring, and E C De Macario. 1999. “Stress Genes and Proteins in the Archaea.” Microbiology and Molecular Biology Reviews 63 (4). doi:https://doi.org/10.1128/MMBR.63.4.923-967.1999.

Madden, C F. 2019. “*Eragrostis nindensis*: unravelling senescence in an African tolerant grass.” Department of Molecular and Cell Biology.

Maraia, R J, S Mattijssen, I Cruz-Gallardo, and M R Conte. 2017. “The LARPs, La and related RNA-binding proteins: Structures, functions and evolving perspectives.” Wiley Interdisciplinary Reviews: RNA 8 (6). doi:10.1002/wrna.1430.

Maris, C, C Dominguez, and F H-T Allain. 2005. “The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression.” The FEBS Journal 272 (9): 2118-2131. doi:10.1111/j.1742-4658.2005.04653.x.

Marks, R A, J M Farrant, D N McLetchie, and R VanBuren. 2021. “Unexplored dimensions of variability in vegetative desiccation tolerance.” American Journal of Botany 108: 1-13. doi:https://doi.org/10.1002/ajb2.1588.

Martinelli, T, A Whittaker, A Bochicchio, C Vazzana, A Suzuki, and C Maclaux-Daubresse. 2007. “Amino acid pattern and glutamate metabolism during dehydration stress in the ‘resurrection’ plant *Sporobolus stapfianus*: a comparison between desiccation-sensitive and desiccation-tolerant leaves.” Journal of Experimental Botany 58 (11): 3037-3046. doi:https://doi.org/10.1093/jxb/erm161.

Martinelli, T, A Whittaker, A Masclaux-Daubresse, J M Farrant, F Brilli, F Lorento, and C Vazzana. 2007. “Evidence for the presence of photorespiration in desiccation-sensitive leaves of the C4 “resurrection” plant *Sporobolus stapfianus* during dehydration stress.” Journal of Experimental Botany 58 (14): 3929-3939. doi:https://doi.org/10.1093/jxb/erm247.

Maruri-López, I, N E Figueroa, I E Hernández-Sánchez, and M Chodasiewicz. 2021. “Plant Stress Granules: Trends and Beyond.” Frontiers in Plant Science. doi: https://doi.org/10.3389/fpls.2021.722643.

Masand, S, and S K Yadav. 2016. “Overexpression of MuHSP70 gene from *Macrotyloma uniflorum* confers multiple abiotic stress tolerance in transgenic Arabidopsis thaliana.” Molecular Biology Reports 43: 53-64. doi:https://doi.org/10.1007/s11033-015-3938-y.

Maynard, D, V Kumar, J Sproß, and K-J Dietz. 2020. “12-Oxophytodienoic Acid Reductase 3 (OPR3) Functions as NADPH-Dependent α,β-Ketoalkene Reductase in Detoxification and Monodehydroascorbate Reductase in Redox Homeostasis.” Plant and Cell Physiology 61 (3): 584-595. doi:doi.org/10.1093/pcp/pcz226.

Meesapyodsuk, D, and X Qiu. 2011. “A Peroxygenase Pathway Involved in the Biosynthesis of Epoxy Fatty Acids in Oat[W][OA].” Plant Physiology 157 (1): 454-463. doi:https://doi.org/10.1104%2Fpp.111.178822.

Mekbib, F, S H Mantell, and V Buchanan-Wollaston. 1996. “Callus Induction and In Vitro Regeneration of Tef [*Eragrostis tef* (Zucc.) Trotter] from Leaf.” Journal of Plant Physiology 151: 368-372. doi:https://doi.org/10.1016/S0176-1617(97)80267-2.

Meng, Q, Z Peng, and J Yang. 2018. “CoABind: a novel algorithm for Coenzyme A (CoA)- and CoA derivatives-binding residues prediction.” Bioinformatics 1 (34). doi:10.1093/bioinformatics/bty162 .

Mittler, R. 2002. “Oxidative stress, antioxidants and stress tolerance.” TRENDS in Plant Science 7 (9). doi:https://doi.org/10.1016/S1360-1385(02)02312-9.

Mizianty, M J, Z Peng, and L Kurgan. 2013. “MFDp2: Accurate predictor of disorder in proteins by fusion of disorder probabilities, content and profiles.” Intrinsically Disordered Proteins 1 (1). doi: 10.4161/idp.24428.

Mock, H P, T Vogt, and D Strack. 1992. “Sinapoylglucose: Malate Sinapoyltransferase acitivity in *Arabidopsis thaliana* and *Brassica rapa*.” Zeitshrift fur Naturforshung C 47: 680-682. doi:doi.org/10.1515/znc-1992-9-1007.

Mongrand, S, JJ Bessoule, F Cabantous, and C Cassagne. 1998. “The C16:3\C18:3 fatty acid balance in photosynthetic tissues from 468 plant species.” Phytochemistry 49 (4): 1049-1064. doi:https://doi.org/10.1016/S0031-9422(98)00243-X.

Moore, J O, E Nguema-Ona, E Vicré-Gibouin, M Sørensen, I Willats, W GT Driouich, and J M Farrant. 2013. “Arabinose-rich polymers as an evolutionary strategy to plasticize resurrection plant cell walls against desiccation.” Planta 237 (3): 739-754. doi:https://doi.org/10.1007/s00425-012-1785-9.

Ncanana, S, W Brandt, G Lindsey, and J M Farrant. 2005. “Development of plant regeneration and transformation protocols for the desiccation-sensitive weeping lovegrass Eragrostis curvula.” Plant Cell Reports 24 (6): 335-340. doi:10.1007/s00299-005-0940-1.

Nester, E W. 2015. “Agrobacterium: Nature's genetic engineer.” Frontiers in Plant Science. doi:https://doi.org/10.3389/fpls.2014.00730.

Nishchal, N, H Mir, R Rani, and A K Pal. 2018. “Effect of Antioxidants in Controlling Phenol Exudation in Micropropagation of Litchi cv. Purbi.” Current Journal of Applied Science and Technology 31 (4): 1-7. doi:10.9734/CJAST/2018/45992 .

Niu, L, and W Liao. 2016. “Hydrogen Peroxide Signalling in Plant Development and Abiotic Responses: Crosstalk with Nitric Oxide and Calcium.” Frontiers in Plant Science 7. doi:https://doi.org/10.3389/fpls.2016.00230.

NOAA. 2023. NOAA National Centers for Environmental information, Climate at a Glance: Global Time Series. Accessed 2023. https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/global/time-series/globe/land_ocean/ytd/12/1880-2023?trend=true&trend_base=10&begtrendyear=1850&endtrendyear=2023.

Ohlrogge, J, and J Browse. 1995. “Lipid Biosynthesis.” The Plant Cell 7 (7): 957-970. doi:10.1105/tpc.7.7.957.

Oliver, M J, J M Farrant, H WM Hilhorst, S Mundree, B Williams, and J D Bewley. 2020. “Desiccation Tolerance: Avoiding Cellular Damage During Drying and Rehydration.” Annual

Review of Plant Physiology 29 (71): 435-460. doi:https://doi.org/10.1146/annurev-arplant-071219-105542.

O'Mahony, P J, and M J Oliver. 1999. “The involvement of ubiquitin in vegetative desiccation tolerance.” Plant Molecular Biology 41: 657-667. doi:doi.org/10.1023/A:1006330623364.

Pardo, J, C M Wai, H Chay, C F Madden, J M Farrant, and R VanBuren. 2020. “Intertwined signatures of desiccation and drought tolerance in grasses.” Proceedings of the National Academy of Sciences 117 (18): 10079-10088. doi:www.pnas.org/cgi/doi/10.1073/pnas.2001928117.

Park, C J, and Y S Seo. 2015. “Heat Shock Proteins: A Review of the Molecular Chaperones for Plant Immunity.” Journal of Plant Pathology 31 (4): 323-333. doi:10.5423/PPJ.RW.08.2015.0150.

Patil, D P, B F Pickering, and S R Jaffrey. 2018. “Reading m6A in the Transcriptome: m6A-Binding Proteins.” Trends in Cell Biology 28 (2): 113-127. doi:doi.org/10.1016/j.tcb.2017.10.001.

Patnaik, D, and P Khurana. 2001. “Germins and germin like proteins: an overview.” Indian Journal of Experimental Biology 39 (3): 191-200.

Patridge, M, and D J Murphy. 2009. “Roles of a membrane-bound caleosin and putative peroxygenase in biotic and abiotic stress responses in Arabidopsis.” Plant Physiology and Biochemistry 47 (9): 796-806. doi:https://doi.org/10.1016/j.plaphy.2009.04.005.

Paul, P, S Simm, A Blaumeiser, K-D Scharf, S Fragkostefanakis, O Mirus, and E Schleiff. 2013. “The protein translocation systems in plants – composition and variability on the example of Solanum lycopersicum.” BMC Genomics 14 (189). doi:https://doi.org/10.1186%2F1471-2164-14-189.

Pei, Y, X Li, Y Zhu, Z Ge, Y Sun, N Liu, Y Jia, F Li, and Y Hou. 2019. “GhABP19, a Novel Germin-Like Protein From *Gossypium hirsutum*, Plays an Important Role in the Regulation of Resistance to Verticillium and Fusarium Wilt Pathogens.” Frontiers in Plant Science 10. doi:doi.org/10.3389/fpls.2019.00583.

Peng, Z, G Serino, and X-W Deng. 2001. “A role of Arabidopsis COP9 signalosome in multifaceted developmental processes revealed by the characterization of its subunit 3.” Development 128 (21): 4277-4288. doi:doi.org/10.1242/dev.128.21.4277.

Peter, J C, and E G Jay. 2014. “The Agrobacterium Ti Plasmids.” Microbiology Spectrum 2 (6). doi:10.1128/microbiolspec.PLAS-0010-2013.

Peyet, H, J K.M Brown, and G P Lomonossoff. 2019. “Improving plant transient expression through the rational design of synthetic 5′and 3′ untranslated regions.” Plant Methods 15 (108). doi:10.1186/s13007-019-0494-9.

Peyret , H, and G P Lomonossoff. 2013. “The pEAQ vector series: the easy and quick way to produce recombinant proteins in plants.” Plant Molecular Biology 83: 51-58. doi:https://doi.org/10.1007/s11103-013-0036-1.

Pospíšilová, J. 2003. “Interaction of cytokinins and abscisic acid during regulation of stomatal opening in bean leaves.” Photosynthetica 41: 49-56. doi:https://doi.org/10.1023/A:1025852210937.

Poxleitner, M, S W Rogers, A L Samuels, J Browse, and J C Rogers. 2006. “A role for caleosin in degradation of oil-body storage lipid during seed germination.” The Plant Journal 47 (6): 917-933. doi:https://doi.org/10.1111/j.1365-313X.2006.02845.x.

Purvis, A C. 1997. “Role of the alternative oxidase in limiting superoxide production by plant mitochondria.” Physiologia Plantarum 100 (1): 165-170. doi:https://doi.org/10.1111/j.1399-3054.1997.tb03468.x.

Rabbani, N, M Al-Motawa, and P J Thornalley. 2020. "Protein Glycation in Plants—An Under-Researched Field with Much Still to Discover." International Journal of Molecular Sciences 21 (11): 3942. doi:https://doi.org/10.3390%2Fijms21113942.

Radermacher, A L, S F du Toit, and J M Farrant. 2019. “Desiccation-Driven Senescence in the Resurrection Plant Xerophyta schlechteri (Baker) N.L. Menezes: Comparison of Anatomical, Ultrastructural, and Metabolic Responses Between Senescent and Non-Senescent Tissues.” Frontiers in Plant Science 10. doi:https://doi.org/10.3389/fpls.2019.01396.

Regnard, G L, R P Halley-Stott, F Tanzer, I I Hitzeroth, and E P Rybicki. 2010. “High level protein expression in plants through the use of a novel autonomously replicating geminivirus shuttle vector.” Plant Biotechnology Journal 8: 38-46. doi:10.1111/j.1467-7652.2009.00462.x.

Reumann, S, and B Bartel. 2016. “Plant peroxisomes: Recent discoveries in functional complexity, organelle homeostasis, and morphological dynamics.” Current Opinion in Plant Biology 34: 17-26. doi:https://doi.org/10.1016%2Fj.pbi.2016.07.008.

Ritossa, F. 1996. “Discovery of the heat shock response.” Cell Stress Chaperones 1 (2): 97-98. doi:10.1379/1466-1268(1996)001<0097:dothsr>2.3.co;2.

Rivas-San, M V, and J Plasencia. 2011. “Salicylic acid beyond defence: its role in plant growth and development.” Journal of Experimental Botany 62 (10): 3321-3338. doi:https://doi.org/10.1093/jxb/err031.

Rivero, R M, M Kojima, A Gepstein, and E Blumwald. 2007. “Delayed leaf senescence induces extreme drought tolerance in a flowering plant.” PNAS 104: 19631-19636. doi:https://doi.org/10.1073/pnas.0709453104.

Rizhsky, L, H Liang, J Shuman, V Shulaev, S Davletova, and R Mittler. 2004. “When Defense Pathways Collide. The Response of Arabidopsis to a Combination of Drought and Heat Stress.” Plant Physiology 134 (4): 1683-1696. doi:https://doi.org/10.1104/pp.103.033431.

Roberts, E H. 1973. “Predicting the storage life of seeds.” Seed Science and Technology 1 (3): 499-514.

Robson, C A, and G C Vanlerberghe. 2002. “Transgenic Plant Cells Lacking Mitochondrial Alternative Oxidase Have Increased Susceptibility to Mitochondria-Dependent and -Independent Pathways of Programmed Cell Death.” Plant Physiology 129 (4): 1908-1920. doi:https://doi.org/10.1104%2Fpp.004853.

Roser, M, H Ritchie, and E Ortiz-Ospina. 2020. World Population Growth. https://ourworldindata.org/world-population-growth.

Roundtree, I A, and C He. 2016. “Nuclear m6A Reader YTHDC1 Regulates mRNA Splicing.” Trends in Genetics 32 (6): 320-321. doi:doi.org/10.1016/j.tig.2016.03.006.

Ruocco, V, and R Strasser. 2022. “Transient Expression of Glycosylated SARS-CoV-2 Antigens in *Nicotiana benthamiana*.” Plants 11 (8). doi:https://doi.org/10.3390/plants11081093.

Saeedipour, S, and F Moradi. 2012. “Stress-induced changes in the free amino acid composition of two wheat cultivars with difference in drought resistance.” African Journal of Biotechnology 11 (40): 9559-9565. doi:10.5897/AJB11.3870.

Sahu, S S, C D Loaiza, and R Kaundal. 2019. “Plant-mSubP: a computational framework for the prediction of single- and multi-target protein subcellular localization using integrated machine-learning approaches.” AoB Plants 12 (3). doi: 10.1093/aobpla/plz068.

Salvador, V H, R B Lima, W D dos Santos, A R Soares, A F Böhm, R Marchiosi, M Ferrarese, and O Ferrarese-Filho. 2013. “Cinnamic acid increases lignin production and inhibits soybean root growth.” PLoS One 8 (7). doi:10.1371/journal.pone.0069105 .

Sarkar, N S, P Kundnani, and A Grover. 2012. “Functional analysis of Hsp70 superfamily proteins of rice (*Oryza sativa*).” Cell stress and Chaperones 18: 427-437. doi:https://doi.org/10.1007/s12192-012-0395-6.

Sasnauskas, G, E Manakova, K Lapėnas, K Kauneckaitė, and V Siksnys. 2018. “DNA recognition by Arabidopsis transcription factors ABI3 and NGA1.” The FEBS Journal 285 (21): 4041-4059. doi:https://doi.org/10.1111/febs.14649.

Savojardo, C, C Martelli, P L Fariselli, P Profiti, and G Casadio. 2018. “BUSCA: an integrative web server to predict subcellular localization of proteins.” Nucleic Acids Research 46 (1): 459-466. doi:https://doi.org/10.1093/nar/gky320.

Sevrioukova, I F. 2011. "Apoptosis-Inducing Factor: Structure, Function, and Redox Regulation." Antioxidants and REDOX Signaling 14 (12): 2545-2579. doi:https://doi.org/10.1089%2Fars.2010.3445.

Schmidt, W, S Thomine, and T J Buckhout. 2020. “Editorial: Iron Nutrition and Interactions in Plants.” Plant Science 10. doi:https://doi.org/10.3389/fpls.2019.01670.

Seo, J S, J Joo, M J Kim, Y K Kim, B H Nahm, S I Song, J J Cheong, J S Lee, J K Kim, and Y D Choi. 2011. “OsbHLH148, a basic helix-loop-helix protein, interacts with OsJAZ proteins in a jasmonate signalling pathway leading to drought tolerance in rice.” Plant Journal 65: 907-921. doi: https://doi.org/10.1111/j.1365-313X.2010.04477.x.

Shao, S. 2022. “Protein biosynthesis at the ER: finding the right accessories.” Molecular Biology of the Cell 34 (1). doi:doi.org/10.1091/mbc.E21-09-0451.

Sharma, N, K P Sharma, R K Guar, and V K Gupta . 2011. “Role of chitinase in plant defense.” Asian Journal of Biochemistry 6 (1): 29-37. doi:10.3923/ajb.2011.29.37.

Sharma, P, A B Jha, R S Dubey, and M Pessarakli. 2012. “Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions.” Journal of Botany 2012. doi:https://doi.org/10.1155/2012/217037.

Sherwin, H W, and J M Farrant. 1998. “Protection mechanisms against excess light in the resurrection plants *Craterostigma wilmsii* and *Xerophyta viscosa*.” Plant Growth Regulation 24 (3): 203-210. doi:https://doi.org/10.1023/A:1005801610891.

Sherwood, S, and Q Fu. 2014. “A drier future?” Science 343 (6172): 737-739. doi:10.1126/science.1247620 .

Shi, H, T Ye, F Chen, Z Cheng, Y Wang, P Yang, Y Zhang, and Z Chan. 2013. “Manipulation of arginase expression modulates abiotic stress tolerance in Arabidopsis: effect on arginine

metabolism and ROS accumulation." Journal of Experimental Botany 64 (5): 1367-1379. doi:https://doi.org/10.1093/jxb/ers400.

Sibéril, Y, P Doireau, and P Gantet. 2001. "Plant bZIP G-box binding factors. Modular structure and activation mechanisms." European Journal of Biochemistry 268 (22): 5655-5666. doi:10.1046/j.0014-2956.2001.02552.x.

Simões, I, and C Faro. 2004. "Structure and function of plant aspartic proteinases." European Journal of Biochemistry 271 (11): 2067-2075. doi:doi.org/10.1111/j.1432-1033.2004.04136.x.

Skalitzky, C A, J R Martin, J H Harwood, J J Bierne, B J Adamczyk, G R Heck, K Cline, and D E Fernandez. 2010. "Plastids Contain a Second Sec Translocase System with Essential Functions." Plant Phsyiology 155: 354-369. doi:10.1104/pp.110.166546.

Song, A, X Zhu, F Chen, H Gao, J Jiang, and S Chen. 2014. "A Chrysanthemum Heat Shock Protein Confers Tolerance to." International Journal of Molecular Sciences 15: 5063-5078. doi:10.3390/ijms15035063.

Song, J T. 2006. "Induction of a salicylic acid glucosyltransferase, AtSGT1, is an early disease response in *Arabidopsis thaliana*." Molecules and Cells 22 (2): 233-238. doi:PMID: 17085977.

Sood, P, A Bhattacharya, and A Sood. 2011. "Problems and possibilities of monocot transformation." Biologia Plantarum 55 (1): 1-15. doi:10.1007/s10535-011-0001-2.

Souer, E, A van Houwelingen, D Kloos, J Mol, and R Koes. 1996. "The No Apical Meristem Gene of Petunia Is Required for Pattern Formation in Embryos and Flowers and Is Expressed at Meristem and Primordia Boundaries." Cell 85 (2): 159-170. doi:https://doi.org/10.1016/S0092-8674(00)81093-4.

Sperschneider, J, A-M Catanzariti, K DeBoer, B Petre, D M Gardiner, K B Singh, P N Dodds, and J M Taylor. 2017. "LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell." Scientific Reports 7. doi:https://doi.org/10.1038/srep44598.

Su, H, M Liu, S Sun, Z Peng, and J Yang. 2019. "Improving the prediction of protein-nucleic acids binding residues via multiple sequence profiles and the consensus of complementary algorithms." Bioinformatics 35 (6). doi:10.1093/bioinformatics/bty756.

Su, H, W Wang, Z Du, Z Peng, S-H Gao, M-M Cheng, and J Yang. 2021. “Improved protein structure prediction using a new multi-scale network and homologous templates.” Advanced Science 8. doi:10.1002/advs.202102592.

Sultana, N, S Islam, A Juhasz, and W Ma. 2021. “Wheat leaf senescence and its regulatory gene network.” The Crop Journal 9 (4): 703-717. doi:https://doi.org/10.1016/j.cj.2021.01.004.

Suprasanna, P, G C Nikalje, and A N Rai. 2016. “Osmolyte Accumulation and Implications in Plant Abiotic Stress Tolerance.” In Osmolytes and Plants Acclimation to Changing Environment: Emerging Omics Technologies, edited by N Iqbal, R Nazar and N Khan. Springer, New Delhi. doi:https://doi.org/10.1007/978-81-322-2616-1_1.

Suzuki, N, S Bajad, J Shuman, V Shulaev, and R Mittler. 2008. "The Transcriptional Co-activator MBF1c Is a Key Regulator of Thermotolerance in Arabidopsis thaliana." Journal of Biological Chemistry 283 (14): 9269-9275. doi:https://doi.org/10.1074/jbc.M709187200.

Takáč, T, D Novák, and J Šamaj. 2019. “Recent Advances in the Cellular and Developmental Biology of Phospholipases in Plants.” Frontiers in Plant Science 10. doi:https://doi.org/10.3389/fpls.2019.00362.

Takagi, D, H Inoue, M Odawara, G Shimakawa, and C Miyake. 2014. "The Calvin Cycle Inevitably Produces Sugar-Derived Reactive Carbonyl Methylglyoxal During Photosynthesis: A Potential Cause of Plant Diabetes." Plant and Cell Physiology 55 (2): 333-340. doi:https://doi.org/10.1093/pcp/pcu007.

Takahashi, S, T Katagiri, K Wamaguchi-Shinozaki, and K Shinozaki. 2000. “An Arabidopsis Gene Encoding a Ca^{2+}-Binding Protein is Induced by Abscisic Acid during Dehydration.” Plant and Cell Physiology 41 (7): 898-903. doi:https://doi.org/10.1093/pcp/pcd010.

Tang, T, A Yu, P Li, H Yang, G Liu, and L Liu. 2016. “Sequence analysis of the Hsp70 family in moss and evaluation of their functions in abiotic stress responses.” Scientific Reports 6. doi:https://doi.org/10.1038/srep33650.

Tavaria, M, T Gabriele, I Kola, and R L Anderson. 1996. “A hitchhiker's guide to the human Hsp70 family.” Cell Stress & Chaperones 1 (1): 23-28. doi:https://doi.org/10.1379%2F1466-1268(1996)001%3C0023%3Aahsgtt%3E2.3.co%3B2.

Tebele, M S, A R Marks, and J M Farrant. 2021. “Two Decades of Desiccation Biology: A Systematic Review of the Best Studied Angiosperm Resurrection Plants.” Plants 10 (12). doi:https://doi.org/10.3390/plants10122784.

Tripathy, B C, and R Oelmüller. 2012. “Reactive oxygen species generation and signalling in plants.” Plant Signalling & Behaviour 7 (12): 1621-1633. doi:https://doi.org/10.4161/psb.22455.

Tshabuse, F, J M Farrant, L Humbert, D Moura, D Rainteau, C Espinasse, A Idrissi, et al. 2018. “Glycerolipid analysis during desiccation and recovery of the resurrection plant *Xerophyta humilis* (Bak) Dur and Schinz.” Plant, Cell, & Environment 41 (3): 533-547. doi: https://doi.org/10.1111/pce.13063.

Tymms, M J, and D F Gaff. 1979. “Proline Accumulation during Water Stress in Resurrection Plants.” Journal of Experimental Botany 30 (1): 165-168. doi:https://doi.org/10.1093/jxb/30.1.165.

Vacic, V, C J Oldfield, A Mohan, P Radivojac, M S Cortese, V N Uversky, and A K Dunker. 2008. “Characterization of molecular recognition features, MoRFs, and their binding partners.” Journal of Proteome Research 6 (6). doi:10.1021/pr0701411.

van der Pas, L, and R Ingle. 2019. “Towards an Understanding of the Molecular Basis of Nickel Hyperaccumulation in Plants.” Plants 8 (1). doi:10.3390/plants8010011.

Van der Willigen, C, N W Pammenter, S Mundree, and J M Farrant. 2001. “Some physiological comparisons between the resurrection grass, *Eragrostis nindensis*, and the related desiccation-sensitive species, *E. curvula*.” Plant Growth Regulation 35 (2): 121-129. doi:https://doi.org/10.1023/A:1014425619913.

Van Oudtshoorn, F. 2012. Guide to Grasses of Southern Africa. Third. Briza Publications.

VanBuren, R, C M Wai, J Pardo, V Giarola, S Ambrosini, X Song, and D Bartels. 2018. “Desiccation Tolerance Evolved through Gene Duplication and Network Rewiring in Lindernia.” The Plant Cell 30 (12): 2943-2958. doi:https://doi.org/10.1105/tpc.18.00517.

VanBuren, R, C M Wai, V Giarola, M Zupunski, J Pardo, M Kalinowski, G Grossmann, and D Bartels. 2023. “Core cellular and tissue-specific mechanisms enable desiccation tolerance in Craterostigma.” The Plant Journal. doi:https://doi.org/10.1111/tpj.16165.

VanBuren, R, J Wai, Q Zhang, X Song, P P Edgar, D Bryant, and D Bartels. 2017. “Seed desiccation mechanisms co-opted for vegetative desiccation in the resurrection grass Oropetium thomeaum.” Plant, Cell & Environment 40. doi:https://doi.org/10.1111/pce.13027.

Varshney, R K, K C Bansal, P K Aggarwal, S K Datta, and P Q Crauford. 2011. “Agricultural biotechnology for crop improvement in a variable climate: hope or hype?” Trends in Plant Science 16 (7): 363-371. doi:10.1016/j.tplants.2011.03.004.

Veljovic-Jovanovic, S, B Kukavica, B Stevanovic, and F Navari-Izzo. 2006. “Senescence- and drought related changes in peroxidase and superoxide dismutase isoforms in leaves of *Ramonda serbica*.” Journal of Experimental Botany 57 (8): 1759-1768. doi:https://doi.org/10.1093/jxb/erl007 .

Vidović, M, I Battisti, A Pantelić, F Morina, G Arrigoni, A Masi, and S V Jovanović. 2022. “Desiccation Tolerance in Ramonda serbica Panc.: An Integrative Transcriptomic, Proteomic, Metabolite and Photosynthetic Study.” Plants 11 (9). doi:https://doi.org/10.3390/plants11091199.

Viviana, E, M Diaz, P Polci, and L Mroginski. 2001. “Embryogenic Cell Suspensions From Different Explants and Cultivars of *Eragrostis curvula* (Schrad.) Nees.” BIOCELL 25 (2): 131-138.

Vlot, A C, D A Dempsey, and D F Klessig. 2009. “Salicylic Acid, a multifaceted hormone to combat disease.” Annual Review Phytopathology 41: 177-206. doi:10.1146/annurev.phyto.050908.135202.

Vogel, C, and E M Marcotte. 2012. “Insights into the regulation of protein abundance from proteomic and transcriptomic analyses.” Nature Reviews Genetics 13: 227-232. doi:https://doi.org/10.1038/nrg3185.

Wang, J, L Song, X Gong, J Xu, and M Li. 2020. “Functions of Jasmonic Acid in Plant Regulation and Response to Abiotic Stress.” International journal of Molecular Science 21 (4). doi:https://doi.org/10.3390%2Fijms21041446.

Wang, M, S Gao, W Zeng, Y Yang, J Ma, and Y Wang. 2020. “Plant Virology Delivers Diverse Toolsets for Biotechnology.” Viruses 12 (11). doi:https://doi.org/10.3390/v12111338.

Wang, W, Z Peng, and J Yang. 2022. “Single-sequence protein structure prediction using supervised transformer protein language models.” Nature Computational Science 2: 804-814. doi:https://doi.org/10.1038/s43588-022-00373-3.

Wang, X, B Yan, M Shi, W Zhou, D Zekria, H Wang, and G Kai. 2016. “Overexpression of a *Brassica campestris* HSP70 in tobacco confers enhanced tolerance to heat stress.” Protoplasma 253: 637-645. doi:https://doi.org/10.1007/s00709-015-0867-5.

Wang, X, M S Rehmani, Q Chen, J Yan, P Zhao, C Li, Z Zhai, N Zhou, B Yang, and Y Q Jiang. 2022. “Rapeseed NAM transcription factor positively regulates leaf senescence via controlling senescence-associated gene expression.” Plant Science 323. doi:https://doi.org/10.1016/j.plantsci.2022.111373.

Wang, X, Y Niu, and Y Zheng. 2021. “Multiple Functions of MYB Transcription Factors in Abiotic Stress Responses.” International Journal of Molecular Science 22 (11). doi:10.3390/ijms22116125.

Wani, S H, V Kumar, V Shiram, and S K Sah. 2016. “Phytohormones and their metabolic engineering for abiotic stress tolerance in crop plants.” The Crop Journal 4: 162-176. doi:http://dx.doi.org/10.1016/j.cj.2016.01.010.

Watanabe, M, S Balazadeh, T Tohge, A Erban, P Giavalisco, J Kopka, B Meuller-Roeber, A R Fernie, and R Hoefgen. 2013. “Comprehensive dissection of spatiotemporal metabolic shifts in primary, secondary, and lipid metabolism during developmental senescence in Arabidopsis.” Plant Physiology 162 (3): 1290-1310. doi:https://doi.org/10.1104/pp.113.217380.

Wheeler, M CG, C L Arias, S Righini, M B Badia, C S Andreo, M F Drincovich, and M Saigo. 2016. “Differential Contribution of Malic Enzymes during Soybean and Castor Seeds Maturation.” PloS One 11 (6). doi:https://doi.org/10.1371%2Fjournal.pone.0158040.

Whittaker, A, A Bochicchio, C Vazzana, G Lindsey, and J M Farrant. 2001. “Changes in leaf hexokinase activity and metabolite levels in response to drying in the desiccation-tolerant species *Sporobolus stapfianus* and *Xerophyta viscosa*.” Journal of Experimental Botany 52 (358): 961-969. doi:https://doi.org/10.1093/jexbot/52.358.961.

Winans, S C. 1991. “An Agrobacterium two-component regulatory system for the detection of chemicals released from plant wounds.” Molecular Microbiology 5: 2345-2350. doi:https://doi.org/10.1111/j.1365-2958.1991.tb02080.x.

Winans, S C. 1990. “Transcriptional induction of an Agrobacterium regulatory gene at tandem promoters by plant-released phenolic compounds, phosphate starvation, and acidic growth media.” Journal of Bacteriology 172: 2433-2438. doi:10.1128/jb.172.5.2433-2438.1990.

Woldemariam, M G, N Onkokesung, I T Baldwin, and I Galis. 2012. “Jasmonoyl-L-isoleucine hydrolase 1 (JIH1) regulates jasmonoyl-L-isoleucine levels and attenuates plant defences against herbivores.” The Plant Journal for Cell and Molecular Biology 72 (5): 758-767. doi:10.1111/j.1365-313X.2012.05117.x. .

Wruck, F, M J Avellaneda, E J Koers, D P Minde, M P Mayer, G Kramer, A Mashaghi, and S J Tans. 2018. “Protein Folding Mediated by Trigger Factor and Hsp70: New Insights from Single-Molecule Approaches.” Journal of Molecular Biology 430 (4): 438-449. doi:https://doi.org/10.1016/j.jmb.2017.09.004.

Wu, Q, Z Peng, Y Zhang, and J Yang. 2018. “COACH-D: improved protein-ligand binding sites prediction with refined ligand-binding poses through molecular docking.” Nucleic Acids Research 2 (46). doi:10.1093/nar/gky439. .

Xu, S M, E Brill, D J Llewellyn, R T Furbank, and Y L Ruan. 2012. “Overexpression of a potato sucrose Synthase gene in cotton accelerates leaf expansion, reduces seed abortion, and enhances fiber production.” Molecular Plant 5 (2): 430-441. doi:https://doi.org/10.1093/mp/ssr090.

Xu, X M, H Lin, J Maple, B Björkblom, G Alves, J P Larsen, and S G Møller. 2010. “The Arabidopsis DJ-1a protein confers stress protection through cytosolic SOD activation.” Journal of Cell Science 123 (10). doi:https://doi.org/10.1242/jcs.063222.

Xu, X, C Zheng, D Lu, CP Song, and L Zhang. 2021. “Phase separation in plants: New insights into cellular compartmentalization.” Journal of Integrative Plant Biology 63 (11): 1835-1855. doi:https://doi.org/10.1111/jipb.13152.

Xu, X, S Legay, K Sergeant, S Zorzan, C C Leclercq, S Charton, V Giarola, et al. 2021. “Molecular insights into plant desiccation tolerance: transcriptomics, proteomics and targeted metabolite profiling in *Craterostigma plantagineum*.” The Plant Journal 107: 377-398. doi: 10.1111/tpj.15294.

Yan, Y, J Gan, Y Tao, T W Okita, and L Tian. 2022. “RNA-Binding Proteins: The Key Modulator in Stress Granule Formation and Abiotic Stress Response.” Frontiers in Plant Science 13. doi:https://doi.org/10.3389/fpls.2022.882596.

Yang, J, A Roy, and Y Zhang. 2013. “Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment.” Bioinformatics 15 (29). doi:10.1093/bioinformatics/btt447.

Yobi, A, K A Schlauch, R L Tillett, W C Yim , C Espinoza, B W.M Wone, J C Cushman, and M J Oliver. 2017. “*Sporobolus stapfianus*: Insights into desiccation tolerance in the resurrection grasses from linking transcriptomics to metabolomics.” BMC Plant Biology 17 (67). doi:10.1186/s12870-017-1013-7.

Yoo, S-D, Y-H Cho, and J Sheen. 2007. “Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene analysis.” Nature Protocols 2 (7): 1565-1572. doi:doi:10.1038/nprot.2007.199.

Yoshiba, Y, T Kiyosue, K Nakashima, K Yamaguchi-Shinozaki, and K Shinozaki. 1997. “Regulation of levels of proline as an osmolyte in plants under water stress.” Plant Cell Physiology 38 (10): 1095-1102. doi:10.1093/oxfordjournals.pcp.a029093 .

Yu, E M, T Yoshinaga, F L Jalufka, H Ehsan, D B Mark Welch, and G Kaneko. 2021. “The complex evolution of the metazoan HSP70 gene family.” Scientific Reports 11. doi:https://doi.org/10.1038/s41598-021-97192-9.

Yusof, N A, M Masnoddin, J Charles, Y Q Thien, F N Nasib, C M,V,L Wong, A Murad, A Munir, N M Mahadi, and I Bharudin. 2022. “Can heat shock protein 70 (HSP70) serve as biomarkers in Antarctica for future ocean acidification, warming and salinity stress?” Polar Biology 45: 371-394. doi:https://doi.org/10.1007/s00300-022-03006-7.

Zamore, P D. 2001. “Thirty-Three Years Later, a Glimpse at the Ribonuclease III Active Site.” Molecular Cell 8 (6): 1158-1160. doi:https://doi.org/10.1016/S1097-2765(01)00418-X.

Zeng, D D, R Qin, M Li, M Alamin, X-L Jin, Y Liu, and C-H Shi. 2017. “The ferredoxin-dependent glutamate synthase (OsFd-GOGAT) participates in leaf senescence and the nitrogen remobilization in rice.” Molecular Genetics and Genomics 292: 385-395. doi:https://doi.org/10.1007/s00438-016-1275-z.

Zhang, H, R W Kumimoto, S Anver, and S Harmer. 2023. "XAP5 CIRCADIAN TIMEKEEPER regulates RNA splicing and the circadian clock by genetically separable pathways." Plant Physiology 192 (3): 2492-2506. doi:https://doi.org/10.1093/plphys/kiad193.

Zhang, C L, K Mao, L J Zhou, G L Wang, Y L Zhang, Y Y Li, and Y J Hao. 2018. “Genome-wide identification and characterization of apple long-chain Acyl-CoA synthetases and expression analysis under different stresses.” Plant Physiology and Biochemistry 132: 320-332. doi:https://doi.org/10.1016/j.plaphy.2018.09.004.

Zhang, Q, X Song, and D Bartels. 2016. “Enzymes and Metabolites in Carbohydrate Metabolism of Desiccation Tolerant Plants.” Proteomes 4 (4). doi:https://doi.org/10.3390/proteomes4040040.

Zhao, H, D K Kosma, and S Lü. 2021. “Functional Role of Long-Chain Acyl-CoA Synthetases in Plant Development and Stress Responses.” Frontiers in Plant Science 12. doi:/doi.org/10.3389/fpls.2021.640996.

Zhao, Z, Z Peng, and J Yang. 2018. “Improving sequence-based prediction of protein-peptide binding residues by introducing intrinsic disorder and a consensus method.” Journal of Chemical Information and Modelling 58. doi:10.1021/acs.jcim.8b00019 .

Zhu, J, P M Oger, B Schrammeijer, P J.J Hooykaas, S K Farrand, and S C Winans. 2000. “The Bases of Crown Gall Tumorigenesis.” Journal of Bacteriology 182 (14): 3885-3895. doi:10.1128/jb.182.14.3885-3895.2000.

Zia, A, J W Berkley, H MO Oung, D Chavuri, P Jahns, A B Courins, J M Farrant, Z Reich, and H Kirchoff. 2016. “Protection of the photosynthetic apparatus against dehydration stress in the resurrection plant *Craterostigma pumilim*.” The Plant Journal 87. doi:10.1111/tpj.13227.

Appendix

Table 8-1: *Eragrostis nindensis* isolated protein with sample ID, concentration, and number of samples per relative water content. NST= non-senescent tissue, ST= senescent tissue. Protein was quantified using BCA protein detection kit and estimated from a BSA standard curve. Label refers to notation used to distinguish tissue and RWC throughout this Chapter

Tissue and RWC	Sample ID	Concentration $\mu g.\mu L^{-1}$	Number of samples	Label
100% NST	100N3	1,38	3	**N/A**
	100N4	1,39		
	100N5	1,27		
100% ST	100S1	2,47	4	**N/A**
	100S3	1,28		
	100S4	1,48		
	100S5	1,54		
25% ST	25S3	3,09	3	**25ST**
	24S4	2,85		
	25S5	2,06		
25% NST	25N1	1,86	5	**25NST**
	25N2	2,32		
	25N3	3,90		
	25N4	4,08		
	25N5	2,85		
20% ST	20S1	1,73	3	**25ST**
	20S2	1,88		
	20S3	2,68		
40% NST	40N1	2,59	6	**40NST**
	40N2	1,90		
	40N3	2,14		
	40N4	2,23		
	40N5	2,17		
	40N6	2,73		
60% NST	60N1	2,52	6	**60NST**
	60N2	2,17		
	60N3	1,97		
	60N4	1,81		
	60N5	1,85		
	60N6	2,72		
60% ST	60S1	2,66	6	**60ST**
	60S2	1,69		
	60S3	1,41		
	60S4	1,96		
	60S5	2,14		
Air-dry NST	AD1	3,10	6	**ADNST**
	AD2	1,57		
	AD3	2,76		
	AD4	2,93		
	AD5	2,14		
	AD6	2,83		
48HRS Post-rehydration	48N1	2,64	6	**48hrsNST**
	48N2	2,23		
	48N3	2,33		
	48N4	3,14		
	48N5	3,46		
	48N6	1,34		
1-week post-rehydration	WN1	1,64	6	**W1NST**
	WN2	1,92		
	WN3	1,69		
	WN4	1,79		
	WN5	2,43		
	WN6	2,37		

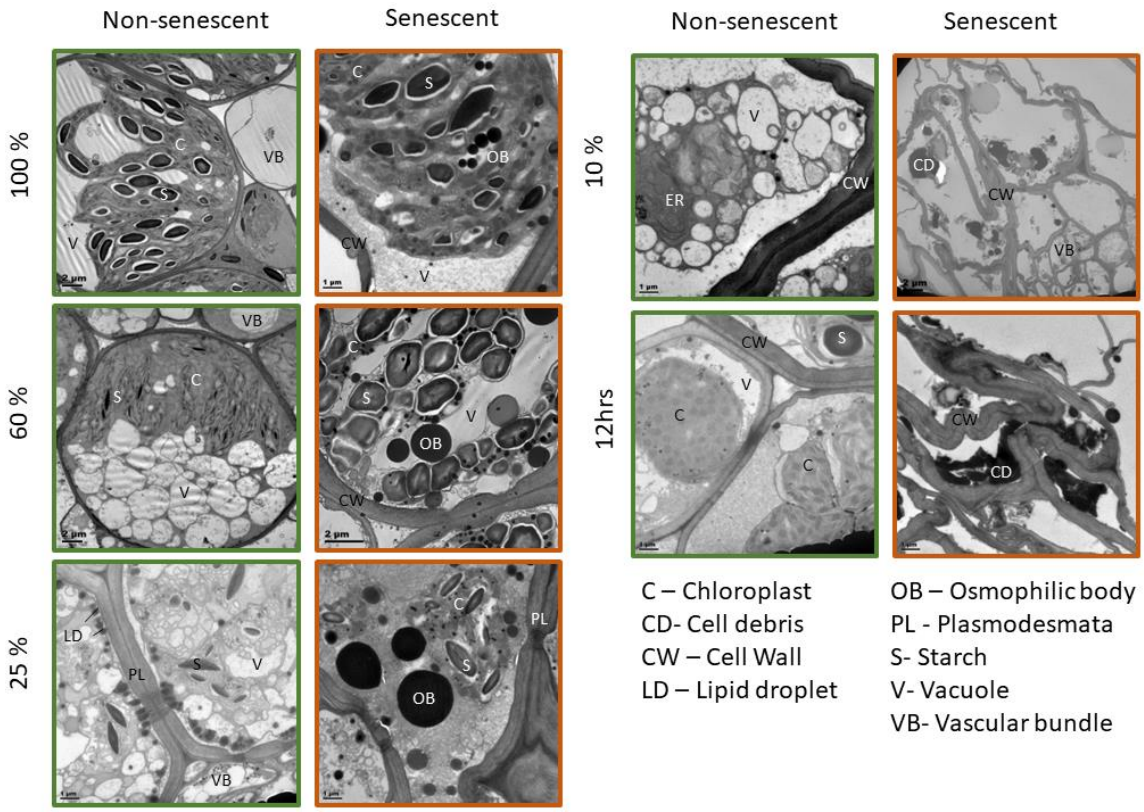


Figure 8-1: E. nindensis TEM images*: Selection of transmission electron microscopy images of E. nindensis NST and ST at different RWCs from Madden (2019).*

Transcriptome and proteome complementation description

Spearman correlation between transcriptome and proteome for matching KEGG matching terms for 60, 25%, and AD NST, and 60 and 25% ST indicated very weak correlation between the two with a mean correlation rank across these RWCs of 0.105 ± 0.08. This is mostly attributed to instances of transcript appearance but absence of protein. Covariance was highest for 25 ST (5.92) and 40 NST (5.55) and lowest for 25 NST (-0.2) and 60 ST (0.02). AD NST and 60 NST had third and fourth highest covariance of 4.12 and 3.28 respectively. The general positive covariance between the two data sets, except for 25 NST and 60 ST is suggestive that the two tissues had similar expression patterns suggesting to some degree that when transcripts are up-regulated proteins are too and *vice versa*. In terms of absolute numbers, 293 pairs of matching KEGG terms were shared between the two sets not taking duplicates into account. By looking at these shared terms across RWCs, there were 862 instances of shared expression pattern of which 621 were up-regulated in both accounting for 49% of all shared KEGG descriptions across all related RWCs. There were only 75 instances of differential expression in either data set direction (i.e., up in the transcriptome but down in the proteome etc) which accounted for 4%. A total of 505 instances of transcript presence but absence of protein was observed and 304 instances of protein presence but transcript

absence. Together the presence-absence observation accounted for 46% compared to 49% of shared expression. The overarching conclusion from these numbers is that there does not appear to be translational delay. While there are instances of transcripts present and no protein product the same can be said for protein presence and no transcript.

Table 8-2: Summary of Spearman correlation and covariance and number of transcript and protein pairs between shared KEGG terms from the transcriptome and proteome per RWC and tissue.

	60 NST	40 NST	25 NST	AD NST	60 ST	25 ST	Total
Spearman	0.189	0.185	0.01	0.098	0.023	0.128	N/A
Covariance	3.28	5.55	-0.20	4.12	0.02	5.92	N/A
Up-regulated in both	113	102	96	89	139	82	621
Down-regulated in both	31	55	10	69	20	56	241
Differential expression - Proteome	10	18	1	6	13	27	75
+Transcriptome + Proteome	40	70	172	96	51	76	505
- Transcriptome	97	46	12	31	68	50	304

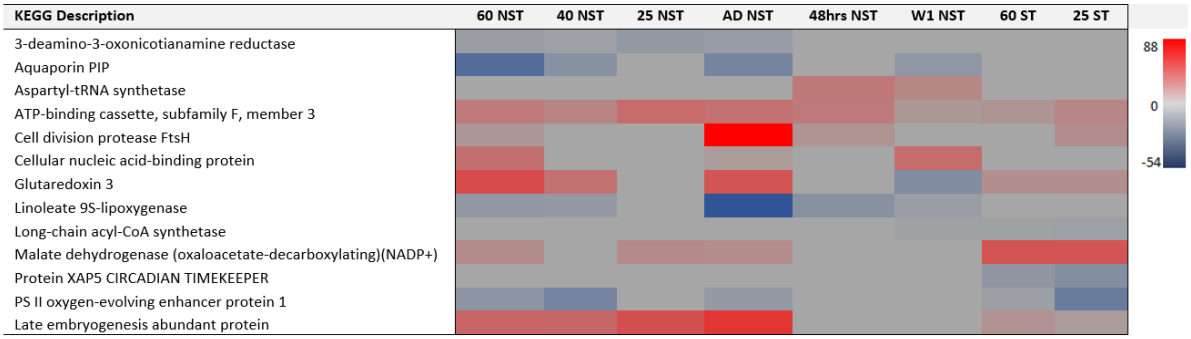


Figure 8-2: Highest and lowest fold change DAPs per RWC: Heatmap showing the highest and lowest fold change DAPs per RWC as annotated by Koala Blast. Values are discussed in Chapter 2 section 2.3.2. All DAPs represent log2 fold change >2, <0.05 FDR relative to fully hydrated leaves. N=non-senescent; S= senescent; AD= airdry; 48hr=48hrs post-rehydration; W1= one week post-rehydration. All other numbers refer to the RWC.

String interaction was compiled by using the associated Arabidopsis homologues used for annotating the transcriptome (Madden, 2019).

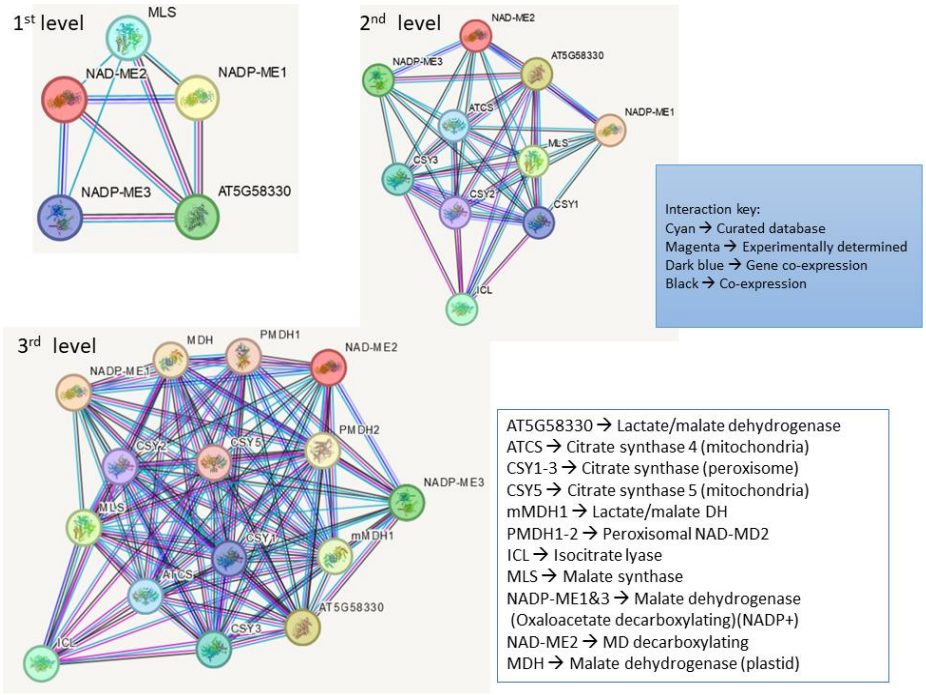


Figure 8-3: String network interaction for selected malate dehydrogenase proteins: String network showing first, second, and third level of interaction of malate dehydrogenase (MD) DAPs. Network interaction indicates the five selected MD DAPs are connected to several proteins involved in the TCA cycle. Relevance is discussed in Chapter 2 section 2.3.2.1.

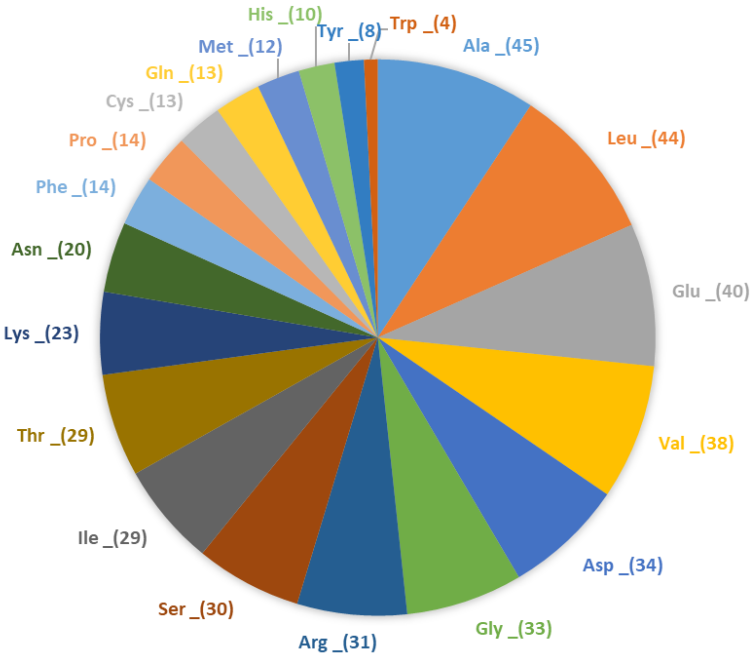


Figure 8-4: EnHSP70 Amino acid composition: Pie chart showing percentage contribution of amino acids to EnHSP70 sequence retrieved from https://web.expasy.org/cgi-bin/protparam/protparam. Numbers represent the total of that particular amino acid.

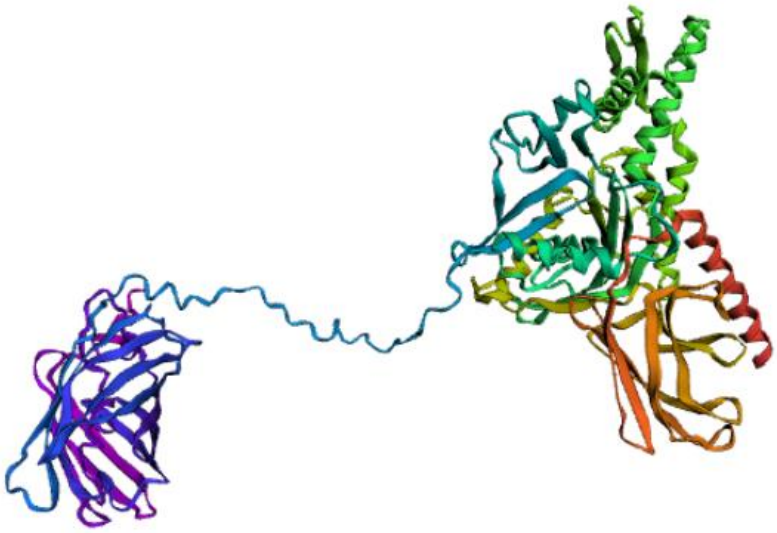


Figure 8-5: mCherry::EnHSP70 three dimensional model: 3-D model showing the mCherry::EnHSP70 fusion product used for the phase separation and lactate dehydrogenase assays. Modelling done using the online tRosetta tool with default parameters (Du, et al., 2021; Wang, et al., 2022; Su, et al., 2021).

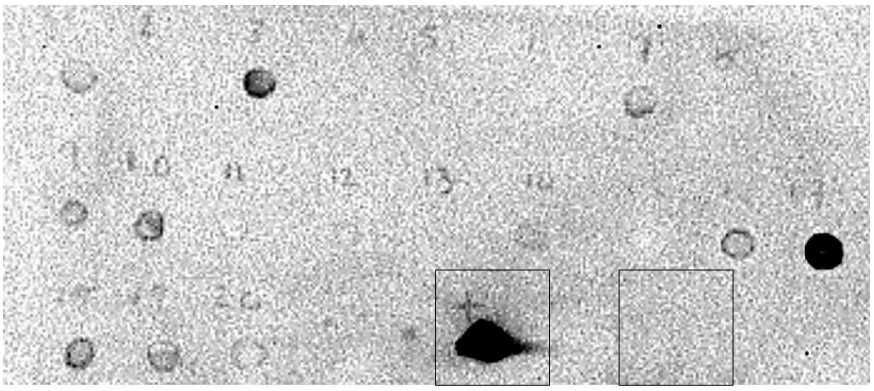


Figure 8-6: Dot blot of EnHSP70 transient expression in tobacco: Dot blot of total protein isolated from 20 different leaves of *N. benthamiana* injected with Agrobacteria harbouring the overexpressing EnHSP70 in pB2WG7 construct generated in Chapter 4 five days post-injection. 6X-His tagged mCherry served as positive control and empty pB2WG7 vector as negative control.

Sequences used for molecular cloning

Areas in bold and underlined are the primer sites as outlined in Table 4-1

EnHSP70

Product size: 302 bp

ATGGCGGCCTCCCAGAGCCAGACCCAAGGAGATGCGCCGGCGATTGGCATCTACCTGGGGACAACCAACTCGTGCGTCG
CCGTGTGTCGGCGCCGGCGTGACCGCGTCGAGATCATAGCCAACGACCGGGGCGACCGCCTCACGCCGTCCTGCGTGG
CATTCACCGACTGTGAGAAGCTCGTCGGCGAGGTGGCCGTGGACCAGGTCTCCTCGAACCCCGCCAACACCATCTACGG
TGAGAGCTCAAGCTTCAGCGCTCGCTCGCATTTACTTTGCCTTCGTCGTCAGGCCACCGTCGATGCCGCCAAGATCGCCG
GCCTGAATGTTCTGCACATTATCAACGAGGCCACCTCTGCGGCTATCGCCTACGGTCTTGACGCTATGCCGGCTGGAAAT
GAAGGGAGGACGGTGCTT**GTGTTTGATCTTGGCGGTGG**TGCCTTGGATGTCTCGCTCCTTGAAGTTGACCGGGGCGAAG
ACATAGATATGGCTCTCTTCGAAGTGAAGGCCATCGCCGGCGACACTCATCTTGGCGGGGTTGATTTCAACAATGAGATG
GTAGAATACCTTCTGCTTGAGTTAATAAGGAAACACAAGAAGGCGACCATCAGAAACAACCCAAGGGCACTCCGGCGGCT
GAGGTTGGCCTGCGAGAGAGTAAAGAAGTCCCTGGCTTTATTGACAGAGACCACAATTGAGG**AGTGCTCGCTACATGATG
GG**ATAGACTTCTGCATTACCATCACCCGCTCCCGATTCGAGGAGCTCAACGAGGCCCACTTCAGCAAGTGCATGGAGGCT
GTCGAGAAGTGCCTCCATGACGTCAAGATGGACAAGAGCGCGGTGCAGGATGTTGTGCTTGTGGGTGGCTCCACCCGCA
TCCCCAAGGTGCAGAGCATGCTCCGGGAATTCTTCGACGGGAAGAGGCTTTGCCAAACCATCAACCCCGACGAGGCAGTT
GCGTATGGAGCGGCCATCCACGCTGCCAATCTCAGCGACGCTAGCGGCAAGAAGGTGTGGGACTTGTTGCTTATGGATGT
CACGCCGTTCTCTCTGGGTGTGGAGGCGGACTGGGGTGACATGACTGTGCTGATACCGAGGAACACCCCTGTTCCCACC
AGGAAGGAGAAAGTATTCAGTACCTACCATGACAATCAAGCTGAGGTGCTCATCCAGGTGTACGAGGGAGAAGGTTGGAG
TGCAGAGGAGAACAATCTGCTCGGCAAATTTGAGCTTGCTGGCATCCATCCGGCGCCCAGAGGTATCCCCGAGATCATTG
TGTGCTTTTACGTGGATGCCAACGGCATTCTGACAGTGTCAGCTGAGGACAAGGAGACGGGGCAGACGAACCAGATCACC
ATTGTTACCGACGATAGTTGCAGGCTGAGCAAGGAGGAGATGGAACGCATGATGCAGGAGTCTGAGAAGGATAGGGGGA
CTTGGGAAGACAGGAAGTCCCATCACCATCACCATCACTAG

EnHSP70 amino acid sequence

MAASQSQTQGDAPAIGIYLGTTNSCVAVCRRRRDRVEIIANDRGDRLTPSCVAFTDCEKLVGEVAVDQVSSNPANTIYGESSSFS
ARSHLLCLRRQATVDAAKIAGLNVLHIINEATSAAIAYGLDAMPAGNEGRTVLVFDLGGGALDVSLLEVDRGEDIDMALFEVKAIA
GDTHLGGVDFNNEMVEYLLLELIRKHKKATIRNNPRALRRLRLACERVKKSLALLTETTIEECSLHDGIDFCITITRSRFEELNEAHF
SKCMEAVEKCLHDVKMDKSAVQDVVLVGGSTRIPKVQSMLREFFDGKRLCQTINPDEAVAYGAAIHAANLSDASGKKVWDLLL
MDVTPFSLGVEADWGDMTVLIPRNTPVPTRKEKVFSTYHDNQAEVLIQVYEGEGWSAEENNLLGKFELAGIHPAPRGIPEIIVCF
YVDANGILTVSAEDKETGQTNQITIVTDDSCRLSKEEMERMMQESEKDRGTWEDRKS

EnHSP70 in pB2WG7 sequence

EnHSP70 in purple, restriction sites in black and underlined, 6XHis tag in capital letters and underlined, and stop codon after 6XHis indicated in red.

ctcccatatggtcgactagagccaagctgatctcctttgccccggagatcaccatggacgactttctctatctctacgatctaggaagaaagttcgacggagaaggtgacgataccat
gttcaccaccgataatgagaagattagcctcttcaatttcagaaagaatgctgacccacagatggttagaggcctacgcggcaggtctcatcaagacgatctacccgagtaata
atctccaggagatcaaataccttcccaagaaggttaaagatgcagtcaaaagattcaggactaactgcatcaagaacacagagaaagatatatttctcaagatcagaagtactatt
ccagtatggacgattcaaggcttgcttcataaaccaaggcaagtaatagagattggagtctctaagaaagtagttcctactgaatcaaaggccatggagtcaaaaattcagatcgag
gatctaacagaactcgccgtgaagactggcgaacagttcatacagagtcttttacgactcaatgacaagaagaaaatcttcgtcaacatggtggagcacgacactctcgtctactcc
aagaatatcaaagatacagtctcagaagaccaaagggctattgagacttttcaacaaagggtaatatcgggaaacctcctcggattccattgcccagctatctgtcacttcatcaaaa
ggacagtagaaaaggaaggtggcacctacaaatgccatcattgcgataaaggaaaggctatcgttcaagatgcccctgccgacagtggtcccaaagatggacccccacccacg
aggagcatcgtggaaaaagaagacgttccaaccacgtcttcaaagcaagtggattgatgtgatatctccactgacgtaagggatgacgcacaatcccactatccttcgcaagacc
cttcctctatataaggaagttcatttcatttggagaggactccggtatttttacaacaattaccacaacaaaacaaacaacaaacaacattacaatttactatttctagtcgacctgcaggc
ggccgcactagtgatatcaccactttgtacaagaaagctgaacgagaaacgtaaaatgatataaatatcaatatattaaattagattttgcataaaaaacagactacataatactgta
aaacacaacatatccagtcactatggtcgacctgcagactggctgtgtataagggagcctgacatttatattccccagaacatcaggttaatggcgtttttgatgtcattttcgcggtggct
gagatcagccacttcttccccgataacggagaccggcacactggccatatcggtggtcatcatgcgccagctttcatccccgatatgcaccaccgggtaaagttcacgggagacttt
atctgacagcagacgtgcactggccaggggatcaccatccgtcgccccccc***ttaattaactagt***aaaATGGCGGCCTCCCAGAGCCAGACCCAAGGAGA
TGCGCCGGCGATTGGCATCTACCTGGGGACAACCAACTCGTGCGTCGCCGTGTGTCGGCGCCGGCGTGACCGCGTCGA
GATCATAGCCAACGACCGGGGCGACCGCCTCACGCCGTCCTGCGTGGCATTCACCGACTGTGAGAAGCTCGTCGGCGAG
GTGGCCGTGGACCAGGTCTCCTCGAACCCCGCCAACACCATCTACGGTGAGAGCTCAAGCTTCAGCGCTCGCTCGCATTT
ACTTTGCCTTCGTCGTCAGGCCACCGTCGATGCCGCCAAGATCGCCGGCCTGAATGTTCTGCACATTATCAACGAGGCCA
CCTCTGCGGCTATCGCCTACGGTCTTGACGCTATGCCGGCTGGAAATGAAGGGAGGACGGTGCTTGTGTTTGATCTTGGC
GGTGGTGCCTTGGATGTCTCGCTCCTTGAAGTTGACCGGGGCGAAGACATAGATATGGCTCTCTTCGAAGTGAAGGCCAT
CGCCGGCGACACTCATCTTGGCGGGGTTGATTTCAACAATGAGATGGTAGAATACCTTCTGCTTGAGTTAATAAGGAAACA
CAAGAAGGCGACCATCAGAAACAACCCAAGGGCACTCCGGCGGCTGAGGTTGGCCTGCGAGAGAGTAAAGAAGTCCCTG

GCTTTATTGACAGAGACCACAATTGAGGAGTGCTCGCTACATGATGGGATAGACTTCTGCATTACCATCACCCGCTCCCGA
TTCGAGGAGCTCAACGAGGCCCACTTCAGCAAGTGCATGGAGGCTGTCGAGAAGTGCCTCCATGACGTCAAGATGGACAA
GAGCGCGGTGCAGGATGTTGTGCTTGTGGGTGGCTCCACCCGCATCCCCAAGGTGCAGAGCATGCTCCGGGAATTCTTC
GACGGGAAGAGGCTTTGCCAAACCATCAACCCCGACGAGGCAGTTGCGTATGGAGCGGCCATCCACGCTGCCAATCTCA
GCGACGCTAGCGGCAAGAAGGTGTGGGACTTGTTGCTTATGGATGTCACGCCGTTCTCTCTGGGTGTGGAGGCGGACTG
GGGTGACATGACTGTGCTGATACCGAGGAACACCCCTGTTCCCACCAGGAAGGAGAAAGTATTCAGTACCTACCATGACA
ATCAAGCTGAGGTGCTCATCCAGGTGTACGAGGGAGAAGGTTGGAGTGCAGAGGAGAACAATCTGCTCGGCAAATTTGAG
CTTGCTGGCATCCATCCGGCGCCCAGAGGTATCCCCGAGATCATTGTGTGCTTTTACGTGGATGCCAACGGCATTCTGAC
AGTGTCAGCTGAGGACAAGGAGACGGGGCAGACGAACCAGATCACCATTGTTACCGACGATAGTTGCAGGCTGAGCAAG
GAGGAGATGGAACGCATGATGCAGGAGTCTGAGAAGGATAGGGGGACTTGGGAAGACAGGAAGTCC**CATCACCATCACC**
ATCACtag***ggatcc***taactcaaaatccacacattatacgagccggaagcataaagtgtaaagcctggggtgcctaatgcggccgccatagtgactggatatgttgtgttttacag
tattatgtagtctgtttttttatgcaaaatctaatttaatatattgatatttatatcattttacgtttctcgttcagctttttgtacaaacttgtgatatcccgcggccatgctagagtccgcaaaaatca
ccagtctctctctacaaatctatctctctctatttttctccagaataatgtgtgagtagttcccagataagggaattagggttcttatagggtttcgctcatgtgttgagcatataagaaaccctt
agtatgtatttgtatttgtaaaatacttctatcaataaaatttctaattcctaaaaccaaaatccagtgacctgcaggcatgcgacgtcgggcccaagcttagcttgagcttggatcagatt
gtcgtttcccgccttcagtttaaactatcagtgtttgacaggatatattggcgggtaaacctaagagaaaagagcgtttattagaataatcggatatttaaaagggcgtgaaaaggtttat
ccgttcgtccatttgtatgtgcatgccaaccacagggttcccctcgggatcaaagtactttaaagtactttaaagtactttaaagtactttgatccaacccctccgctgctatagtgcagtcg
gcttctgacgttcagtgcagccgtcttctgaaaacgacatgtcgcacaagtcctaagttacgcgacaggctgccgccctgccctttttcctggcgttttcttgtcgcgtgttttagtcgcataa
agtagaatacttgcgactagaaccggagacattacgccatgaacaagagcgccgccgctggcctgctgggctatgcccgcgtcagcaccgacgaccaggacttgaccaaccaa
cgggccgaactgcacgcggccggctgcaccaagctgttttccgagaagatcaccggcaccaggcgcgaccgcccggagctggccaggatgcttgaccacctacgccctggcg
acgttgtgacagtgaccaggctagaccgcctggcccgcagcacccgcgacctactggacattgccgagcgcatccaggaggccggcgcgggcctgcgtagcctggcagagcc
gtgggccgacaccaccacgccggccggccgcatggtgttgaccgtgttcgccggcattgccgagttcgagcgttccctaatcatcgaccgcacccggagcgggcgcgaggccgc
caaggcccgaggcgtgaagtttggcccccgccctaccctcaccccggcacagatcgcgcacgcccgcgagctgatcgaccaggaaggccgcaccgtgaaagaggcggctgc
actgcttggcgtgcatcgctcgaccctgtaccgcgcacttgagcgcagcgaggaagtgacgcccaccgaggccaggcggcgcggtgccttccgtgaggacgcattgaccgagg
ccgacgccctggcggccgccgagaatgaacgccaagaggaacaagcatgaaaccgcaccaggacggccaggacgaaccgttttcattaccgaagagatcgaggcggaga
tgatcgcggccgggtacgtgttcgagccgcccgcgcacgtctcaaccgtgcggctgcatgaaatcctggccggtttgtctgatgccaagctggcggcctggccggccagcttggcc
gctgaagaaaccgagcgccgccgtctaaaaaggtgatgtgtatttgagtaaaacagcttgcgtcatgcggtcgctgcgtatatgatgcgatgagtaaataaacaaatacgcaaggg
gaacgcatgaaggttatcgctgtacttaaccagaaaggcgggtcaggcaagacgaccatcgcaacccatctagcccgcgccctgcaactcgccggggccgatgttctgttagtcg
attccgatccccagggcagtgcccgcgattgggcggccgtgcgggaagatcaaccgctaaccgttgtcggcatcgaccgcccgacgattgaccgcgacgtgaaggccatcggc
cggcgcgacttcgtagtgatcgacggagcgccccaggcggcggacttggctgtgtccgcgatcaaggcagccgacttcgtgctgattccggtgcagccaagcccttacgacatatg
ggccaccgccgacctggtggagctggttaagcagcgcattgaggtcacggatggaaggctacaagcggcctttgtcgtgtcgcgggcgatcaaaggcacgcgcatcggcggtg
aggttgccgaggcgctggccgggtacgagctgcccattcttgagtcccgtatcacgcagcgcgtgagctacccaggcactgccgccgccggcacaaccgttcttgaatcagaacc
cgagggcgacgctgcccgcgaggtccaggcgctggccgctgaaattaaatcaaaactcatttgagttaatgaggtaaagagaaaatgagcaaaagcacaaacacgctaagtg
ccggccgtccgagcgcacgcagcagcaaggctgcaacgttggccagcctggcagacacgccagccatgaagcgggtcaactttcagttgccggcggaggatcacaccaagct
gaagatgtacgcggtacgccaaggcaagaccattaccgagctgctatctgaatacatcgcgcagctaccagagtaaatgagcaaatgaataaatgagtagatgaattttagcggc
taaaggaggcggcatggaaaatcaagaacaaccaggcaccgacgccgtggaatgccccatgtgtggaggaacgggcggttggccaggcgtaagcggctgggttgtctgccgg
ccctgcaatggcactggaaccccc aagcccgaggaatcggcgtgagcggtcgcaaaccatccggcccggtacaaatcggcgcggcgctgggtgatgacctggtggagaagtt
gaaggccgcgcaggccgcccagcggcaacgcatcgaggcagaagcacgccccggtgaatcgtggcaagcggccgctgatcgaatccgcaaagaatcccggcaaccgccg
gcagccggtgcgccgtcgattaggaagccgcccaagggcgacgagcaaccagatttttcgttccgatgctctatgacgtgggcacccgcgatagtcgcagcatcatggacgtgg
ccgttttccgtctgtcgaagcgtgaccgacgagctggcgaggtgatccgctacgagcttccagacgggcacgtagaggtttccgcagggccggccggcatggccagtgtgtgggat
tacgacctggtactgatggcggtttcccatctaaccgaatccatgaaccgatacccgggaagggaagggagacaagcccggccgcgtgttccgtccacacgttgcggacgtactca
agttctgccggcgagccgatggcggaaagcagaaagacgacctggtagaaacctgcattcggttaaacaccacgcacgttgccatgcagcgtacgaagaaggccaagaacg
gccgcctggtgacggtatccgagggtgaagccttgattagccgctacaagatcgtaaagagcgaaaccgggcggccggagtacatcgagatcgagctagctgattggatgtacc
gcgagatcacagaaggcaagaacccggacgtgctgacggttcaccccgattacttttgatcgatcccggcatcggccgttttctctaccgcctggcacgccgcgccgcaggcaag
gcagaagccagatggttgttcaagacgatctacgaacgcagtggcagcgccggagagttcaagaagttctgtttcaccgtgcgcaagctgatcgggtcaaatgacctgccggagt
acgatttgaaggaggaggcggggcaggctggcccgatcctagtcatgcgctaccgcaacctgatcgagggcgaagcatccgccggttcctaatgtacggagcagatgctaggg
caaattgccctagcaggggaaaaaggtcgaaaaggtctctttcctgtggatagcacgtacattgggaacccaaagccgtacattgggaaccggaacccgtacattgggaaccca
aagccgtacattgggaaccggtcacacatgtaagtgactgatataaaagagaaaaaaggcgatttttccgcctaaaactctttaaaacttattaaaactcttaaaacccgcctggcct
gtgcataactgtctggccagcgcacagccgaagagctgcaaaaagcgcctaccc ttcggtcgctgcgctccctacgccccgccgcttcgcgtcggcctatcgcggccgctggccg
ctcaaaaatggctggcctacggccaggcaatctaccagggcgcggacaagccgcgccgtcgccactcgaccgccggcgcccacatcaaggcaccctgcctcgcgcgtttcggt
gatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgg
gtgtcggggcgcagccatgacccagtcacgtagcgatagcggagtgtatactggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaatacc
gcacagatgcgtaaggagaaaataccgcatcaggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggt
aatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccatagg
ctccgccccccgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctc
ctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgg
gctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaaca
ggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttc
ggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttg
atcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgcatgatatatctcccaatttgtgtagggcttattatgcacgcttaaaaataataaagc
agacttgacctgatagtttggctgtgagcaattatgtgcttagtgcatctaacgcttgagttaagccgcgccgcgaagcggcgtcggcttgaacgaatttctagctagacattatttgccg
actaccttggtgatctcgcctttcacgtagtggacaaattcttccaactgatctgcgcgcgaggccaagcgatcttcttcttgtccaagataagcctgtctagcttcaagtatgacgggctg
atactgggccggcaggcgctccattgcccagtcggcagcgacatccttcggcgcgattttgccggttactgcgctgtaccaaatgcgggacaacgtaagcactacatttcgctcatcg
ccagcccagtcgggcggcgagttccatagcgttaaggtttcatttagcgcctcaaatagatcctgttcaggaaccggatcaaagagttcctccgccgctggacctaccaaggcaac
gctatgttctcttgcttttgtcagcaagatagccagatcaatgtcgatcgtggctggctcgaagatacctgcaagaatgtcattgcgctgccattctccaaattgcagttcgcgcttagctg
gataacgccacggaatgatgtcgtcgtgcacaacaatggtgacttctacagcgcggagaatctcgctctctccaggggaagccgaagtttccaaaaggtcgttgatcaaagctcgc
cgcgttgtttcatcaagccttacggtcaccgtaaccagcaaatcaatatcactgtgtggcttcaggccgccatccactgcggagccgtacaaatgtacggccagcaacgtcggttcga
gatggcgctcgatgacgccaactacctctgatagttgagtcgatacttcggcgatcaccgcttcccccatgatgtttaactttgttttagggcgactgccctgctgcgtaacatcgttgctg
ctccataacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatagactgtaccccaaaaaaacagtcataacaagccatgaaaaccgccactgcg
ccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacagcttacgaaccgaacaggcttatgtccactgggttcgtgcccgaattgat
cacaggcagcaacgctctgtcatcgttacaatcaacatgctaccctccgcgagatcatccgtgtttcaaacccggcagcttagttgccgttcttccgaatagcatcggtaacatgagca

aagtctgccgccttacaacggctctcccgctgacgccgtcccggactgatgggctgcctgtatcgagtggtgattttgtgccgagctgccggtcggggagctgttggctggctggtggc
aggatatattgtggtgtaaacaaattgacgcttagacaacttaataacacattgcggacgtttttaatgtactgaattaacgccgaattgaattatcagcttgcatgccggtcgatctagta
acatagatgacaccgcgcgcgataatttatcctagtttgcgcgctatatttgttttctatcgcgtattaaatgtataattgcgggactctaatcataaaaacccatctcataaataacgtcat
gcattacatgttaattattacatgcttaacgtaattcaacagaaattatatgataatcatcgcaagaccggcaacaggattcaatcttaagaaactttattgccaaatgtttgaacgatctg
cttgactctaggggtcatcagatttcggtgacgggcaggaccggacggggcggcaccggcaggctgaagtccagctgccagaaacccacgtcatgccagttcccgtgcttgaag
ccggccgcccgcagcatgccgcgggggggcatatccgagcgcctcgtgcatgcgcacgctcgggtcgttgggcagcccgatgacagcgaccacgctcttgaagccctgtgcctcc
agggacttcagcaggtgggtgtagagcgtggagcccagtcccgtccgctggtggcgggggggagacgtacacggtcgactcggccgtccagtcgtaggcgttgcgtgccttccagg
gacccgcgtaggcgatgccggcgacctcgccgtccacctcggcgacgagccagggatagcgctcccgcagacggacgaggtcgtccgtccactcctgcggttcctgcggctcg
gtacggaagttgaccgtgcttgtctcgatgtagtggttgacgatggtgcagaccgccggcatgtccgcctcggtggcacggcggatgtcggccgggcgtcgttctgggctcatggtag
atcccctcgatcgagttgagagtgaatatgagactctaattggataccgaggggaatttatggaacgtcagtggagcatttttgacaagaaatatttgctagctgatagtgaccttaggc
gacttttgaacgcgcaataatggtttctgacgtatgtgcttagctcattaaactccagaaacccgcggctcagtggctccttcaacgttgcggttctgtcagttccaaacgtaaaacggctt
gtcccgcgtcatcggcgggggtcataacgtgactcccttaattctccgctcatgatcgataattcgagct

mCherry::EnHSP70 sequence in RY8686

mCherry indicated in red, linker region in blue italics, EnHSP70 in capital purple, and restriction sites in black and underlined.

tggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttccct
tcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatggttca
cgtagtgggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcggtctattcttttga
tttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaactagtaacgtttacaatttcaggtggcacttttcggggaaat
gtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgaattaattcttagaaaaactcatcgagcatcaaatgaaactgcaatttattcatatcaggattat
caataccatatttttgaaaaagccgtttctgtaatgaaggagaaaactcaccgaggcagttccataggatggcaagatcctggtatcggtctgcgattccgactcgtccaacatcaata
caacctattaatttcccctcgtcaaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatggcaaaagtttatgcatttctttccagacttgttcaaca
ggccagccattacgctcgtcatcaaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgcctgagcgagacgaaatacgcgatcgctgttaaaaggacaattacaaaca
ggaatcgaatgcaaccggcgcaggaacactgccagcgcatcaacaatgttttcacctgaatcaggatattcttctaatacctggaatgctgttttcccggggatcgcagtggtgagta
accatgcatcatcaggagtacggataaaatgcttgatggtcggaagaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgccat
gtttcagaaacaactctggcgcatcgggcttcccatacaatcgatagattgtcgcacctgattgcccgacattatcgcgagcccatttatacccatataaatcagcatccatgttggaatt
taatcgcggcctagagcaagacgtttcccgttgaatatggctcataacacccttgtattactgtttatgtaagcagacagttttattgttcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccg
gatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcacc
gcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggct
gaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaagg
cggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttga
gcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatg
cggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgtgactg
ggtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtc
agaggttttcaccgtcatcaccgaaacgcgcgaggcagctgcggtaaagctcatcagcgtggtcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagctcgttgagtttc
tccagaagcgttaatgtctggcttctgataaagcgggccatgttaagggcggttttttcctgtttggtcactgatgcctccgtgtaagggggatttctgttcatgggggtaatgatataccgatg
aaacgagagaggatgctcacgatacgggttactgatgatgaacatgcccggttactggaacgttgtgagggtaaacaactggcggtatggatgcggcgggaccagagaaaaat
cactcagggtcaatgccagcgcttcgttaatacagatgtaggtgttccacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgcagggcgctgacttccgcgt
ttccagacttacgaaacacggaaaccgaagaccattcatgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatcggtgattcattctgctaacca
gtaaggcaacccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggggccgccatgccggcgataatggcctgcttctcgccgaaacgtttggtg
gcgggaccagtgacgaaggcttgagcgagggcgtgcaagattccgaataccgcaagcgacaggccgatcatcgtcgcgctccagcgaaagcggtcctcgccgaaaatgacc
cagagcgctgccggcacctgtcctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcgcccaccggaaggagctgactgggttgaaggct
ctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaacttacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatc
ggccaacgcgcggggagaggcggtttgcgtattgggcgccagggtggttttcttttcaccagtgagacgggcaacagctgattgcccttcaccgcctggccctgagagagttgcag
caagcggtccacgctggtttgccccagcaggcgaaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcac
caacgcgcagcccggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcatcgcagtgggaacgatgccctcattcagcatttgcatggtttgttgaa
aaccggacatggcactccagtcgccttcccgttccgctatcggctgaatttgattgcgagtgagatatttatgccagccagccagacgcagacgcgccgagacagaacttaatggg
cccgctaacagcgcgatttgctggtgacccaatgcgaccagatgctccacgcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgtctggtcagagacatca
agaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctggtcatccagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccg
ccgctttacaggcttcgacgccgcttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagatttaatcgccgcgacaatttgcgacggcgcgtgcagggcca
gactggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgtgccacgcggttgggaatgtaattcagctccgccatcgccgcttccacttttcccgcgttttcgcag
aaacgtggctggcctggttcaccacgcgggaaacggtctgataagagacaccggcatactctgcgacatcgtataacgttactggtttcacattcaccaccctgaattgactctcttcc
gggcgctatcatgccatacccgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgt
tgagcaccgccgccgcaaggaatggtgcatgcaaggagatggcgcccaacagtccccggccacggggcctgccaccatacccacgccgaaacaagcgctcatgagcccg
aagtggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgatgcgtccggcgtagaggatcgagatc
tcgatcccgcgaaattaatacgactcactatagggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggagatataccatgcaccatcaccatcac
catgggatcgaggaaaacctgtacttccaatccggttctggA

GATATTACATCCCATAACGAAGACTATACCATCGTGGAACAGTACGAGCGGGCCGAGGGCAGACATAGCACAGGGGGCAT
GGATGAATTGTACAAG*ggtgcaccgggttctgcgggttctgccgcaggt****ggatcc***cATGGCGGCCTCCCAGAGCCAGACCCAAGGAGATGCGC
CGGCGATTGGCATCTACCTGGGGACAACCAACTCGTGCGTCGCCGTGTGTCGGCGCCGGCGTGACCGCGTCGAGATCAT
AGCCAACGACCGGGGCGACCGCCTCACGCCGTCCTGCGTGGCATTCACCGACTGTGAGAAGCTCGTCGGCGAGGTGGC
CGTGGACCAGGTCTCCTCGAACCCCGCCAACACCATCTACGGTGAGAGCTCAAGCTTCAGCGCTCGCTCGCATTTACTTT
GCCTTCGTCGTCAGGCCACCGTCGATGCCGCCAAGATCGCCGGCCTGAATGTTCTGCACATTATCAACGAGGCCACCTCT
GCGGCTATCGCCTACGGTCTTGACGCTATGCCGGCTGGAAATGAAGGGAGGACGGTGCTTGTGTTTGATCTTGGCGGTG
GTGCCTTGGATGTCTCGCTCCTTGAAGTTGACCGGGGCGAAGACATAGATATGGCTCTCTTCGAAGTGAAGGCCATCGCC
GGCGACACTCATCTTGGCGGGGTTGATTTCAACAATGAGATGGTAGAATACCTTCTGCTTGAGTTAATAAGGAAACACAAG
AAGGCGACCATCAGAAACAACCCAAGGGCACTCCGGCGGCTGAGGTTGGCCTGCGAGAGAGTAAAGAAGTCCCTGGCTT
TATTGACAGAGACCACAATTGAGGAGTGCTCGCTACATGATGGGATAGACTTCTGCATTACCATCACCCGCTCCCGATTCG
AGGAGCTCAACGAGGCCCACTTCAGCAAGTGCATGGAGGCTGTCGAGAAGTGCCTCCATGACGTCAAGATGGACAAGAG
CGCGGTGCAGGATGTTGTGCTTGTGGGTGGCTCCACCCGCATCCCCAAGGTGCAGAGCATGCTCCGGGAATTCTTCGAC
GGGAAGAGGCTTTGCCAAACCATCAACCCCGACGAGGCAGTTGCGTATGGAGCGGCCATCCACGCTGCCAATCTCAGCG
ACGCTAGCGGCAAGAAGGTGTGGGACTTGTTGCTTATGGATGTCACGCCGTTCTCTCTGGGTGTGGAGGCGGACTGGGG
TGACATGACTGTGCTGATACCGAGGAACACCCCTGTTCCCACCAGGAAGGAGAAAGTATTCAGTACCTACCATGACAATCA
AGCTGAGGTGCTCATCCAGGTGTACGAGGGAGAAGGTTGGAGTGCAGAGGAGAACAATCTGCTCGGCAAATTTGAGCTTG
CTGGCATCCATCCGGCGCCCAGAGGTATCCCCGAGATCATTGTGTGCTTTTACGTGGATGCCAACGGCATTCTGACAGTG
TCAGCTGAGGACAAGGAGACGGGGCAGACGAACCAGATCACCATTGTTACCGACGATAGTTGCAGGCTGAGCAAGGAGG
AGATGGAACGCATGATGCAGGAGTCTGAGAAGGATAGGGGGACTTGGGAAGACAGGAAGTCCTAG***gcggcc***gcactcgagcacc
accaccaccaccactgagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataacccttggggcctctaaacgggtctt
gaggggttttttgctgaaaggaggaactatatccggat

mCherry::EnHSP70 sequence in pB2WG7

mCherry indicated in red, linker region in blue italics, EnHSP70 in capital purple, and restriction sites in black and underlined.

ctcccatatggtcgactagagccaagctgatctcctttgccccggagatcaccatggacgactttctctatctctacgatctaggaagaaagttcgacggagaaggtgacgataccat
gttcaccaccgataatgagaagattagcctcttcaatttcagaaagaatgctgacccacagatggttagaggcctacgcggcaggtctcatcaagacgatctacccgagtaata
atctccaggagatcaaatacctttcccaagaaggttaaagatgcagtcaaaagattcaggactaactgcatcaagaacacagagaaagatatatttctcaagatcagaagtactatt
ccagtatggacgattcaaggcttgcttcataaaccaaggcaagtaatagagattggagtctctaagaaagtagttcctactgaatcaaaggccatggagtcaaaaattcagatcgag
gatctaacagaactcgccgtgaagactggcgaacagttcatacagagtcttttacgactcaatgacaagaagaaaatcttcgtcaacatggtggagcacgacactctcgtctactcc
aagaatatcaaagatacagtctcagaagaccaaagggctattgagacttttcaacaaagggtaatatcgggaaacctcctcggattccattgcccagctatctgtcacttcatcaaaa
ggacagtagaaaaggaaggtggcacctacaaatgccatcattgcgataaaggaaaggctatcgttcaagatgcctctgccgacagtggtcccaaagatggacccccacccacg
aggagcatcgtggaaaaagaagacgttccaaccacgtcttcaaagcaagtggattgatgtgatatctccactgacgtaagggatgacgcacaatcccactatccttcgcaagacc
cttcctctatataaggaagttcatttcatttggagaggactccggtatttttacaacaatacccacaacaaaacaaacaacaaacaacattacaatttactattctagtcgacctgcaggc
ggccgcactagtgatatcaccactttgtacaagaaagctgaacgagaaacgtaaaatgatataaatatcaatatattaaattagattttgcataaaaaacagactacataatactgta
aaacacaacatatccagtcactatggtcgacctgcagactggctgtgtataagggagcctgacatttatattccccagaacatcaggttaatggcgtttttgatgtcattttcgcggtggct
gagatcagccacttcttccccgataacggagaccggcacactggccatatcggtggtcatcatgcgccagctttcatccccgatatgcaccaccgggtaaagttcacgggagacttt
atctgacagcagacgtgcactggccaggggatcaccatccgtcgccc**GGGAAA**atggtgtctaaaggcgaggaggataacatggctataatcaaggaatttatgaggttca
aggtgcatatggagggatctgtgaacggtcacgagttcgaaatcgagggcgaaggcgaggggcgaccctatgaagggacacagacggccaaactgaaggtgaccaagggtg
gcccccgccttccgcctgggacatcctgtcacccagtttatgtacggctctaaggcttacgttaagcacccgccgatattccggactacttgaaactgagcttccctgaagggttcaa
gtgggaacgggtaatgaattttgaggatggcggtgtagtcacagttacccaggacagctctcttcaggacggagaatttatctataaggttaaacttcggggcactaacttcccatcag
acggccccgtgatgcagaagaaaactatgggctgggaagccagctcagaacgcatgtatcccgaggacggagcactgaagggagaaatcaagcagcgactcaagctgaag
gatgggggacattatgatgctgaggtcaaaaccacctataaggccaagaaacccgttcagcttcctggggcatacaacgtgaatatcaaactggatattacatcccataacgaaga
ctataccatcgtggaacagtacgagcgggccgagggcagacatagcacaggggcatggatgaattgtacaag*GGTGCACCGGGTTCTGCGGGTTCTGCC*
*GCAGGT***GGATCC**ATGGCGGCCTCCCAGAGCCAGACCCAAGGAGATGCGCCGGCGATTGGCATCTACCTGGGGACAACC
AACTCGTGCGTCGCCGTGTGTCGGCGCCGGCGTGACCGCGTCGAGATCATAGCCAACGACCGGGGCGACCGCCTCACG
CCGTCCTGCGTGGCATTCACCGACTGTGAGAAGCTCGTCGGCGAGGTGGCCGTGGACCAGGTCTCCTCGAACCCCGCCA
ACACCATCTACGGTGAGAGCTCAAGCTTCAGCGCTCGCTCGCATTTACTTTGCCTTCGTCGTCAGGCCACCGTCGATGCC
GCCAAGATCGCCGGCCTGAATGTTCTGCACATTATCAACGAGGCCACCTCTGCGGCTATCGCCTACGGTCTTGACGCTAT
GCCGGCTGGAAATGAAGGGAGGACGGTGCTTGTGTTTGATCTTGGCGGTGGTGCCTTGGATGTCTCGCTCCTTGAAGTTG
ACCGGGGCGAAGACATAGATATGGCTCTCTTCGAAGTGAAGGCCATCGCCGGCGACACTCATCTTGGCGGGGTTGATTTC
AACAATGAGATGGTAGAATACCTTCTGCTTGAGTTAATAAGGAAACACAAGAAGGCGACCATCAGAAACAACCCAAGGGCA
CTCCGGCGGCTGAGGTTGGCCTGCGAGAGAGTAAAGAAGTCCCTGGCTTTATTGACAGAGACCACAATTGAGGAGTGCTC
GCTACATGATGGGATAGACTTCTGCATTACCATCACCCGCTCCCGATTCGAGGAGCTCAACGAGGCCCACTTCAGCAAGT
GCATGGAGGCTGTCGAGAAGTGCCTCCATGACGTCAAGATGGACAAGAGCGCGGTGCAGGATGTTGTGCTTGTGGGTGG
CTCCACCCGCATCCCCAAGGTGCAGAGCATGCTCCGGGAATTCTTCGACGGGAAGAGGCTTTGCCAAACCATCAACCCCG
ACGAGGCAGTTGCGTATGGAGCGGCCATCCACGCTGCCAATCTCAGCGACGCTAGCGGCAAGAAGGTGTGGGACTTGTT
GCTTATGGATGTCACGCCGTTCTCTCTGGGTGTGGAGGCGGACTGGGGTGACATGACTGTGCTGATACCGAGGAACACC
CCTGTTCCCACCAGGAAGGAGAAAGTATTCAGTACCTACCATGACAATCAAGCTGAGGTGCTCATCCAGGTGTACGAGGG
AGAAGGTTGGAGTGCAGAGGAGAACAATCTGCTCGGCAAATTTGAGCTTGCTGGCATCCATCCGGCGCCCAGAGGTATCC
CCGAGATCATTGTGTGCTTTTACGTGGATGCCAACGGCATTCTGACAGTGTCAGCTGAGGACAAGGAGACGGGGCAGACG
AACCAGATCACCATTGTTACCGACGATAGTTGCAGGCTGAGCAAGGAGGAGATGGAACGCATGATGCAGGAGTCTGAGAA
GGATAGGGGGACTTGGGAAGACAGGAAGTCCTAG**ACGCGT**ttacgccccgccctgccactcatcgcagtactgttgtaattcattaagcattctgccgac
atggaagccatcacagacggcatgatgaacctgaatcgccagcggcatcagcaccttgtcgccttgcgtataatatttgcccatggtgaaaacgggggcgaagaagttgtccatatt

ggccacgtttaaatcaaaactggtgaaactcacccagggattggctgagacgaaaaacatattctcaataaaccctttagggaaataggccaggttttcaccgtaacacgccacat
cttgcgaatatatgtgtagaaactgccggaaatcgtcgtggtattcactccagagcgatgaaaacgtttcagtttgctcatggaaaacggtgtaacaagggtgaacactatcccatatc
accagctcaccgtctttcattgccatacggaattccggatgagcattcatcaggcgggcaagaatgtgaataaaggccggataaaacttgtgcttatttttctttacggtctttaaaaagg
ccgtaatatccagctgaacggtctggttataggtacattgagcaactgactgaaatgcctcaaaatgttctttacgatgccattgggatatatcaacggtggtatatccagtgatttttttctc
cattttagcttccttagctcctgaaaatctcgccggatcctaactcaaaatccacacattatacgagccggaagcataaagtgtaaagcctggggtgcctaatgcggccgccatagtg
actggatatgttgtgttttacagtattatgtagtctgtttttttatgcaaaatctaatttaatatattgatatttatatcattttacgtttctcgttcagctttttgtacaaacttgtgatatcccgcggccat
gctagagtccgcaaaaatcaccagtctctctctacaaatctatctctctatttttctccagaataatgtgtgagtagttcccagataagggaattagggttcttatagggtttcgctcatgtg
ttgagcatataagaaacccttagtatgtatttgtatttgtaaaatacttctatcaataaaatttctaattcctaaaaccaaaatccagtgacctgcaggcatgcgacgtcgggcccaagctt
agcttgagcttggatcagattgtcgtttcccgccttcagtttaaactatcagtgtttgacaggatatattggcgggtaaacctaagagaaaagagcgtttattagaataacggatatttaaa
agggcgtgaaaaggtttatccgttcgtccatttgtatgtgcatgccaaccacagggttcccctcgggatcaaagtactttgatccaacccctccgctgctatagtgcagtcggcttctgac
gttcagtgcagccgtcttctgaaaacgacatgtcgcacaagtcctaagttacgcgacaggctgccgccctgccctttttcctggcgttttcttgtcgcgtgttttagtcgcataaagtagaat
acttgcgactagaaccggagacattacgccatgaacaagagcgccgccgctggcctgctgggctatgcccgcgtcagcaccgacgaccaggacttgaccaaccaacgggccg
aactgcacgcggccggctgcaccaagctgttttccgagaagatcaccggcaccaggcgcgaccgcccggagctggccaggatgcttgaccaccacgccctggcgacgttgtg
acagtgaccaggctagaccgcctggcccgcagcacccgcgacctactggacattgccgagcgcatccaggaggccggcgcgggcctgcgtagcctggcagagccgtgggcc
gacaccaccacgccggccggccgcatggtgttgaccgtgttcgccggcattgccgagttcgagcgttccctaatcatcgaccgcacccggagcgggcgcgaggccgccaaggc
ccgaggcgtgaagtttggcccccgccctaccctcaccccggcacagatcgcgcacgcccgcgagctgatcgaccaggaaggccgcaccgtgaaagaggcggctgcactgctt
ggcgtgcatcgctcgaccctgtaccgcgcacttgagcgcagcgaggaagtgacgcccaccgaggccaggcggcgcggtgccttccgtgaggacgcattgaccgaggccgacg
ccctggcggccgccgagaatgaacgccaagaggaacaagcatgaaaccgcaccaggacggccaggacgaaccgtttttcattaccgaagagatcgaggcggagatgatcgc
ggccgggtacgtgttcgagccgcccgcgcacgtctcaaccgtgcggctgcatgaaatcctggccggtttgtctgatgccaagctggcggcctggccggccagcttggccgctgaag
aaaccgagcgccgccgtctaaaaaggtgatgtgtatttgagtaaaacagcttgcgtcatgcggtcgctgcgtatatgatgcgatgagtaaataaacaaatacgcaaggggaacgc
atgaaggttatcgctgtacttaaccagaaaggcgggtcaggcaagacgaccatcgcaacccatctagcccgcgccctgcaactcgccggggccgatgttctgttagtcgattccga
tccccagggcagtgcccgcgattgggcggccgtgcgggaagatcaaccgctaaccgttgtcggcatcgaccgcccgacgattgaccgcgacgtgaaggccatcggccggcgc
gacttcgtagtgatcgacggagcgccccaggcggcggacttggctgtgtccgcgatcaaggcagccgacttcgtgctgattccggtgcagccaagcccttacgacatatgggccac
cgccgacctggtggagctggttaagcagcgcattgaggtcacggatggaaggctacaagcggccttttgtcgtgtcgcgggcgatcaaaggcacgcgcatcggcggtgaggttgc
cgaggcgctggccgggtacgagctgcccattcttgagtcccgtatcacgcagcgcgtgagctacccaggcactgccgccgccggcacaaccgttcttgaatcagaacccgaggg
cgacgctgcccgcgaggtccaggcgctggccgctgaaattaaatcaaaactcatttgagttaatgaggtaaagagaaaatgagcaaaagcacaaacacgctaagtgccggccg
tccgagcgcacgcagcagcaaggctgcaacgttggccagcctggcagacacgccagccatgaagcgggtcaactttcagttgccggcggaggatcacaccaagctgaagatg
tacgcggtacgccaaggcaagaccattaccgagctgctatctgaatacatcgcgcagctaccagagtaaatgagcaaatgaataaatgagtagatgaattttagcggctaaagga
ggcggcatggaaaatcaagaacaaccaggcaccgacgccgtggaatgccccatgtgtggaggaacgggcggttggccaggcgtaagcggctgggttgtctgccggccctgca
atggcactggaaccccaagcccgaggaatcggcgtgacggtcgcaaaccatccggcccggtacaaatcggcgcggcgctgggtgatgacctggtggagaagttgaaggccg
cgcaggccgcccagcggcaacgcatcgaggcagaagcacgccccggtgaatcgtggcaagcggccgctgatcgaatccgcaaagaatcccggcaaccgccggcagccgg
tgcgccgtcgattaggaagccgcccaagggcgacgagcaaccagatttttcgttccgatgctctatgacgtgggcacccgcgatagtcgcagcatcatggacgtggccgttttccgt
ctgtcgaagcgtgaccgacgagctggcgaggtgatccgctacgagcttccagacgggcacgtagaggtttccgcagggccggccggcatggccagtgtgtgggattacgacctg
gtactgatggcggtttcccatctaaccgaatccatgaaccgataccgggaagggaagggagacaagcccggccgcgtgttccgtccacacgttgcggacgtactcaagttctgcc
ggcgagccgatggcggaaagcagaaagacgacctggtagaaacctgcattcggttaaacaccacgcacgttgccatgcagcgtacgaagaaggccaagaacggccgcctg
gtgacggtatccgagggtgaagccttgattagccgctacaagatcgtaaagagcgaaaccgggcggccggagtacatcgagatcgagctagctgattggatgtaccgcgagatc
acagaaggcaagaacccggacgtgctgacggttcaccccgattacttttgatcgatcccggcatcggccgttttctctaccgcctggcacgccgcgccgcaggcaaggcagaag
ccagatggttgttcaagacgatctacgaacgcagtggcagcgccggagagttcaagaagttctgtttcaccgtgcgcaagctgatcgggtcaaatgacctgccggagtacgatttga
aggaggaggcggggcaggctggcccgatcctagtcatgcgctaccgcaacctgatcgagggcgaagcatccgccggttcctaatgtacggagcagatgctagggcaaattgcc
ctagcaggggaaaaaggtcgaaaaggtctctttcctgtggatagcacgtacattgggaacccaaagccgtacattgggaaccggaacccgtacattgggaacccaaagccgtac
attgggaaccggtcacacatgtaagtgactgatataaaagagaaaaaaggcgatttttccgcctaaaactctttaaaacttattaaaactcttaaaacccgcctggcctgtgcataact
gtctggccagcgcacagccgaagagctgcaaaaagcgcctacccttcggtcgctgcgctccctacgccccgccgcttcgcgtcggcctatcgcggccgctggccgctcaaaaatg
gctggcctacggccaggcaatctaccagggcgcggacaagccgcgccgtcgccactcgaccgccggcgcccacatcaaggcaccctgcctcgcgcgtttcggtgatgacggtg
aaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggc
gcagccatgacccagtcacgtagcgatagcggagtgtatactggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacagatgc
gtaaggagaaaataccgcatcaggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttat
ccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccc
tgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgacc
ctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcac
gaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcag
agcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaaga
gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacg
gggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgcatgatatatctcccaatttgtgtagggcttattatgcacgcttaaaaataataaaagcagacttgacct
gatagtttggctgtgagcaattatgtgcttagtgcatctaatcgcttgagttaacgccggcgaagcggcgtcggcttgaacgaatttctagctagacattatttgccgactaccttggtgat
ctcgcctttcacgtagtggacaaattcttccaactgatctgcgcgcgaggccaagcgatcttcttcttgtccaagataagcctgtctagcttcaagtatgacgggctgatactgggccgg
caggcgctccattgcccagtcggcagcgacatccttcggcgcgattttgccggttactgcgctgtaccaaatgcgggacaacgtaagcactacatttcgctcatcgccagcccagtc
gggcggcgagttccatagcgttaaggtttcatttagcgcctcaaatagatcctgttcaggaaccggatcaaagagttcctccgccgctggacctaccaaggcaacgctatgttctcttg
cttttgtcagcaagatagccagatcaatgtcgatcgtggctggctcgaagatacctgcaagaatgtcattgcgctgccattctccaaaattgcagttcgcgcttagctggataacgccacg
gaatgatgtcgtcgtgcacaacaatggtgacttctacagcgcggagaatctcgctctctccaggggaagccgaagtttccaaaaggtcgttgatcaaagctcgccgcgttgtttcatc
aagccttacggtcaccgtaaccagcaaatcaatatcactgtgtggcttcaggccgccatccactgcggagccgtacaaatgtacggccagcaacgtcggttcgagatggcgctcg
atgacgccaactacctctgatagttgagtcgatacttcggcgatcaccgcttcccccatgatgtttaactttgttttagggcgactgccctgctgcgtaacatcgttgctgctccataacatc
aaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatagactgtaccccaaaaaaacatgtcataacaagaagccatgaaaaccgccactgcgccgttacc
accgctgcgttcggtcaaggttctggaccagttgcgtgacggcagttacgctacttgcattacagcttacgaaccgaacgaggcttatgtccactgggttcgtgcccgaattgatcaca
ggcagcaacgctctgtcatcgttacaatcaacatgctaccctccgcgagatcatccgtgtttcaaacccggcagcttagttgccgttcttccgaatagcatcggtaacatgagcaaagt
ctgccgccttacaacggctctcccgctgacgccgtcccggactgatgggctgcctgtatcgagtggtgattttgtgccgagctgccggtcggggagctgttggctggctggtggcagg
atatattgtggtgtaaacaaattgacgcttagacaacttaataacacattgcggacgtttttaatgtactgaattaacgccgaattgaattatcagcttgcatgccggtcgatctagtaacat
atagatgacaccgcgcgcgataatttatcctagtttgcgcgctatattttgttttctatcgcgtattaaatgtataattgcgggactctaatcataaaaacccatctcataaataacgtcatgc
attacatgttaattattacatgcttaacgtaattcaacagaaattatatgataatcatcgcaagaccggcaacaggattcaatcttaagaaacttattgccaaatgtttgaacgatctgctt
gactctaggggtcatcagatttcggtgacgggcaggaccggacggggcggcaccggcaggctgaagtccagctgccagaaacccacgtcatgccagttcccgtgcttgaagcc
ggccgcccgcagcatgccgcgggggggcatatccgagcgcctcgtgcatgcgcacgctcgggtcgttgggcagcccgatgacagcgaccacgctcttgaagccctgtgcctccag

ggacttcagcaggtgggtgtagagcgtggagcccagtcccgtccgctggtggcgggggagacgtacacggtcgactcggccgtccagtcgtaggcgttgcgtgccttccaggga
cccgcgtaggcgatgccggcgacctcgccgtccacctcggcgacgagccagggatagcgctcccgcagacggacgaggtcgtccgtccactcctgcggttcctgcggctcggta
cggaagttgaccgtgcttgtctcgatgtagtggttgacgatggtgcagaccgccggcatgtccgcctcggtggcacggcggatgtcggccgggcgtcgttctgggctcatggtagatc
ccctcgatcgagttgagagtgaatatgagactctaattggataccgagggaatttatggaacgtcagtggagcatttttgacaagaaatatttgctagctgatagtgaccttaggcga
cttttgaacgcgcaataatggtttctgacgtatgtgcttagctcattaaactccagaaacccgcggctcagtggctccttcaacgttgcggttctgtcagttccaaacgtaaaacggcttgt
cccgcgtcatcggcgggggtcataacgtgactcccttaattctcatgtatgataattcgagct

EnHSP70 multiple sequence alignment

```
E.nindensis      MAA--------------------SQSQTQGDAPAIGIYLGTTNSCVAVCRRRRDRVEII
A.comosus        MA-------------------------GKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
E.ventricosum    MA-------------------------GKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
A.thaliana       MA-------------------------GKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
G.barbadense     MA-------------------------GKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
T.turgidum       MA-----------------------PTKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
C.nucifera       MA-----------------------KGEGPAIGIDLGTTYSCVGVW--QHDRVEII
P.hallii         MA--------------------SKNGGKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
Z.officinale     MA-----------------------GGKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
D.carota         MA-------------------------GKGEGAAIGIDLGTTYSCVGVW--QHDRVEII
H.annuus         MA-------------------------GKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
T.aestivum       MA-------------------------GKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
C.annuum         MA-------------------------GKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
Z.mays           MA-----------------------KGEGPAIGIDLGTTYSCVGVW--QHDRVEII
C.sativa         MA-----------------------TKSD-KAIGIDLGTTYSCVGVW--LNDRVEII
G.max            MA-----------------------TK-EGKAIGIDLGTTYSCVGVW--QNDRVEII
V.angularis      MA-----------------------SKVEGKSIGIDLGTTYSCVGVW--QNDRVEII
S.intermedia     MA-----------------------GKGEGPAIGIDLGTTYSCVGVW--QHDRVEII
J.curcas         MA-----------------------GKGKVPAIGIDLGTTYSCVGVW--QHDRVEII
L.angustifolius  MA-----------------------KNDESIAIGIDLGTTYSCVAVWEKQHNRVEII
N.tabacum        MA---------------------------PAVGIDLGTTYSCVGIF-REIGKCDII
D.exilis         MA-----------------------ASKDDGPAVGIDLGTTYSCVAVW--RHDRGEVI
E.curvula        MAA--------------------ASSKADGPAVGIDLGTTYSCVAVW--QGDRGEVI
O.sativa         MDRVRGCAFLLGVLLAGSLFAFSVAKEETKKLGTVIGIDLGTTYSCVGVY--KNGHVEII
                 *                              :** **** ***.:     .: ::*

E.nindensis      ANDRGDRLTPSCVAFT-DCEKLVGEVAVDQVSSNPANTIY--------------------
A.comosus        ANDQGNRTTPSYVAFT-DSERLIGDAAKNQVAMNPTNTVF--------------------
E.ventricosum    ANDQGNRTTPSYVAFT-DSERLIGDAAKNQVAMNPINTVF--------------------
A.thaliana       ANDQGNRTTPSYVAFT-DSERLIGDAAKNQVAMNPINTVFDAKRLIGRRFTDSSVQSDIK
G.barbadense     ANDQGNRTTPSYVGFT-DTERLIGDAAKNQVAMNPLNTVF--------------------
T.turgidum       ANDQGNRTTPSYVAFT-DSERLIGDAAKNQVAMNPINTVLSAYEPCG-------------
C.nucifera       ANDQGNRTTPSYVAFT-DTERLIGDAAKNQVAMNPTNTVF--------------------
P.hallii         ANDQGNRTTPSYVAFT-DTERLIGDAAKNQVAMNPTNTVFDAKRLIGRRFSDPSVQADMK
Z.officinale     ANDQGNRTTPSYVAFT-DTERLIGDAAKNQVAMNPTNTVFDAKRLIGRRYSDPSVQGDIK
D.carota         ANDQGNRTTPSYVAFT-DTERLIGDAAKNQVAMNPINTVF--------------------
H.annuus         ANDQGNRTTPSYVAFT-DGERLIGDAAKNQVAMNPMNTVFDAKRLIGRRFSDTSVQDDMK
T.aestivum       ANDQGNRTTPSYVAFT-DSERLIGDAAKNQVAMNPVNTVFDAKRLIGRRVSDPSVQSDMK
C.annuum         ANDQGNRTTPSYVGFT-DSERLIGDAAKNQVAMNPINTVFDAKRLIGRRFSDASVKEDMK
Z.mays           ANDQGNRTTPSYVAFT-DTERLIGDAAKNQVAMNPTNTVFDAKRLIGRRFSSPAVQSSMK
C.sativa         ANDQGNRTTPSYVAFT-DTERLIGDAAKNQ-------------------------------
G.max            PNDQGNRTTPSYVAFT-DTERLIGDAAKNQVAMNPQNTVFDAKRLIGRRFSDSSVQNDMK
V.angularis      ANDQGNRTTPSYVAFT-ETERLIGDAAKNQ-------------------------------
S.intermedia     ANDQGNRTTPSYVAFT-DTERLIGDAAKNQVAMNPTNTVFDAKRLIGRRISDPSVQSDIK
J.curcas         ANDQGNRTTPSHVAFT-DSERLIGDAAKNQVGMNPVNTVF--------------------
L.angustifolius  HNDHGNKITPSFVAFT-DDQRLIGDAAKNQAAINPTNTIF--------------------
N.tabacum        ANDQGNRTTPSFVGFT-DTERLIGDAAKNQVAMNPQNTVFDAKRLIGRKFADPEVQADMK
D.exilis         PNDQGNRLTPSCVAFTDDAERLVGDAAVNQAALNPTNTIF--------------------
E.curvula        PNDQGNRLTPSCVAFT-DTECFVGDAAVNQAALNPTNTIFEVKRLVGRRFSDKTVKEDIK
O.sativa         ANDQGNRITPSWVAFT-DSERLIGEAAKNQAAVNPERTIFDVKRLIGRKFEDKEVQRDMK
                  **:*:: *** *.** : : ::*:.* :*

E.nindensis      ----------------------GESSSFSAR----------------------------
A.comosus        ------------------------------------------------VYLFIVFEQRLDKH
E.ventricosum    ----------GECPSSMMIYA----------HLLAMLLLK---------GLTFGDSA---
A.thaliana       LWPFTLKSGPAEKPMIVVNYK-GEDKEFSAEEISSMILIKMREIAEAYLGTTIKNAVVTV
G.barbadense     ----------------------------------------MREIAEAYLGFTVKNAVVTV
T.turgidum       ----------------------------------------REIAEAYLGVTIKNAVVTV
C.nucifera       ------------------------------------------------AYLGSTIKNAVVTV
P.hallii         MWPFKVVPGPADKPMIVVTYK-GEEKKFSAEEISSMVLTKMKEIAEAYLSTTIKNAVITV
Z.officinale     LWPFKVIAGPGDKPMIVVQYK-GEEKQFSAEEISSMVLTKMKEIAEAYLGTTVKNAVVTV
D.carota         ------------------------------------------------ETFLGTVVKNAVVTV
H.annuus         LWPFKVSSGPAEKPMIGVSYK-GEDKQFSAEEISSMVLTKMREIAEAFVGSTVKNAVVTV
T.aestivum       LWPFKVIAGLGDKPMI-VQYK-GEEKEFAAEEISSMVLIKMREIAEAYLGSTIKNVVVTV
C.annuum         LWPFKVVAGPGDKPMIVVNYK-GEEKQFAAEEISSMVLTKMKEIAEAFLGSTVKNAVVTV
Z.mays           LWPSRHL-GLGDKPMIVFNYK-GEEKQFAAEEISSMVLIKMKEIAEAYLGSTIKNAVVTV
C.sativa         ----------------------------------------MREVSEAFLGHSVKNAVVTV
G.max            LWPFKVGGSPCDKPMIVVNYK-GEEKKFSAEEISSMVLVKMREVAEAFLGHAVKNAVVTV
V.angularis      ------------------------------------------------SVKNAVITV
S.intermedia     LWPFKVVPGPGEKPMIAVQYK-GEEKQFAAEEISSMVLVKMREIAE--------------
J.curcas         ----------------------------------------------------VKNAVVTV
L.angustifolius  ----------------------------------------ESYLESPVKNAVVTV
N.tabacum        HFPFKIV-DKGGKPNIEVEFK-GETKTFTPEEISAMILTKMRETAESYLGETVTNAVVTV
D.exilis         ------------------------------------------------GENYTRAIHHL
E.curvula        LWPFKVVPGCEDRPMIMVQYK-GEEKQFSAEEISSMVLAKLRKTAEVYLGTTVTNVVITV
O.sativa         LVPYKIV-NKDGKPYIQVKIKDGENKVFSPEEVSAMILGKMKETAEAYLGKKINDAVVTV
```

```
E.nindensis      -SHLLCLRRQATVDAAKIAGLNVLHIINEATSAAIAYGLDAMPAGNEGRTVLVFDLGGGA
A.comosus        IMVGNTSERQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKAESVGEKNVLIFDLGGGT
E.ventricosum    ------SDRQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKASSAGEKNVLIFDLGGGT
A.thaliana       PAYFNDSQRQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKATSVGEKNVLIFDLGGGT
G.barbadense     PAYFNDSQRQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKATSVGEKNVLIFDLGGGT
T.turgidum       PAYFNDSQRQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKATSVGEKNVLIFDLGGGT
C.nucifera       PAYFNDSQRQATKDAGVISGLNVMRIINEPTAAAIAYGLDKKAGSSAEKNVLIFDLGGGT
P.hallii         PAYFNDSQRQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKAASTGEKSVLIFDLGGGT
Z.officinale     PAYFNDSQRQATKDAGVIAGLNVLRIINEPTAAAIAYGLDKKGSSSGEKNVLIFDLGGGT
D.carota         PAYFNDSQRQATKDAGVISGLNVMRIINEPTAAAIAYGLDKKATSVGEKNVLIFDLGGGT
H.annuus         PAYFNDSQRQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKASSAGEKNVLIFDLGGGT
T.aestivum       PAYFNDSQRQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKATSVGEKNVLIFDLGGGT
C.annuum         PAYFNDSQRQATKDAGVISGLNVLRIINEPTAAAIAYGLDKKATSVGEKNVLIFDLGGGT
Z.mays           PAYFNDSQRQATKDAGVIAGLNVMRIINEPTAAAIAYGLDKKATSSGEKNVLIFDLGGGT
C.sativa         PAYFNDSQRQATKDAGAISGLNVLRIINEPTAAAIAYGLDKKASRVGEQNVLIFDLGGGT
G.max            PAYFNDSQRQATKDAGAISGLNVLRIINEPTAAAIAYGLDKKASRKGEQNVLIFDLGGGT
V.angularis      PAYFNDSQRQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKKGSKSGEKNVLIFDLGGGT
S.intermedia     ----------PTKDAGVIAGLNVLRIINEPTAAAIAYGLDKKGGGAGEKNVLIFDLGGGT
J.curcas         PAYFNDFQRQATKDAGVIAGLNVMRIINEPTAAAITYGLDKMATSDDEKNVLIFDLGGGT
L.angustifolius PAYFNDSQRQATKDAATIAGLNVMRITNEPTAAALAYGLDKRANCVDERNVFVFDLGGGT
N.tabacum        PAYFNDSQRQATKDAGLIAGLNVLRIINEPTAAAIAYGLDKKV--EGERNVLIFDLGGGT
D.exilis         VCLYSNSQRQATIDAGTIAGLNVLRIINEPTAAALAYGLEKMPVTNEGRTVLVFDLGGGT
E.curvula        PVYFSNSQRQATIDAGAIAGLNVMRIINEPTAAAIAYGLEKMPVSNKGRTVLVFDLGGGT
O.sativa         PAYFNDAQRQATKDAGVIAGLNVARIINEPTAAAIAYGLDKKG---GEKNILVFDLGGGT
                           .* **. *:**** :* **.*:**::***:        :.:::******:

E.nindensis      LDVSLLEVDRGEDIDMALFEVKAIAGDTHLGGVDFNNEMVEYLLLELIRKHKKATIRNNP
A.comosus        FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKH-KKDITGNP
E.ventricosum    FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKH-KKDISGNP
A.thaliana       FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKN-KKDISGNP
G.barbadense     FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKN-KKDISGNP
T.turgidum       FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKH-KKDISGNP
C.nucifera       FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKH-KKDISGNP
P.hallii         FDVSILTI------EEGIFEVKATAGDTHLGGEDFDNRLVNHFVQEFKRKH-KKDITGNP
Z.officinale     FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKH-KKDISGNP
D.carota         FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKN-KKDISGNP
H.annuus         FDVSILTI------EEGIFEVKSTAGDTHLGGEDFDNRMVNHFVQEFKRKH-KKEINGSP
T.aestivum       FDVSLLTI------EEGIFEVKSTAGDTHLGGEDFDNRMVNHFVQEFKRKN-KKDISGNP
C.annuum         FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKH-KKDISGNP
Z.mays           FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVQEFKRKN-KKDISGNP
C.sativa         FDVSILTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVSEFRRKN-KKDISGNA
G.max            FDVSILTI------EEGIFEVKATAGDTHLGGEDFDNRMVNHFVSEFKRKN-KKDISGNA
V.angularis      FDVSLLTI------EEGIFEVKATAGDTHLGGEDFDNRLVNHFVAEFKRKH-KKDISSSA
S.intermedia     FDVSLLTI------EEGIFEVKSTAGDTHLGGEDFDNRLVNHFVQEFKRKH-KKDISGNP
J.curcas         FDVSLLTI------KDGIFEVKAMAGDTHLGGEDIDNRVVNHLVQEFKRKN-KKDISGNP
L.angustifolius FDVSLLTI------KGKIFEVKATAGDTHLGGLDFDNRMVNYFVQEFKRKN-NVNISGNP
N.tabacum        FDVSLLTI------EEGIFEVKSTAGDTHLGGEDFDNRLVNHFVNEFKRKH-KKDLSTNV
D.exilis         FDVSILNIDPGKNMDKGLFEVKAIAGDTHLGGADFDNEMVKYALQEFKRKHGNMDINSNQ
E.curvula        LDVSLLNIDPGIDIDMGLFEVKAVSGDTHLGGADFDNEMVKFFLREFIRKHEKADIRSNQ
O.sativa         FDVSILTI------DNGVFEVLATNGDTHLGGEDFDQRIMEYFIKLIKKKY-SKDISKDN
                 :***:* :      .  :*** :  ******* *:::.:::. :  : :*  .  :  .

E.nindensis       RALRRLRLACERVKKSLALLTETTIEECSLHDGIDFCITITRSRFEELNEAHFSKCMEAV
A.comosus         RALRRLRTSCERAKRTLSSTAQTTIEIDSLYEGVDFYSTITRARFEELNMDLFRKCMEPV
E.ventricosum     RALRRLRTACERAKRTLSSTAQTTIEIDSLFEGVDFYSPITRARFEELNMDLFRKCMEPV
A.thaliana        RALRRLRTACERAKRTLSSTAQTTIEIDSLFDGIDFYAPITRARFEELNIDLFRKCMEPV
G.barbadense      RALRRLRTSCERAKRTLSSTAQTTIEIDSLYEGIDFYSTITRARFEELNMDLFRKCMEPV
T.turgidum        RSLRRLRTSCERAKRTLSSTAQTTIEIDSLFEGVDFYSTITRARFEELNMDLFRKCMEPV
C.nucifera        RALRRLRTACERAKRTLSSTAQTTIEIDSLYEGIDFYSTITRARFEELNMDLFRKCMEPV
P.hallii          RALRRLRTACERAKRTLSSTAQTTIEIDSLYEGIDFYATITRARFEELNMDLFRKCMEPV
Z.officinale      RALRRLRTACERAKRTLSSTAQTTIEIDSLFEGIDFYSTITRARFEELNMDLFKKCMEPV
D.carota          RALRRLRTSCERAKRTLSSTAQTTIEIDSLFEGIDFYSTITRARFEEMNMDLFRKCMEPV
H.annuus          RALRRLRTACERAKRTLSSTAQTTIEIDSLYEGVDFYTTMTRARFEELNMDLFRKCMEPV
T.aestivum        RVLRGLRTACERAKRTLSSIAQTTIEIDSLYEGIDFYTTITRARFEELNMDLFRKCMEPV
C.annuum          RALRRLRTACERAKRTLSSTAQTTIEIDSLYEGVDFYSTITRARFEELNMDLFRKCMEPV
Z.mays            RALRRLRTACERAKRTLSSTAQTTIEIDSLFEGIDFTPRSSRARFEELNMDLFRKCMEPV
C.sativa          RSLRRLRTACERAKRTLSSTTQTTIEIDSLYEGIDFYATITRARFEELNMDLFKRCMEPV
G.max             RALRRLRTACERAKRTLSSTAQTTIEIDSLYEGIDFYATITRARFEEMNMDLFRKCMEPV
V.angularis       RALRRLRTACERAKRTLSSTSQTTIEIDSLYEGIDFYSTITRARFEELNMDLFRKCMEPV
S.intermedia      RALRRLRTACERAKRTLSSTAQTTIEIDSLSEGIDFFSTITRARFEELNMDLFKRCMEPV
J.curcas          RALGRLRTACERAKRALSFTAVATIEIDFLYEGTDFYSTMTRAKFEELNMDLFRKCMEHV
L.angustifolius RALRRLRSACERAKRVLSFASNTTIEVDSLLQGIDFCSSVTRAKFEQLNIDLFKKCMEPV
N.tabacum         RALRRLRTACERAKRTLSSSAQTSIEIDSLFEGIDFYTSITRARFEELCQDLFRSTIQPV
D.exilis          KALRRLRTACERAKRMLSFTAQTTIEVDSLHDGIDFCTIITRSRFEELNKDLFDKCMKAV
E.curvula         KALRRLRTACERAKRMLSSTSQTTIEVDSLHDGIDFCTTITRSRFEELNKDLFSKCMKAV
O.sativa          RALGKLRREAERAKRALSNQHQVRVEIESLFDGTDFSEPLTRARFEELNNDLFRKTMGPV
                  : *  **  .**.*: *:    . :*   * :* **    :*::**::    *   :  *
```

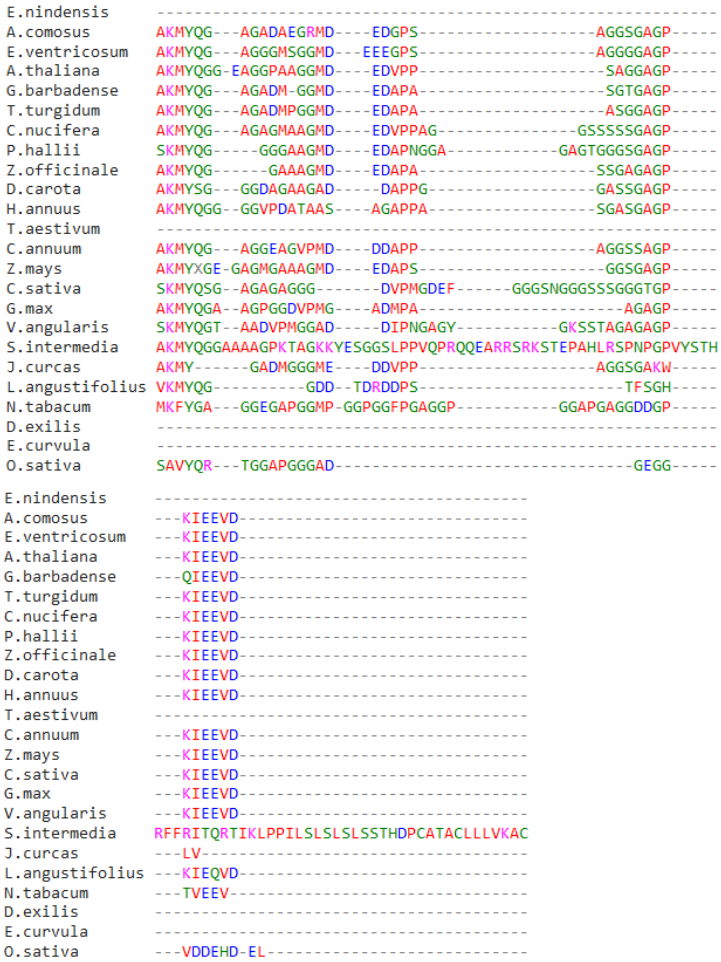
```
E.nindensis      EKCLHDVKMDKSAVQDVVLVGGSTRIPKVQSMLREFFDGKRLCQTINPDEAVAYGAAIHA
A.comosus        EKCLRDAKMDKSTVHDVVLVGGSTRIPRVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
E.ventricosum    EKCLRDAKMDKSSIHDVVLVGGSTRIPRVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
A.thaliana       EKCLRDAKMDKNSIDDVVLVGGSTRIPKVQQLLVDFFNGKELCKSINPDEAVAYGAAVQA
G.barbadense     EKCLRDAKMDKSSVHDVVLVGGSTRIPKVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
T.turgidum       EKCLRDAKMDKSTIHDVVLVGGSTRIPRVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
C.nucifera       EKCLRDAKMDKSSIHDVVLVGGSTRIPKVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
P.hallii         EKCLRDAKMDKSQIHDVVLVGGSTRIPKVQQLLQDFFNGKELCKSINPDEAVAYGASVQA
Z.officinale     EKCMRDAKMDKSSVDDVVLVGGSTRIPKVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
D.carota         EKCLRDAKMDKSTVHDVVLVGGSTRIPKVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
H.annuus         EKCLRDAKMDKKSVHEIVLVGGSTRIPKVQNLLQDFFNGKELCKSINPDEAVAYGAAVQA
T.aestivum       EKCLRDAKMDKSTVHDVVLVGGSTRIPRVQQLLQDFFNGKELCKNINPDEAVAYGAAVQA
C.annuum         EKCLRDAKMDKSSVHDVVLVGGSTRIPKVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
Z.mays           EKCLRDAKMDKSSVHDVVLVGGSTRIPKVQQ-LQDFFNGKELCKSINPDEAVAYGAAVQA
C.sativa         EKCLRDAKIDKSQIHEVVLVGGSTRIPKVQQLLQDLFNGKELCKSINPDEAVAYGAAVQA
G.max            EKCLRDAKIDKSQVHEVVLVGGSTRIPKVHQLLQDFFNGKELCKSINPDEAVAYGAAVQA
V.angularis      EKCLRDSKIDKSHVHDVVLVGGSTRIPKVQQLLQDFFNGKELCKSINPDEAVAYGAAVQA
S.intermedia     EKCLRDAKMDKSSIHDVVLVGGSTRIPKVQQLLQEFFNGKELCKSINPDEAVAYGAAVQA
J.curcas         EKCLGDAKMDNSSVHDVVLVGGSTRIPKVQQLLQDFFNGKELCKSINPDEAVAFGAAVQA
L.angustifolius  EQCLRFAKMEKSSVHDVVLVGGSTRIPKVQKLLQDFFNGKELCNSINPDEAVAYGAAVQA
N.tabacum        DRVLTDAKIDKSLVHEIVLVGGSTRILRVQKLITDYFNGKEPNKSINPDEAVAYGAAVQA
D.exilis         EKCLQDAKMDKSSIHDVVLVGGSTRIPQVQSMLRDFFQGKELCRNINPDEAVAYGAAIQA
E.curvula        EKCLHDAKMDKSSVDDVVLVGGSTRIPKVQSMLQKFFDGKELRRNINPDEAVAYGAAIQA
O.sativa         KKAMDDAGLEKSQIHEIVLVGGSTRIPKVQQLLRDYFEGKEPNKGVNPDEAVAYGAAVQG
                 .: :    :::. :.::********** :*:. : . *:**.  . :*******:**:::.

E.nindensis      ANLSDASGKK-VWDLLLMDVTPFSLGVEADW-------GDMTVLIPRNTPVPTRKEKVFS
A.comosus        AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
E.ventricosum    AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIQRNTTIPTKKEQVFS
A.thaliana       AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIQRNTTIPTKKEQVFS
G.barbadense     AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
T.turgidum       AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLITRNTTIPTKKEQVFS
C.nucifera       AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
P.hallii         AILSGEGNQK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
Z.officinale     AILSGEGNEK-VHDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
D.carota         AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIQRNTTIPTKKEQVFS
H.annuus         AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
T.aestivum       AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
C.annuum         AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
Z.mays           AILSGEGNER--SDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQVFS
C.sativa         AILSGEGNEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQIFS
G.max            AILSGQGDEK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQIFS
V.angularis      AILSGEGNEK-VQDLLLLDVTPLSLGIETAG-------GVMAVLIPRNTTIPTKKEQIFS
S.intermedia     AILTGEGNQK-VQDLLLLDVTPLSLGLETAG-------GVMTVLIPRNTTIPTKKEQIFS
J.curcas         AISSGEGNEK-VQDLVLLDVTPLSLGLKIIG-------GVMSVLIPRNTTIPAKKERFFT
L.angustifolius AMLSGEGNEM-VQDLCLLDVTPLSLGLETSG-------GLMTVLIPRNTPIPTKRELVFS
N.tabacum        AILSGDTSSKATNEILLLDVAPLSLGIETAG-------GMMTKLIPRNTTIPTKKSEVFS
D.exilis         SILCGQADDGRLLDMLLRDVTPLSLGVEIRDD------NTMSVVIPRNTAIPTKMVKGFT
E.curvula        SILNGGNGDGRLVDMLLRDVTPLSLGIKTATGFVVTIFGVMSVLIPRNTAIPTRKEKVFT
O.sativa         SILSGEGGDE-TKDILLLDVAPLTLGIETVG-------GVMTKLIPRNTVIPTKKSQVFT
                 :   .  ..    :: * **:*::**::          . *: :* *** :*::    *:

E.nindensis      TYHDNQAEVLIQVYEGEGWSAEENNLLGKFELAGIHPAPRGIPEIIVCFYVDANGILTVS
A.comosus        TYSDNQPGVLIQVYEGERTRTRDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
E.ventricosum    TYSDNQPGVLIQVYEGERTRTRDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
A.thaliana       TYSDNQPGVLIQVYEGERARTKDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
G.barbadense     TYSDNQPGVLIQVYEGERARTRDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
T.turgidum       TYSDNQPGVLIQVFEGERTRTRDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
C.nucifera       TYSDNQPGVLIQVYEGERTRTRDNNLLGKFELSGIPPAPRGVPQINVCFDMDANGILNVT
P.hallii         TYSDNQPGVLIQVYEGERTRTKDNNLLGKFELTGIPPAPRGVPQINVTFDIDANGILNVS
Z.officinale     TYSDNQPGVLIQVYEGERARTKDNNLLGKFELSGIPPAPRGVPQINVCFDLDANGILNVS
D.carota         TYSDNQPGVLIQVFEGERTRTRDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
H.annuus         TYSDNQPGVLIQVYEGERTRTRDNNLLGKFELSGIPPAPRGVPQINVCFDIDANGILNVS
T.aestivum       TYFDNQPGVLIQVYEGERTRTRDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
C.annuum         TYSDNQPGVLIQVFEGERARTRDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
Z.mays           TYSDNQPGVLIQVYEGERARTKDNNLLGKFELSGIPPAPRGVPQITVTFDIDVNNILNVS
C.sativa         TYSDNQPGVLIQVYEGERARTRDNNLLGKFELTGIPPAPRGVPQVNVVFDIDANGILNVS
G.max            TYSDNQPGVLIQVFEGERARTKDNNLLGKFELTGIPPAPRGVPQVNVCFDIDANGILNVS
V.angularis      TYADNQAGVLIQVYEGERARTKDNNLLGKFELAGIPPAPRGVPQINVCFDIDANGILNVS
S.intermedia     TYSDNQPGVLIQVFEGERTRTKDNNLLGKFELSGIPPAPRGVPQITVCFDIDANGILNVS
J.curcas         TVIDDQPKVLIQVYEGERSRTKDNNLLGKFELCGIQPAPRGVPEITVCFCIDANGILNVS
L.angustifolius TYSDYQSSVLIQVYEGERTRAKDNNLLGKFEFSGIQLVPRGVPQITVCFDIDANGILNVS
N.tabacum        TFSDNQPGVLIQVYEGERQRTKDNNLMGKFELTGIPPAPRGVPQIEVTFDLDANGIMNVS
D.exilis         TRYDNQTIVSFPVYEGESASTKNNNLLGEFELTGITPALKGVPKFDVTFDIDANGVMNVS
E.curvula        TLSDNQLGVPIRVYEGESAWTKDNNLLGEFTLEGIQLAPRGVPKINVTFDIDANGIMDVS
O.sativa         TYQDQQTTVSIQVFEGERSMTKDCRLLGKFDLSGIPAAPRGTPQIEVTFEVDANGILNVK
                 *  * *  * : *:***   :.: .*:*:* : **  . :* *:. * * :*.*.:: *.
```

```
E.nindensis      AEDKETGQTNQITIVTDDSCRLSKEEMERMMQESEKDRG--------------------
A.comosus        AEDKTTGQKNKITI-TNDKGRLSKDEIEKMVQDAEKYKSEDEEHKKKVESKNALENYAYN
E.ventricosum    AEDKTTGQKNKITI-TNDKGRLSKEEIERMVQEAEKYKSEDEDHKKKVEAKNALENYSYN
A.thaliana       AEDKTTGQKNKITI-TNDKGRLSKDEIEKMVQEAEKYKSEDEEHKKKVDAKNALENYAYN
G.barbadense     AEDKTTGQKNKITI-TNDKGRLSKEEIEKMVQEAEKYKSEDEEHKKKVDAKNALENYAYN
T.turgidum       AEDKTTGQKNKITI-TNDKGRLSKDDIEKMVQDAEKYKSEDEEHKKKVDAKNSLENYAYN
C.nucifera       AEDKTTGQKNKITI-TDDKGRLSKEEIEKMVQEAEKYKAEDEEHKKKIEAKNALENYAYN
P.hallii         AEDKTTGKKNKITI-TNDKGRLSKEEIERMVQEAEKYKAEDEEVKRKVEARNALENYAYN
Z.officinale     AEDKTTGQKNKITI-TNDKGRLSKAEIEKMVQEAEKYKAED--HKKKVEAKNALENYAYN
D.carota         AEDKTTGQKNKITI-TNDKGRLSKEEIEKMVLEAEKYKAEDEEHKKKVEAKNSLENYAYN
H.annuus         AEDKVGGKKNKITI-TNDKGRLSKEDIEKMVQEAEKYKAEDEEHKKKVEAKNSLENYAYN
T.aestivum       AEDKTTGQKNKITI-TNDKGRLSKDEIEKMVQEAEKYKSEDEEHKKKVESKNALENY---
C.annuum         AEDKTTGQKNKITI-TNDKGRLSKEEIEKMVQDAEKYKAEDEELKKKVGG----------
Z.mays           AEDKTTGQKNKITI-TNDKGRLSKEEIEKMVQEAEKYKAEDEEVKKKVDAKNALENYAYN
C.sativa         AEDKTAGVKNKITI-TNDKGRLSKEDIEKMVKDAERYKAEDEEVKRKVEAKNGLENYAYN
G.max            AEDKTAGVKNKITI-TNDKGRLSKEEIEKMVKDAERYKAEDEEVKKKVEAKNSLENYAYN
V.angularis      AEDKTAGVKNKITI-TNDKGRLSKEDIERMVEEAERYKAEDEDVKKKVEAKNSLENYAYN
S.intermedia     AEDKTTGQKNKITI-TNDKGRLSKEEIERMVREADKYKAEDEEIKKKVEAKNALENYAYN
J.curcas         AEDKTTGQKNNISI-TSDKGRLSKEEIEKMILEAQKYKAEDEEHKKNVEAMNALEYHAYN
L.angustifolius AENKTTSQKNMITI-TNDKGRLSKEEIERMIQEAEKYKFEDEEHKKKIEAKNALEKYAYN
N.tabacum        AVEKGTGKSNKIVI-TNDKGRLSKEEIERMLSDAEKYKEEDEAEGKRVAAKNGLESYAYS
D.exilis         AEDRDTGRKNHITI-TNHSGRLGKKEIEGMGQEAERYRR------KRIM-----------
E.curvula        AEDKTTGQKNKITV-TTKRGRLTKQEIERMALEVKAIAL--------------------
O.sativa         AEDKGTGKSEKITI-TNEKGRLSQEEIDRMVREAEEFAEEDKKVKERIDARNQLETYVYN
                 * ::  . .: * : * .  ** : ::: *  : .

E.nindensis      ------------------------------------------TWEDRKS----------
A.comosus        MRNTVKD-EKISSKLSAADKKKIEDAIEQAIQWLDGNQLAEADEFEDKMKELEGICNPII
E.ventricosum    MRNTVKD-DKIASKLPAADKKKIEDAVDQAIQWLDANQLAEVDEFEDKMKELENLCNPII
A.thaliana       MRNTIRD-EKIGEKLAGDDKKKIEDSIEAAIEWLEANQLAECDEFEDKMKELESICNPII
G.barbadense     MRNTVKD-EKIGAKLAAADKKKVEDAIDEAIQWLDNNQLAESDEFEDKMKELESICNPII
T.turgidum       MRNTIQD-EKIASKLPADDKKKIEDAVDAAIQWLDANQLGEVDEFEDKMKELEGLCNPII
C.nucifera       MRNTIRD-EKIAAKLPPADKKKIEDSIEAAINWLDGNQLAEADEFEDKMKELESICNPII
P.hallii         MRNTVRD-DKIASKLPAEDKKKIEDTIEDAIKWLDGNQLAEAEEFEDKMKELEGICNPII
Z.officinale     MRNTIKD-DKIAGKLPPADKKTIEDAVERAIRWLDGNQLAEAEEFEDKMKELEGICNPII
D.carota         MRNTIKD-EKIGSKLSPADKKTIEDSIDQAITWLDNNQLAEVEEFEDKMKELEGICNPII
H.annuus         MRNTISD-EKVSEKIAAGDKKKMEDAIEQTINWLDANQLGEVDEFEDKMKELEAICNPII
T.aestivum       -----------------GQKSQI-------------------------------------
C.annuum         --------QKFSGELRLQHEEH-----------NSNQLAEVDEFEDKMKELEGICNPII
Z.mays           MRNTIKD-DKIASKLPAEDKKKIEDAVDGAISWLDSNQLAEVEEFEDKMKELEGICNPII
C.sativa         MRNTVRD-EKFGEKLDKSDKEKIEKAVEEAIRWLEGNQLAEVDELEDKLKELEGVCNPII
G.max            MRNTIKD-EKIGGKLSPDEKQKIEKAVEDAIQWLEGNQMAEVDEFEDKQKELEGICNPII
V.angularis      MRNTVKD-EKFAGKLNPADKQKIEKAVDETIEWLDGNQLAEVEEFEDRLKDLEGLCNPII
S.intermedia     MRNTIRD-EQIASKLTPGEKKKIEDAVDGAIRWLDGNQLAEAEEFEHKMKELESVCNPIL
J.curcas         MRNTVKD-EKISSKLAAGDKKKIKDAVYTTIQWLGRNQLAKAEEFEDKMKALKSICNPII
L.angustifolius MTNTIED-YKISSKLSATDKAKIEDAIQKAFQWLDANQHAEVDEFDGKVKELDSICNPIM
N.tabacum        LRNTLSD-PKVEEKIEASDKETLTAEIDKVVQWLDDNQQATREEYEEHQKELEGKANPIM
D.exilis         ------------------------------------------EREAGQN----------
E.curvula        ------------------------------------------------------------
O.sativa         MKNTVGDKDKLADKLESEEKEKVEEALKEALEWLDENQTAEKEEYEEKLKEVEAVCNPII
```


E.nindensis
A.comosus
E.ventricosum
A.thaliana
G.barbadense
T.turgidum
C.nucifera
P.hallii
Z.officinale
D.carota
H.annuus
T.aestivum
C.annuum
Z.mays
C.sativa
G.max
V.angularis
S.intermedia
J.curcas
L.angustifolius
N.tabacum
D.exilis
E.curvula
O.sativa

Global proteome for *Eragrostis nindensis*

The following table lists the differentially abundant proteins as annotated from KEGG Koala blast using the methods as describes in Chapter 2 section 2.3.2. Values displayed here are the fold change values for each annotated protein per RWC relative to the fully hydrated control. All proteins were deemed as significant with a fold change of <2 and FDR of >0.05. A selection of the proteins below is used in Chapter 2 and Chapter 3 as part of exploring the proteomic signature of *E. nindensis* during dehydration and rehydration. Different KEGG metabolic categories are indicated by different colours. Heatmaps generated in Chapter 2 and Chapter 3 are summary heat maps from related and repeated terms shown below. Green boxes indicate DAPs which have a positive fold change relative to the fully hydrated control whereas red boxes are those which have a negative fold change relative to the fully hydrated control. Black areas indicate absence of the protein at that RWC.

Table 8-3 : Differentially abundant proteins as annotated by KEGG Koala Blast

KEGG Category	Protein	60 NST	40 NST	25 NST	AD NST	48hrs NST	W1 NST	60 ST	25 ST
Amino Acid Metabolism	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase							3,01	
	3-deoxy-7-phosphoheptulonate synthase	-6,11	-6,58						
	4-hydroxyphenylalaninenylpyruvate dioxygenase					-2,00	-2,89		
	5-methylthioribose kinase	2,98	2,58		4,31				
	5-methylthioribose kinase				4,31				
	5-oxoprolinase (ATP-hydrolysing)						-6,39		
	Alanine transaminase	-2,31			-2,23		-2,30	-2,67	-2,82
	Alanine transaminase	-2,45	-2,13		-2,29	-2,06	-2,44	-2,49	-3,31
	Alanine-glyoxylate transaminase / (R)-3-amino-2-methylpropionate-pyruvate transaminase								2,64
	Alanine-glyoxylate transaminase / serine-glyoxylate transaminase / serine-pyruvate transaminase	-2,45	-2,30					-2,58	-2,67
	Alanyl-tRNA synthetase	-2,90	-3,75		-2,50	-2,29	-2,17	-2,31	-3,03
	Aminoacylase								2,18
	Aminocyclopropanecarboxylate oxidase	4,74	3,11		2,97			2,28	2,14
	Aminopeptidase N						-2,13		
	Aralkylamine N-acetyltransferase	-4,17			-4,69			-4,59	-3,48
	Arginase		3,67						2,61
	Arginine decarboxylase				-5,70		-7,61		
	Arginyl-tRNA synthetase						-2,48		
	Arogenate/prephenylalaninenate dehydratase	-3,68			-4,81			-2,79	-3,21
	Asparagine synthase (glutamine-hydrolysing)					-2,09	-2,20		
	Asparaginyl-tRNA synthetase			-2,04					
	Aspartate aminotransferase, cytoplasmic	-2,13							
	Aspartate aminotransferase, mitochondrial	-2,27						-2,44	
	Aspartyl-tRNA synthetase					24,54	16,59		
	Aspartyl-tRNA(Asn)/glutamyl-tRNA(Gln) amidotransferase subunit A		-2,05		-2,98			-2,11	-2,44
	Aspartyl-tRNA(Asn)/glutamyl-tRNA(Gln) amidotransferase subunit B							-2,75	-3,69
	Bifunctional aspartate aminotransferase and glutamate/aspartate-prephenylalaninenate aminotransferase								-2,50
	Branched-chain amino acid aminotransferase		-2,09		-3,72		-2,30		-2,92

Carbamoyl-phosphate synthase large subunit				-3,52		-2,72		-2,51
D-3-phosphoglycerate dehydrogenase / 2-oxoglutarate reductase	-2,04					-2,22		
Fumaryl-acetoacetase	2,17	2,19		2,48				
Glutamate decarboxylase	5,83	8,19		4,89	3,38		2,81	2,94
Glutamate decarboxylase	2,28	2,04	2,82	2,89				
Glutamate dehydrogenase (NAD(P)+)						-2,59		
Glutamate synthase (NADH)							-2,05	-2,73
Glutamate--glyoxylate aminotransferase	-3,08	-2,50				-3,70	-2,63	-3,15
Glutamine synthetase	2,13	2,30		2,16				2,02
Glutamine synthetase	-2,90	-2,30						
Glutaminyl-tRNA synthetase						-2,09		
Glutamyl-tRNA synthetase	-2,08	-2,49		-2,58			-2,31	-2,49
Glutamyl-tRNA synthetase	-11,21	-11,73		-3,28			-7,40	-6,86
Glutamyl-tRNA synthetase	-11,21	-11,73		-3,28			-7,40	-6,86
Glutathione peroxidase		2,33						2,87
Glutathione peroxidase	2,69	3,10						2,36
Glutathione reductase (NADPH)		2,35						
Glutathione S-transferase								3,13
Glutathione S-transferase								10,50
Glutathione S-transferase		4,18						
Glutathione S-transferase		2,92						
Glutathione synthase						-2,07		
Glycerate dehydrogenase	-3,51	-2,95					-2,14	-2,14
Glycine cleavage system H protein								2,24
Glycine dehydrogenase	-2,54	-2,86		-2,36		-2,29	-2,84	-4,48
Glycine hydroxymethyltransferase		-3,17		-4,20		-3,35		-2,31
Glycyl-tRNA synthetase						-2,57		
Glycyl-tRNA synthetase				-2,36		-2,74		
Histidyl-tRNA synthetase	-2,04			-3,17		-2,42	-3,05	-2,68

Homocysteine S-methyltransferase						2,11			
Homogentisate 1,2-dioxygenase	4,29	3,73		2,15			4,36	3,37	
Homoserine kinase				-5,49					
Hydroxymethylglutaryl-CoA lyase		2,54							
Hydroxymethylglutaryl-CoA synthase		2,10		2,16					
Indole-3-glycerol phosphate synthase				-3,47					
Indole-3-pyruvate monooxygenase	8,95	7,22	6,27	6,47	3,01			5,15	
Isoleucyl-tRNA synthetase								-2,10	
Isovaleryl-CoA dehydrogenase						-2,33			
Isovaleryl-CoA dehydrogenase		6,20							
L-3-cyanoalanine synthase/ cysteine synthase	-2,08						-2,01		
L-ascorbate peroxidase	6,28			15,69				9,21	
L-ascorbate peroxidase						2,55		-3,83	
L-ascorbate peroxidase							-2,71		
L-cysteine desulfhydrase				-2,70					
LL-diaminopimelate aminotransferase		-2,31						-2,02	
Lysyl-tRNA synthetase, class II								-2,20	
Lysyl-tRNA synthetase, class II				-2,42				-2,32	
Maleylacetoacetate isomerase								2,24	
Methionine S-methyltransferase						-2,07			
Methionyl-tRNA formyltransferase				-2,37				-5,25	
Methionyl-tRNA synthetase							-2,63	-4,64	
Methionyl-tRNA synthetase		-2,65		-3,17		-2,42		-3,10	
Methylthioribose-1-phosphate isomerase	9,12	11,04		5,84					
N-acetyl-gamma-glutamyl-phosphate reductase		2,22							
Ornithine-oxo-acid transaminase					-2,07	-2,11			
Peptide-methionine (S)-S-oxide reductase				-2,29					
Peptidylprolyl isomerase	4,54			5,19				5,11	
Peptidylprolyl isomerase		-2,22				2,55		-4,64	
Peptidyl-prolyl isomerase D								-3,37	

Peroxiredoxin 6	12,11	11,74	16,30	12,43	8,02	3,47	8,17	11,94
Phenylalanine ammonia-lyase	-3,10	-3,76		-2,22				-2,92
Phenylalanyl-tRNA synthetase alpha chain	2,01							
Phenylalanyl-tRNA synthetase beta chain				-3,79				
Polyamine oxidase	3,71							
Polyphenol oxidase			-2,26					
Primary-amine oxidase		-2,98	-2,14				-2,16	-3,26
Proline iminopeptidase	2,04							
Prolyl 4-hydroxylase				2,37				
Prolyl 4-hydroxylase						-2,08		
Prolyl-tRNA synthetase	-7,88	-9,38		-2,47			-5,26	-17,86
Protein phosphatase		-2,28		-3,13				-3,76
Protein phosphatase 1G	2,26			2,83				
Protein phosphatase 1L				-2,09				
Protein phosphatase 1L	2,43			3,22				2,50
Protein phosphatase 1L						2,06		
S-adenosylmethionine synthetase	-2,89	-2,98		-2,27			-2,18	-2,73
S-adenosylmethionine synthetase	-4,36			-3,17				
Sarcosine oxidase / L-pipecolate oxidase		3,25		2,83				
Sarcosine oxidase / L-pipecolate oxidase	-2,07					-3,43		
Spermidine synthase							-2,22	
Spermidine synthase	-3,14			-2,31		-2,63	-3,55	-2,99
Thioredoxin reductase (NADPH)	3,89	17,35	4,61	6,18		2,80		
Thioredoxin reductase (NADPH)		2,17						
Thioredoxin reductase (NADPH)	-2,11		-2,40	-2,26		-2,34	-2,62	-2,94
Threonine synthase		-3,06	-2,27	-2,96		-2,41		-3,32
Threonyl-tRNA synthetase						-2,42		
Tryptophanyl-tRNA synthetase	-2,02			-2,09		-2,12	-2,33	
Late embryogenesis abundant protein D-29	7,54	5,12	9,02	23,37			3,81	11,01
Anthranilate synthase component I	-3,94	-3,85		-2,05			-5,85	-6,21

Biosynthesis Secondary Metabolites	Caffeic acid 3-O-methyltransferase / acetylserotonin O-methyltransferase		-3,54	-3,49			4,37	5,95		-2,11
	Caffeoylshikimate esterase	2,43			2,43	2,45				
	Chalcone isomerase								4,91	
	Chalcone synthase				3,45			2,35		
	Chalcone synthase							2,35		
	Cinnamyl-alcohol dehydrogenase								-2,52	
	Coniferyl-aldehyde dehydrogenase	10,06	12,23	11,39	10,15	4,29			8,26	10,64
	Coniferyl-aldehyde dehydrogenase	3,06	3,26		2,39				2,74	3,30
	Coniferyl-aldehyde dehydrogenase		5,25							
	Flavonoid 3'-monooxygenase	4,63							4,55	3,39
	Peroxidase	4,02						6,84		
	Peroxidase	11,05		6,11	5,23	5,86	3,60	14,21		7,69
	Peroxidase	2,57	4,50					8,35		
	Peroxidase		-2,17		-3,39					
	Peroxidase	-5,09		-5,78	-7,13	-6,68				
	Peroxidase									-2,35
	Scopoletin glucosyltransferase		4,73		-2,51					
	Shikimate O-hydroxycinnamoyltransferase							-2,32		
Carbohydrate Metabolism	1,4-alpha-glucan branching enzyme	-2,13								-2,14
	2,3-bisphosphoglycerate-independent phosphoglycerate mutase								3,44	
	2-hydroxyflavanone C-glucosyltransferase		2,58							2,58
	2-hydroxyflavanone C-glucosyltransferase									
	2-oxoglutarate dehydrogenase E1 component		2,02							
	2-oxoglutarate dehydrogenase E1 component		2,02							
	2-oxoisovalerate dehydrogenase E2 component (dihydrolipoyl transacylase)		2,43							
	3-isopropylmalate dehydrogenase				-2,30					
	4-aminobutyrosineate---pyruvate transaminase	2,02								
	4-coumarate--CoA ligase	-3,22	-3,49		-3,09				-3,65	-5,95
	6-phosphofructo-2-kinase / fructose-2,6-biphosphatase 3				-3,46			-2,73		

6-phosphofructokinase 1	2,70	2,76		2,21				
6-phosphofructokinase 1				2,21				
6-phosphofructokinase 1	2,64							
6-phosphofructokinase 1	2,64							
6-phosphofructokinase 1	-2,13			-2,22			-2,75	-3,88
6-phosphofructokinase 1						-2,28		
Acetate/butyrosineate---CoA ligase						-2,20		
Acetolactate synthase I/III small subunit	-3,28	-3,97		-2,37		-2,39		
Acetyl-CoA synthetase								-2,05
Aconitate hydratase		2,28						
Aconitate hydratase		2,28						
Aconitate hydratase		2,76						
Alcohol dehydrogenase (NADP+)							2,74	
Alcohol dehydrogenase (NADP+)	-2,86	-3,64					-3,88	-4,38
Alcohol dehydrogenase class-P							6,08	
Alpha-1,4-galacturonosyltransferase	-2,05	-2,75		-3,12				
Alpha-amylase	2,15							
Alpha-amylase		3,15						
Alpha-D-xyloside xylohydrolase						2,07		
Alpha-galactosidase	-2,00	-3,00		-3,12				-2,40
Alpha-glucosidase								-4,30
Alpha-L-arabinofuranosidase	-2,81	-2,48	-3,38	-2,33	-2,18	-2,05		-3,01
Alpha-L-arabinofuranosidase		-2,30		-2,06				-2,24
ATP citrate (pro-S)-lyase								-3,03
Basic endochitinase B	2,90	2,61	3,53				4,03	7,25
Basic endochitinase B				-3,02	-2,56			
Basic endochitinase B							5,61	
Basic endochitinase B	-4,49			-3,94				
Beta-amylase				-2,93				-15,97

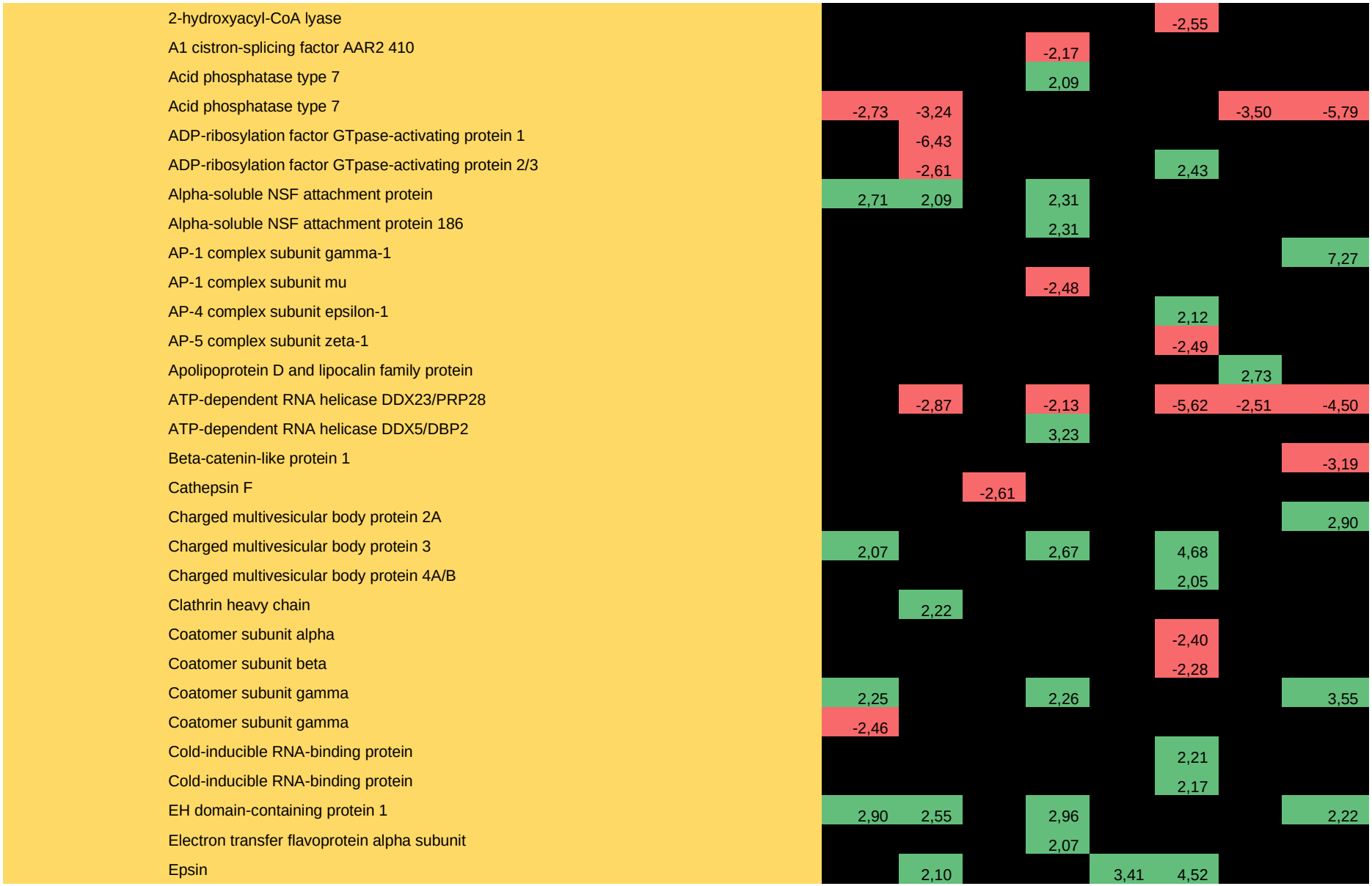
Beta-amylase							-11,70		
Beta-fructofuranosidase				-2,56					
Beta-fructofuranosidase	-2,72			-2,56					
Beta-fructofuranosidase			-2,72		-2,43	-2,11			
Beta-galactosidase							2,03		
Beta-glucosidase	3,99	2,67		3,43			2,49		
Beta-glucosidase								-2,19	
Beta-glucosidase	-6,87	-6,64	-3,32	-4,87		-2,02	-5,50	-6,83	
Cellulose synthase A		2,07					2,01	2,96	
Chitinase	5,97	4,84		2,97		2,49	4,15	5,39	
Chitinase	2,33	3,96					2,63	3,47	
CTP synthase	2,25			2,11		-2,32			
D-glycerate 3-kinase		-2,61		-3,36			-2,67		
Dihydrolipoamide dehydrogenase							-2,03		
Diphosphate-dependent phosphofructokinase									
Diphosphate-dependent phosphofructokinase	2,28					-2,57			
Formate dehydrogenase	3,04	4,05	3,59	2,04			2,13	2,87	
Fructokinase	3,82	3,72		-2,23	-2,81	-2,78		2,46	
Fructokinase							-3,01	-2,45	
Fructose-1,6-bisphosphatase I	-5,48	-4,21	-3,52	-5,13	-2,16		-5,73	-7,30	
Fructose-1,6-bisphosphatase I	-6,21	-6,40		-6,07		-2,46	-7,08	-9,33	
Fructose-1,6-bisphosphatase I	-2,28	-2,40					-2,94	-3,49	
Fructose-bisphosphate aldolase, class I	-4,06	-3,46					-3,23		
Fructose-bisphosphate aldolase, class I		2,61							
Fructose-bisphosphate aldolase, class I		2,50							
Fructose-bisphosphate aldolase, class I	-4,65	-3,93		-3,64			-2,88	-2,03	
Fructose-bisphosphate aldolase, class I	-3,60	-2,62					-2,49	-2,40	
Fructose-bisphosphate aldolase, class I	-13,73	-9,37		-4,10			-6,32	-7,32	
Fumarate hydratase, class II						-2,10	-2,36		

Galacturan 1,4-alpha-galacturonidase	3,15			4,47				2,64
GDP-D-mannose 3', 5'-epimerase				2,21				2,96
GDP-mannose 4,6-dehydratase	-2,67	-4,17						
Gephyrin	3,24	4,15	3,36	2,98			3,19	4,57
Glucose-1-phosphate adenylyltransferase	-2,68	-2,56		-2,79		-2,05	-2,33	-3,29
Glucose-1-phosphate adenylyltransferase	3,12	3,48	3,25	2,80			2,42	2,51
Glucose-1-phosphate adenylyltransferase	2,52			2,30				
Glucose-6-phosphate 1-dehydrogenase	-3,39	-4,30		-3,29			-2,90	-3,08
Glucose-6-phosphate 1-dehydrogenase						-2,01		
Glucose-6-phosphate 1-dehydrogenase	2,79							
Glucose-6-phosphate 1-epimerase								-8,21
Glutamate-1-semialdehyde 2,1-aminomutase	-2,38	-2,61		-2,40				
Glyceraldehyde 3-phosphate dehydrogenase (phosphorylating)				2,53				
Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosphorylating)	-2,44	-2,59					-2,17	-2,62
Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosphorylating)	-5,91	-8,02		-3,02			-4,06	-4,26
Glycogen phosphorylase							-2,13	
Granule-bound starch synthase	-2,79	-4,01					-5,20	-7,10
Hexokinase						-3,66		
Hexokinase							-2,09	
Hexokinase	2,30	4,03		3,70			2,55	3,24
Hexosaminidase		-2,36						
Hydroquinone glucosyltransferase						2,03		
Inositol-pentakisphosphate 2-kinase		2,70						
Isocitrate dehydrogenase (NAD+)						-2,02		
Lactoylglutathione lyase	2,50							
Lactoylglutathione lyase	2,27							
Lactoylglutathione lyase								-2,04
Magnesium-protoporphyrin IX monomethyl ester (oxidative) cyclase	-2,62	-2,53		-3,50	2,05		-4,23	-4,08
Magnesium-protoporphyrin O-methyltransferase		-6,68					-3,54	-3,93
Malate dehydrogenase (decarboxylating)	-2,45				-2,08	-2,85		

Malate dehydrogenase (NADP+)	-2,23	-2,68		-2,67				-2,43
Malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	14,45		15,65	14,00			45,28	43,92
Malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	3,73	3,94	3,15	2,70			3,32	3,59
Malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)		6,58				-5,03		
Malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	-2,17	-2,07		-2,71			-2,97	-2,88
Malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)						-2,15		
Malate synthase		2,20				-2,82		
Mannose-1-phosphate guanylyltransferase						-2,06		
Monodehydroascorbate reductase (NADH)						-2,29		
Monodehydroascorbate reductase (NADH)	-2,35					-5,00		
Multiple inositol-polyphosphate phosphatase / 2,3-bisphosphoglycerate 3-phosphatase	11,14			6,83			5,66	6,13
Myo-inositol-1-phosphate synthase		-2,09			-5,49	-9,68		-2,98
Oxalate---CoA ligase	2,51	2,39	3,65	2,60				
Pathogen-inducible salicylic acid glucosyltransferase					-2,34			
Pathogen-inducible salicylic acid glucosyltransferase	2,37	5,74				-2,67		
Pathogen-inducible salicylic acid glucosyltransferase						-32,36		
Pheophorbide a oxygenase	2,52	2,29		2,17				
Phosphoenolpyruvate carboxykinase (ATP)	-2,23					-2,42		
Phosphoenolpyruvate carboxylase				-2,28	-4,10	-2,59		
Phosphoenolpyruvate carboxylase							-2,75	-3,33
Phosphoenolpyruvate carboxylase	-3,48	-3,63		-4,07		-3,51	-2,92	-3,58
Phosphoglucan, water dikinase							-2,10	-3,06
Phosphoglucomutase	-2,58	-2,30		-2,07			-2,56	-3,45
Phosphoglycerate kinase		2,26	2,89	2,56				3,08
Phosphoglycerate kinase	-4,33	-4,38		-2,58				
Phosphoglycerate kinase	-2,96	-3,28					-2,77	-3,02
Phosphoglycolate phosphatase		-2,09						
Phosphoribosylglycinamide formyltransferase							-2,34	
Phosphoribulokinase	-3,38	-3,40		-2,21			-2,52	-4,11

Phosphoribulokinase	-3,43	-3,79		-2,27			-3,01	-4,69
Protochlorophyllide reductase					17,89	11,05		
Protochlorophyllide reductase						2,22		-2,06
Protoporphyrin/coproporphyrin ferrochelatase		2,15		2,85				
Protoporphyrinogen/coproporphyrinogen III oxidase		-2,94		-2,05				-2,27
Pyruvate decarboxylase	8,76	10,51	6,55	7,18	5,12			2,60
Pyruvate decarboxylase	5,06	4,55	4,81	2,53			2,61	2,42
Pyruvate dehydrogenase E1 component alpha subunit			-2,30		-2,13			
Pyruvate kinase						-3,07		
Pyruvate kinase	-2,18	-2,72		-2,21			-2,14	-2,08
Pyruvate kinase							-2,16	
Pyruvate, orthophosphate dikinase	-7,49	-8,93	-4,61	-11,45		-2,82	-5,66	-12,90
Pyruvate, phosphate dikinase-phosphate phosphotransferase /	-3,32	-6,77	-3,69		-2,04		-3,40	-9,76
Raffinose synthase				2,91		-2,24		
Ribokinase	-2,26						-2,02	-2,49
Ribose 5-phosphate isomerase A						2,68		3,44
Ribose-phosphate pyrophosphokinase	-8,22	-8,83		-6,00			-5,79	-8,72
Ribulose-bisphosphate carboxylase large chain	-3,97	-4,66		-2,63		-2,30	-4,01	-6,48
Ribulose-bisphosphate carboxylase small chain	-3,45	-2,31						-2,73
Ribulose-phosphate 3-epimerase	-3,93	-3,28					-3,47	-3,34
S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase	-4,35	-5,10	-3,95	-6,14		-2,60	-4,29	-4,49
Starch synthase				3,50				4,54
Starch synthase							2,33	
Starch synthase					2,02			
Sterol 3beta-glucosyltransferase				-2,24				
Succinate dehydrogenase (ubiquinone) flavoprotein subunit		2,17						
Succinate-semialdehyde dehydrogenase, mitochondrial		2,39	2,69					2,33
Succinyl-CoA synthetase alpha subunit	-6,46	-4,55		-5,13			-3,68	-3,91
Sucrose synthase	2,77	2,58					2,59	

	Sucrose synthase	-2,29	-2,68		-2,93		3,21			
	Sucrose-6-phosphatase									-2,08
	Sucrose-phosphate synthase					-2,57	-4,07			-2,14
	Sucrose-phosphate synthase						-2,52			-2,33
	Transaldolase	-3,10			-2,49					
	Trans-cinnamate 4-monooxygenase	-2,93	-2,86		-2,79					
	Transketolase	-6,39	-4,17		-3,90		-2,12	-4,17		-4,47
	Transketolase	3,92		5,19	3,38	2,92		3,25		3,83
	Trehalose 6-phosphate synthase/phosphatase	-4,07	-6,11	-3,92	-7,10	-2,33	-2,15	-3,50		-4,35
	Triosephosphate isomerase (TIM)	-2,88	-2,91					-3,40		-3,59
	UDP-glucose 4-epimerase		3,05		8,27			2,72		5,57
	UDP-glucose 4-epimerase	-3,74	-3,55		-2,71					-4,26
	UDP-glucose 6-dehydrogenase		-2,42		-2,52					
	UDP-glucuronate decarboxylase									2,50
	UDP-glucuronate decarboxylase						2,18			
	Xylan 1,4-beta-xylosidase		-2,02							
Cell Motility	Cofilin	2,27								
Cellular Process	(S)-2-hydroxy-acid oxidase							-2,83		-2,22
	(S)-2-hydroxy-acid oxidase	-3,33								
	26S proteasome non-ATpase regulatory subunit 5						2,73	2,33		
	26S proteasome regulatory subunit N1				2,77					
	26S proteasome regulatory subunit N1						-2,33			
	26S proteasome regulatory subunit N11				2,94					
	26S proteasome regulatory subunit N11				2,94					
	26S proteasome regulatory subunit N4	2,05			2,68					
	26S proteasome regulatory subunit T4	6,81		9,80	6,39		2,05			3,49
	2-alkenal reductase (NADP+)		2,16							
	2-alkenal reductase (NADP+)		2,16							
	2-carboxy-D-arabinitol-1-phosphatase						-16,84			-2,11


2-hydroxyacyl-CoA lyase
-2,55
A1 cistron-splicing factor AAR2 410
-2,17
Acid phosphatase type 7
2,09
Acid phosphatase type 7
-2,73
-3,24
-3,50
-5,79
ADP-ribosylation factor GTpase-activating protein 1
-6,43
ADP-ribosylation factor GTpase-activating protein 2/3
-2,61
2,43
Alpha-soluble NSF attachment protein
2,71
2,09
2,31
Alpha-soluble NSF attachment protein 186
2,31
AP-1 complex subunit gamma-1
7,27
AP-1 complex subunit mu
-2,48
AP-4 complex subunit epsilon-1
2,12
AP-5 complex subunit zeta-1
-2,49
Apolipoprotein D and lipocalin family protein
2,73
ATP-dependent RNA helicase DDX23/PRP28
-2,87
-2,13
-5,62
-2,51
-4,50
ATP-dependent RNA helicase DDX5/DBP2
3,23
Beta-catenin-like protein 1
-3,19
Cathepsin F
-2,61
Charged multivesicular body protein 2A
2,90
Charged multivesicular body protein 3
2,07
2,67
4,68
Charged multivesicular body protein 4A/B
2,05
Clathrin heavy chain
2,22
Coatomer subunit alpha
-2,40
Coatomer subunit beta
-2,28
Coatomer subunit gamma
2,25
2,26
3,55
Coatomer subunit gamma
-2,46
Cold-inducible RNA-binding protein
2,21
Cold-inducible RNA-binding protein
2,17
EH domain-containing protein 1
2,90
2,55
2,96
2,22
Electron transfer flavoprotein alpha subunit
2,07
Epsin
2,10
3,41
4,52

Ferritin heavy chain	5,98	5,36	5,62	6,46	4,06		2,50	
Heterogeneous nuclear ribonucleoprotein A1/A3	2,64			2,17		2,21		
Heterogeneous nuclear ribonucleoprotein R	-4,40	-5,63		-4,73				-2,77
Htra serine peptidase 2	3,18		2,98				2,22	3,61
IK cytokine	3,35	3,11						3,63
Legumain		2,84						
Nitric oxide synthase-interacting protein		-2,75						-2,44
Peroxin-10		-4,84		-3,26				-3,11
Peroxin-5	2,16							
Phytanoyl-CoA hydroxylase	2,48	2,54						
Plasminogen activator inhibitor 1 RNA-binding protein	4,09		4,05	2,79		4,64	4,37	7,35
Plasminogen activator inhibitor 1 RNA-binding protein	2,19	2,68	4,33	3,62	4,30	5,79	2,55	3,99
Plasminogen activator inhibitor 1 RNA-binding protein	3,39			2,41	4,33	5,84		
Pre-mRNA-processing factor 39						-2,39		
Pre- mRNA -splicing factor ISY1								3,14
Pre- mRNA -splicing factor SYF1				-2,17		-2,23		
Pre- mRNA -splicing factor SYF2	3,88	3,80	6,13	5,11		4,98		3,28
Pre- mRNA -splicing helicase BRR2		-2,43		-3,37				
Programmed cell death protein 5	-3,19	-3,83					-4,05	-6,33
Proteasome assembly chaperone 4								3,59
Proteasome component ECM29	-5,16	-5,86		-4,72		-3,43		-2,78
Protein CWC15	4,14			4,01		3,67		
Protein disulfide-isomerase A1								2,62
Protein FAM50 (Protein XAP5 CIRCADIAN TIMEKEEPER)							-10,55	-15,68
Rab GDP dissociation inhibitor	2,07			2,22				
Serine/arginine-rich splicing factor 7	2,27					2,30		
Spartin	6,74	7,99		16,63			4,04	10,52
Splicing factor 1	-2,13	-3,32		-2,07				
Splicing factor 3A subunit 2				2,03		2,58		

	Splicing factor 3B subunit 1	2,16							2,02
	Splicing factor 3B subunit 2								10,15
	Splicing factor U2AF 65 kDa subunit								-2,86
	Stromal membrane-associated protein		4,38		3,00				
	Superoxide dismutase, Cu-Zn family	9,32		7,60	8,51	8,60	5,07	7,45	6,90
	Survival of motor neuron-related-splicing factor 30				2,67		2,45		2,75
	THO complex subunit 2	3,91							
	THO complex subunit 4						2,22		
	Transmembrane 9 superfamily member 2/4	-2,42							
	Tubulin alpha		-3,27						-4,58
	Tuftelin-interacting protein 11								-2,70
	U2 small nuclear ribonucleoprotein A'						2,02		
	U4/U6,U5 tri-snrnp-associated protein 2	-2,16			-2,24	-2,70	-2,73		
	Vam6/Vps39-like protein vacuolar protein sorting-associated protein 39	2,10	2,13		2,25				
	Vesicle transport protein SEC22							2,04	
	WD repeat-containing protein 45		2,16						2,47
Chaperone And Folding Catalyst	ATP-dependent Clp protease ATP-binding subunit clpb	2,41	2,11		3,05				
	ATP-dependent Clp protease ATP-binding subunit clpb	3,63		3,55	3,55	4,07	2,73		
	ATP-dependent Clp protease ATP-binding subunit clpb	3,50							
	ATP-dependent Clp protease ATP-binding subunit clpc	3,63		3,55	3,55	4,07	2,73		
	ATP-dependent Clp protease ATP-binding subunit clpx		-2,15						
	DnaJ homolog subfamily B member 4								4,11
	DnaJ homolog subfamily B member 4	2,78			2,65				2,53
	DnaJ homolog subfamily B member 4				2,04				
	DnaJ homolog subfamily C member 13								3,64
	Glutaredoxin 3	48,77	29,09		44,10		-16,27	12,42	12,87
	Glutaredoxin 3	3,23	4,01	4,75	4,29	2,64	2,34	2,27	2,92
	Glutaredoxin 3								-2,04
	Heat shock 70kda protein 1/2/6/8		-2,06						-2,22

Heat shock 70kda protein 4	12,05	10,27		7,55			11,82	
Heat shock protein 90kda beta	2,30							
HSP20 family protein	14,28	8,38	5,86	5,79			3,82	4,32
HSP20 family protein							4,36	
HSP20 family protein	3,32		3,69					2,03
HSP20 family protein	17,19		9,73	6,80				
HSP20 family protein	4,42	3,81		2,95				
HSP20 family protein								-8,01
Hsp70-interacting protein	2,35							
Late embryogenesis abundant protein	35,16	33,55	46,05	59,82			11,32	5,99
Late embryogenesis abundant protein	4,76	4,02		3,86			3,55	6,54
Late embryogenesis abundant protein D-29	15,08		9,02	4,97				3,42
Late embryogenesis abundant protein D-29	5,60	7,12	4,76	7,71		2,44	4,02	3,96
Late embryogenesis abundant protein D-34-like	5,44	4,13		13,16			4,14	4,14
Late embryogenesis abundant protein D-34-like	-2,18							-2,62
Late embryogenesis abundant protein Dc3-like	39,15			28,07	6,24	3,14	31,86	30,83
Late embryogenesis abundant protein_B	23,26	10,84	13,11	22,63			11,32	5,99
Molecular chaperone DnaK		-5,93						-4,10
Molecular chaperone grpe	2,04	2,01						
Molecular chaperone grpe								-2,38
Molecular chaperone hscb	2,12	2,19						
Molecular chaperone htpg							-2,12	-2,08
Nascent polypeptide-associated complex subunit alpha	4,42	2,97		3,15		2,05		
Nascent polypeptide-associated complex subunit alpha	2,18			2,12				
Nascent polypeptide-associated complex subunit beta	2,17					2,23	2,87	3,64
PAT complex subunit CCDC47	-4,38	-4,68		-3,15	-2,44	-2,81	-3,07	-3,72
Peptidyl-prolyl cis-trans isomerase B (cyclophilin B)		-2,77						-5,20
Prefoldin subunit 2		-2,70						
Protein disulfide-isomerase		-2,84						
Protein disulfide-isomerase A6	2,76	2,53		2,23			2,17	2,18

	Stress-induced-phosphoprotein 1	-2,11								
	Suppressor of tumorigenicity protein 13	2,63								
	Thioredoxin 1									3,02
	Thioredoxin 1									-2,09
	Thioredoxin 1	-2,33	-2,96						-2,15	-3,37
	Trigger factor	-2,25	-2,66		-2,42				-3,84	-4,12
Chromosome	Activating signal cointegrator								3,50	
	Centromeric protein E	10,76			48,30	5,16	2,81	11,95		21,77
	Centromeric protein E	5,55		6,06	5,13		-2,10			
	Chromatin remodeling protein	-5,49	-7,70		-4,52					-3,86
	Chromodomain-helicase-DNA-binding protein 4				-2,44					
	COMPASS component SWD1									2,31
	High mobility group B protein, plant		-3,29							
	Histone H2A								3,00	
	Histone H2A						2,91			4,10
	Histone H2A				2,53		4,67			
	Histone H2A						2,58			
	Histone H3						3,76			
	Histone H3				4,59					
	Histone H4						2,32			
	N-alpha-acetyltransferase 15/16, nata auxiliary subunit	-3,60	-4,31		-3,04		2,63	-3,52		-5,22
	Nucleosome assembly protein 1-like 1	2,46								
	Nucleosome assembly protein 1-like 1		-2,06					-3,45		-2,96
	Paired amphipathic helix protein Sin3a				2,14	2,43		-2,78		
	PAX-interacting protein 1		4,44							
	PAX-interacting protein 1	5,42			9,81		4,69			
	SAGA-associated factor 29	8,37	11,01		7,86		3,11			
	Structural maintenance of chromosome 1	2,04			2,88					
	3,4-dihydroxy 2-butanone 4-phosphate synthase / GTP cyclohydrolase II	-11,96	-13,31		-21,69	-4,87	-3,95	-4,97		-5,79

	4a-hydroxytetrahydrobiopterin dehydratase	-8,91	-11,42		-6,32			-5,79	-11,46
	5-formyltetrahydrofolate cyclo-ligase		-2,96			-3,64	-2,56		
	Isochorismate synthase / 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase / 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase / o-succinylbenzoate synthase						2,42		
	MPBQ/MSBQ methyltransferase		-2,02					-2,50	-2,71
	NAD(P)H dehydrogenase (quinone)						2,56		
	Phosphopantothenate---cysteine ligase (ATP)						-2,46		
	Pyridoxal 5'-phosphate synthase / NAD(P)H-hydrate epimerase		2,82						2,17
	Pyridoxal 5'-phosphate synthase pdxs subunit							-2,15	
	Pyridoxal phosphate phosphatase PHOSPHO2				3,21		2,32		2,02
	Pyridoxamine 5'-phosphate oxidase								2,18
	Pyridoxine kinase	-3,00	-3,55		-4,73		-2,69		-4,24
	Type II pantothenate kinase		4,72						
	Ubiquinone biosynthesis monooxygenase Coq6	9,99	5,75		5,60			6,81	6,00
	Ubiquinone biosynthesis protein COQ9								2,33
	Uroporphyrinogen decarboxylase	-2,95	-2,78		-2,89				
Cofactors And Vitamins Metabolism	Uroporphyrinogen decarboxylase	-3,64	-4,55		-3,32				
	Uroporphyrinogen-III synthase							-2,27	-2,33
Cytoskeleton	Actin, other eukaryote							2,50	
	Actin, other eukaryote								3,56
	Actin-related protein 7, plant								2,15
	Katanin p60 ATpase-containing subunit A1	3,50		3,28	2,98			3,22	2,97
	Myosin V						2,14		
	Tubulin--Tyrosine ligase-like protein 12								-2,37
Energy Metabolism	Aarf domain-containing kinase	-2,13			-2,02		-2,15		
	Aarf domain-containing kinase	-4,19			-2,84			-4,39	-3,79
	Aarf domain-containing kinase							-3,66	-4,61
	Aarf domain-containing kinase	-5,21	-6,27		-8,02		-4,12	-5,19	-4,85
	Adenylyl-sulfate reductase (glutathione)	-3,05			-4,89			-2,84	-5,28
	ATP synthase protein I	-2,93	-4,04		-5,05		2,66		

ATpase family AAA domain-containing protein 3A/B				-2,28				
ATpase family AAA domain-containing protein 3A/B				-2,28				
Bola-like protein 1	2,21						2,21	
Carbonic anhydrase	-2,88	-4,59					-4,14	-5,40
Complement component 1 Q subcomponent-binding protein, mitochondrial 79	2,46			4,57				
Cytochrome b6-f complex iron-sulfur subunit	-6,25	-8,98				2,45		-12,12
Cytochrome c	2,53	2,57		2,05				
Electron-transferring-flavoprotein dehydrogenase	4,87	2,45		2,86			3,81	2,50
Ferredoxin--NADP+ reductase				-3,03		-2,24	-2,36	-2,97
Ferredoxin--NADP+ reductase	-3,56	-3,96		-3,52			-3,98	-5,62
Ferredoxin--NADP+ reductase	-3,94	-2,64		-3,06			-5,64	-6,14
Ferredoxin-nitrite reductase				-3,34		-3,61		
Ferredoxin-thioredoxin reductase catalytic chain								-3,02
F-type H+-transporting ATpase subunit b	-7,76	-7,98		-7,36				
F-type H+-transporting ATpase subunit delta				-2,26		3,24		
Glutamate synthase (ferredoxin)	5,55		6,06	5,13		-2,10		
H+-transporting ATpase	-3,80		-2,64	-2,72				
H+-transporting ATpase	10,39	6,85	11,02	14,69	9,39	2,75	4,79	5,31
H+-transporting ATpase	-3,44	-5,20		-2,85				
Inorganic pyrophosphatase	-2,79	-3,78		-3,59				-3,56
Inorganic pyrophosphatase	-3,22							
Light-harvesting complex I chlorophyll a/b binding protein 1	-3,10	-4,25		-3,53		2,03	-2,79	-5,80
Light-harvesting complex I chlorophyll a/b binding protein 3	-2,06							
Light-harvesting complex I chlorophyll a/b binding protein 4	-3,10	-3,42		-2,18		2,30	-3,55	-7,07
Light-harvesting complex II chlorophyll a/b binding protein 1	-11,45	-5,90		-3,77		2,12	-3,35	-3,94
Light-harvesting complex II chlorophyll a/b binding protein 1	-20,48	-11,34		-3,11		2,25	-4,20	-5,57
Light-harvesting complex II chlorophyll a/b binding protein 4	-12,27	-18,87		-9,30				-13,22
Light-harvesting complex II chlorophyll a/b binding protein 5	-3,11	-2,97		-2,22		2,02		-2,97
Light-harvesting complex II chlorophyll a/b binding protein 6	-5,75	-5,33				2,20		-6,04

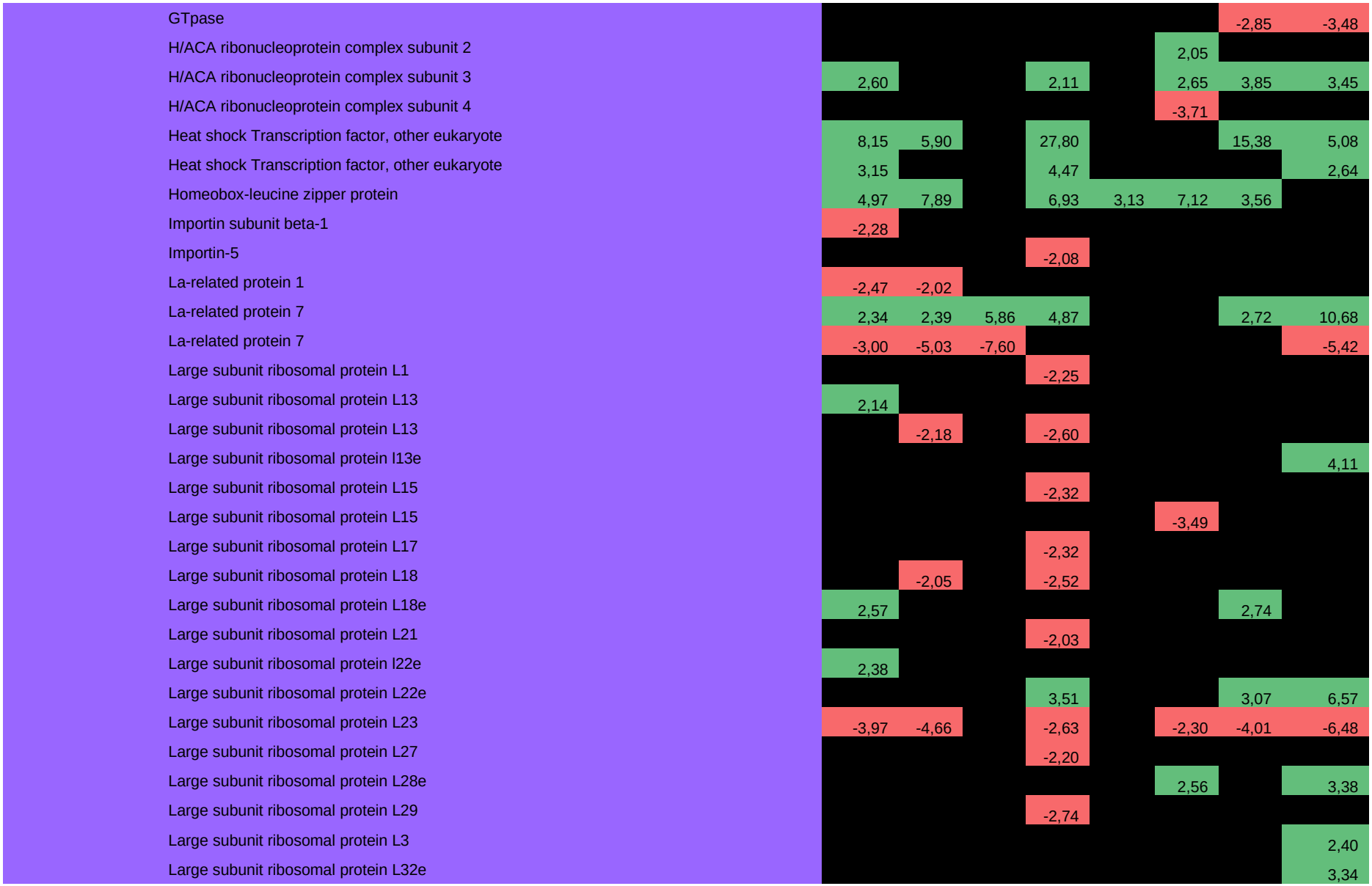
Methanethiol oxidase		2,06						
Mitochondrial FAD-linked sulfhydryl oxidase	16,42	10,02	15,97	12,29			13,91	11,38
Mitochondrial-processing peptidase subunit alpha	-8,02	-7,73		-7,15			-4,03	-5,58
MYG1 exonuclease	-4,02	-5,71		-3,46	-3,15	-2,23		
NAD(P)H-quinone oxidoreductase subunit I		-2,69						-2,07
NADH dehydrogenase (ubiquinone) Fe-S protein 6						2,21		
NADH:quinone reductase (non-electrogenic)						-2,15		
Nitrate reductase (NAD(P)H)	-4,13			-2,82			-2,92	-3,75
Outer membrane protein insertion porin family	-2,30	-2,05				-2,26		
Photosystem I subunit II	-7,78	-7,28		-2,77		2,17	-4,38	-9,02
Photosystem I subunit III	-12,39	-14,18		-8,27				-7,15
Photosystem I subunit IV	-5,67	-4,99					-4,03	-12,31
Photosystem I subunit psan	-7,13	-8,82		-5,68			-3,31	-6,54
Photosystem II 22kda protein	-2,17							
Photosystem II oxygen-evolving enhancer protein 1	-11,82	-22,65		-8,04			-4,33	-27,69
Photosystem II oxygen-evolving enhancer protein 2	-7,42	-9,45		-2,65			-2,88	-12,45
Photosystem II oxygen-evolving enhancer protein 3	-2,67	-4,40				2,15		-4,61
Photosystem II oxygen-evolving enhancer protein 3	-7,08	-7,61						
Photosystem II oxygen-evolving enhancer protein 3	-10,58	-10,21						
Photosystem II Psb27 protein	-4,17	-5,58		-2,39				-5,53
Prohibitin 1								4,54
Prohibitin 2		12,44						12,51
Prohibitin 2 290	-3,52			-2,19				
Reticulon-4-interacting protein 1, mitochondrial 283								2,02
Sedoheptulose-bisphosphatase	-8,22	-7,83		-5,36	-2,07		-6,80	-11,12
Sulfite oxidase		-2,46		-2,33	-2,35	-2,19		
Sulfite reductase (ferredoxin)	-2,00	-2,56					-2,82	-4,02
Transferase CAF17, mitochondrial	4,21		4,05				3,26	4,59
Ubiquinol oxidase		3,47						

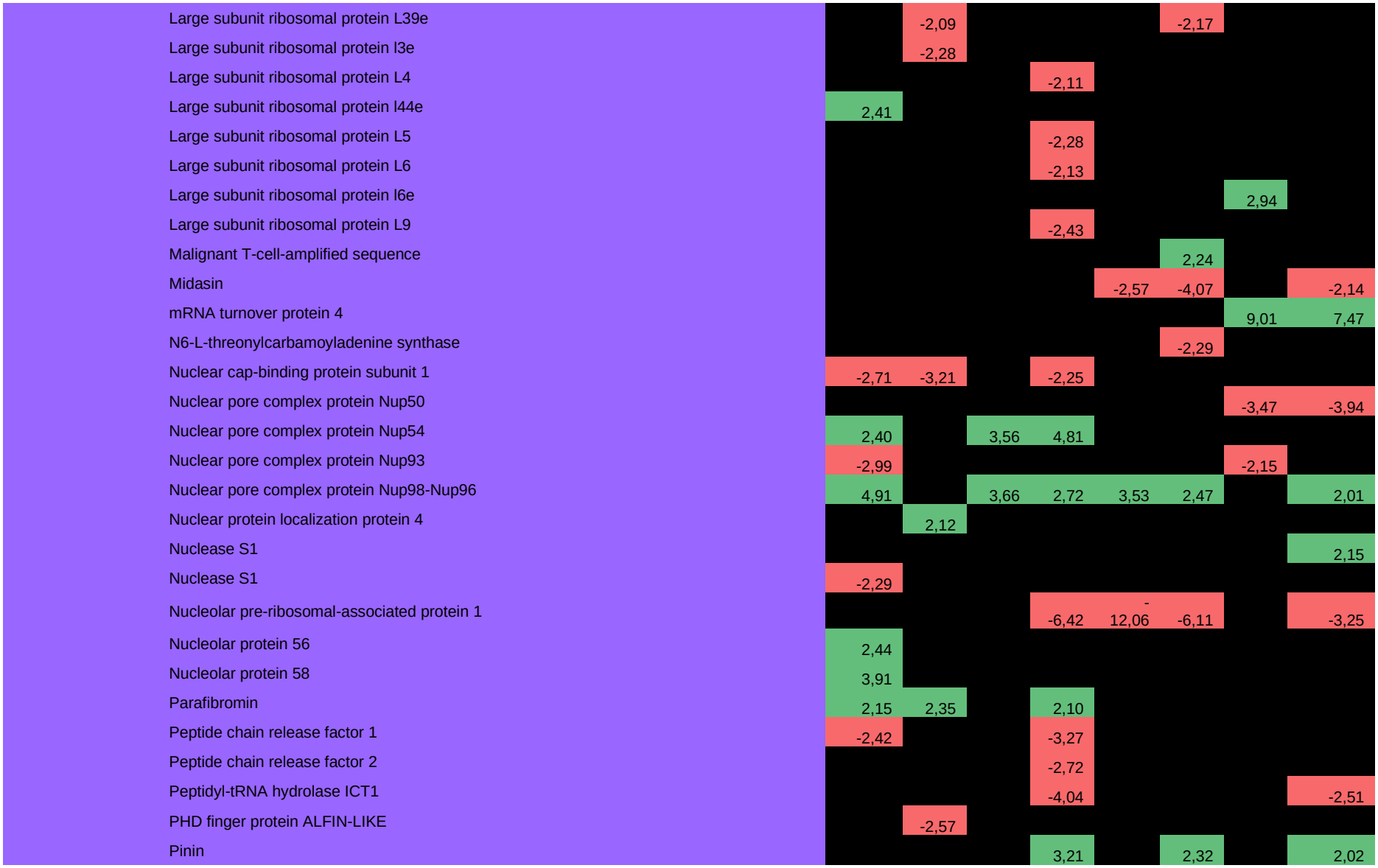
	Ubiquinol oxidase		-2,02					-2,27	-2,77
	V-type H+-transporting ATpase subunit A	-2,34			-2,39				
	V-type H+-transporting ATpase subunit A		2,48						
	V-type H+-transporting ATpase subunit D								3,28
	V-type H+-transporting ATpase subunit D	3,44	3,99	6,96	5,94	2,99		3,09	4,27
	V-type H+-transporting ATpase subunit d						-2,12		
	V-type H+-transporting ATpase subunit G								2,43
Environmental Adaptation	Calcium-binding protein CML					-3,13	-3,22		
	Calcium-binding protein CML	-3,14	-4,07		-3,50	-3,52	-3,31	-3,07	-5,08
	Calcium-binding protein CML				2,93				
	Calcium-binding protein CML				2,93				
	Calcium-dependent protein kinase	7,31	4,96		5,33			7,14	8,52
	Calcium-dependent protein kinase							-7,02	
	Disease resistance protein RPM1		2,02						
	Disease resistance protein RPM1				-2,22			-2,08	-2,66
	Disease resistance protein RPM1		-3,19		-2,76				-3,22
	Disease resistance protein RPM1	-2,25	-2,66		-2,42			-3,84	-4,12
	NADH dehydrogenase		2,09						
	Pto-interacting protein 1	-2,28							
	Pto-interacting protein 1				4,97	-2,23			
	Ubiquinol-cytochrome c reductase subunit 6				-3,52		-2,72		-2,51
Enzyme	ATpase family AAA domain-containing protein 2	3,62	4,71		14,72	2,19			4,02
	Dehydrogenase/reductase SDR family member 7				-2,19				
	Deoxyhypusine synthase	-5,55	-5,23		-2,88				
	Ferric-chelate reductase	-9,50			-25,43		-4,12	-6,80	
	Protein adenylyltransferase	2,19							
Folding, Sorting And Degradation	Serine carboxypeptidase 1			-2,60					
	Ubiquitin-conjugating enzyme E2 J1		-2,03	-2,24					-2,43
	Stromal processing peptidase, chloroplastic			-2,12	-2,28		-2,57		-4,86

Preprotein translocase subunit seca			-2,10	-2,66				
Abhydrolase domain-containing protein 17				-2,22				-3,57
Aspartyl aminopeptidase							2,38	
ATP-dependent Lon protease						-2,80		
B-cell receptor-associated protein 31	3,76	4,66					5,05	9,16
B-cell receptor-associated protein 31				2,70				
Blocked early in transport 1	2,13			2,75		2,76		
Calnexin	2,04						17,04	
Carboxyl-terminal processing protease								2,29
Carboxyl-terminal processing protease				3,92				
Cell division protease ftsh	7,66			88,47	10,31			13,38
Cell division protease ftsh	-3,38	-4,29		-3,30			-2,75	-3,45
Cell division protease ftsh	-2,72	-3,16					-2,81	-4,75
Cell division protease ftsh								-2,38
CO(2)-response secreted protease	-8,62			-5,41				
COP9 signalosome complex subunit 3	-10,88	-15,61		-6,57			-7,07	-21,42
COP9 signalosome complex subunit 7							-2,09	
Cullin						-2,13		
Cullin-associated NEDD8-dissociated protein 1						-2,03		
Cysteine proteinase RD21						3,10		
F-box and leucine-rich repeat protein 2/20	-3,68			-4,81			-2,79	-3,21
Fused signal recognition particle receptor				-2,12				
Glucose-induced degradation protein 8						2,10		
IAA-amino acid hydrolase	2,25						2,47	
Insulysin				-3,26				
KDEL-tailed cysteine endopeptidase				-5,49				
Methionyl aminopeptidase	4,06							2,63
Methionyl aminopeptidase							-2,26	-2,07
Phospholipase A-2-activating protein							-2,74	

Phytepsin				-2,70							
Polyribonucleotide nucleotidyltransferase				-2,49			-2,36	-2,05			
Preprotein translocase subunit seca							-2,47				
Presequence protease							-2,23				
Prolyl oligopeptidase	-4,08	-4,02		-2,55			-2,66	-2,76	-5,95		
Protease Do-like 2, chloroplastic	-3,05	-2,92		-3,08				-2,79	-3,85		
Protein deglycase									-3,25		
Protein deglycase				-6,42	-12,06		-6,11		-3,25		
Protein kinase C substrate 80K-H	2,66										
Protein-L-isoaspartate(D-aspartate) O-methyltransferase	7,31	4,34		7,84					3,65		
Puromycin-sensitive aminopeptidase	3,62	4,71		14,72	2,19				4,02		
Serine carboxypeptidase-like clade I									3,08		
Serine carboxypeptidase-like clade II	2,28	2,89									
Serpin B	5,77	5,43		3,30	4,06			7,97	3,86		
Synaptobrevin homolog YKT6				2,22			2,10				
Syntaxin 1B/2/3	-2,47	-2,21						-3,43	-3,65		
Syntaxin 1B/2/3							2,23				
Syntaxin 5	-2,26			-2,14							
Syntaxin 7				2,47				2,06	3,03		
Syntaxin of plants SYP5									2,00		
Syntaxin of plants SYP7		-2,77									
Tether containing UBX domain for GLUT4	2,74	2,23		2,38			2,91				
Transitional endoplasmic reticulum ATpase	2,01										
Transitional endoplasmic reticulum ATpase	2,03										
Translocation protein SEC63							2,39				
Ubiquitin conjugation factor E4 B									-2,10		
Ubiquitin-like 1-activating enzyme E1 A								-2,03	-2,72		
UBX domain-containing protein 1							2,19				
Xaa-Pro aminopeptidase									-2,34		

	Zinc metalloprotease EGY, chloroplastic						-2,53			
	Zinc metalloprotease EGY, chloroplastic						-3,43	-2,64	-3,97	
	Phytepsin	2,64	2,65	2,07	2,44			2,37		
	OTU domain-containing protein 6	2,14	2,40	2,16				2,08		
	Jasmonoyl-L-amino acid hydrolase	3,12	2,92	2,42						
	Phytepsin	2,50		2,52	2,19	3,26	4,03			
	Protein deglycase	2,52	2,00	2,59	2,45					
	Chaperonin groel			3,83	2,65		2,89	2,16	2,39	
	Dnaj homolog subfamily A member 2	3,58	4,23	5,22	7,66	3,22	2,33			
	Serine carboxypeptidase-like clade II	16,12	12,09	16,54	15,53			3,71	5,59	
Genetic Information Processing	30S ribosomal protein S31								4,85	
	50S ribosomal protein 6				-3,11					
	Apoptotic chromatin condensation inducer in the nucleus	2,02						2,34	2,88	
	ATP-binding cassette, subfamily F, member 3	23,94	18,51	31,99	29,27	23,89	7,69	9,63	16,96	
	ATP-dependent RNA helicase DHX29	-2,05			-2,25		-3,11		-2,43	
	Cellular nucleic acid-binding protein	29,70			6,29		31,35			
	Cyclin-dependent kinase 12/13	-2,09					3,13			
	Density-regulated protein	2,24		2,45			2,67			
	DNA repair protein RAD7	-3,22	-3,49		-3,09			-3,65	-5,95	
	DNA topoisomerase II								2,92	
	DNA-directed RNA polymerase II subunit RPB3				-3,06		-4,58	-2,11	-2,65	
	DNA-directed RNA polymerases I, II, and III subunit RPABC1						2,11			
	Elongation factor 2						-2,77			
	Elongation factor G						-3,03			
	Elongation factor G	-3,23	-4,06		-4,57			-3,50	-4,33	
	Elongation factor Ts				-2,24			-2,12	-2,35	
	Enhancer of mRNA-decapping protein 4						-2,18			
	Eukaryotic translation initiation factor 2C		2,50							
	Eukaryotic translation initiation factor 2C	-4,54	-4,06		-3,06			-2,75	-2,90	
	Exonuclease 3'-5' domain-containing protein 1						-2,63			


GTpase
-2,85 -3,48
H/ACA ribonucleoprotein complex subunit 2
2,05
H/ACA ribonucleoprotein complex subunit 3
2,60 2,11 2,65 3,85 3,45
H/ACA ribonucleoprotein complex subunit 4
-3,71
Heat shock Transcription factor, other eukaryote
8,15 5,90 27,80 15,38 5,08
Heat shock Transcription factor, other eukaryote
3,15 4,47 2,64
Homeobox-leucine zipper protein
4,97 7,89 6,93 3,13 7,12 3,56
Importin subunit beta-1
-2,28
Importin-5
-2,08
La-related protein 1
-2,47 -2,02
La-related protein 7
2,34 2,39 5,86 4,87 2,72 10,68
La-related protein 7
-3,00 -5,03 -7,60 -5,42
Large subunit ribosomal protein L1
-2,25
Large subunit ribosomal protein L13
2,14
Large subunit ribosomal protein L13
-2,18 -2,60
Large subunit ribosomal protein l13e
4,11
Large subunit ribosomal protein L15
-2,32
Large subunit ribosomal protein L15
-3,49
Large subunit ribosomal protein L17
-2,32
Large subunit ribosomal protein L18
-2,05 -2,52
Large subunit ribosomal protein L18e
2,57 2,74
Large subunit ribosomal protein L21
-2,03
Large subunit ribosomal protein l22e
2,38
Large subunit ribosomal protein L22e
3,51 3,07 6,57
Large subunit ribosomal protein L23
-3,97 -4,66 -2,63 -2,30 -4,01 -6,48
Large subunit ribosomal protein L27
-2,20
Large subunit ribosomal protein L28e
2,56 3,38
Large subunit ribosomal protein L29
-2,74
Large subunit ribosomal protein L3
2,40
Large subunit ribosomal protein L32e
3,34


Large subunit ribosomal protein L39e
-2,09
-2,17
Large subunit ribosomal protein l3e
-2,28
Large subunit ribosomal protein L4
-2,11
Large subunit ribosomal protein l44e
2,41
Large subunit ribosomal protein L5
-2,28
Large subunit ribosomal protein L6
-2,13
Large subunit ribosomal protein l6e
2,94
Large subunit ribosomal protein L9
-2,43
Malignant T-cell-amplified sequence
2,24
Midasin
-2,57
-4,07
-2,14
mRNA turnover protein 4
9,01
7,47
N6-L-threonylcarbamoyladenine synthase
-2,29
Nuclear cap-binding protein subunit 1
-2,71
-3,21
-2,25
Nuclear pore complex protein Nup50
-3,47
-3,94
Nuclear pore complex protein Nup54
2,40
3,56
4,81
Nuclear pore complex protein Nup93
-2,99
-2,15
Nuclear pore complex protein Nup98-Nup96
4,91
3,66
2,72
3,53
2,47
2,01
Nuclear protein localization protein 4
2,12
Nuclease S1
2,15
Nuclease S1
-2,29
Nucleolar pre-ribosomal-associated protein 1
-6,42
-12,06
-6,11
-3,25
Nucleolar protein 56
2,44
Nucleolar protein 58
3,91
Parafibromin
2,15
2,35
2,10
Peptide chain release factor 1
-2,42
-3,27
Peptide chain release factor 2
-2,72
Peptidyl-tRNA hydrolase ICT1
-4,04
-2,51
PHD finger protein ALFIN-LIKE
-2,57
Pinin
3,21
2,32
2,02

Pleiotropic regulator 1	-2,82	-2,93				-2,20	-2,00	-2,68
Poly								2,33
Poly(rc)-binding protein 3/4								11,76
Poly(rc)-binding protein 3/4	-2,66			-2,36				
Polyadenylate-binding protein 2	17,30		11,29	7,06		4,27	12,00	8,46
Pre-60S factor REI1	2,39	2,55		4,16				
Protein LSM14				2,18		2,29		3,15
Protein TIF31				2,06				
Protein TIF31								-2,33
Protein TIF31							-3,60	-3,39
Pumilio RNA-binding family	2,08	2,32						-2,30
Putative Transcription factor	2,13		2,50			3,79		
Ras GTpase-activating protein-binding protein 2				2,11		2,10		3,43
Ras GTpase-activating protein-binding protein 2						2,07		
Regulator of nonsense transcripts 1							-2,23	-2,54
Regulator of Ribosome biosynthesis	2,43							
Replication factor A1							5,40	4,93
Replication factor A2	3,71							2,71
Ribonuclease T2				-3,74	-3,33	-2,64		
Ribonuclease T2				-8,86	-7,37	-9,80		
Ribosome-binding factor A				-3,06				-3,08
RNA polymerase II-associated factor 1								4,43
RNA-binding protein FUS	3,83			3,86	2,24	3,54		
RNA-binding protein Musashi	2,59			2,39	2,07	2,99	2,61	
RNA-dependent RNA polymerase	-7,76	-7,98		-7,36				
rRNA 2'-O-methyltransferase fibrillarin	-6,68	-9,83		-5,81			-5,18	-7,30
rRNA small subunit pseudouridine methyltransferase Nep1		-3,27				-2,50	-2,85	-2,20
SAP domain-containing ribonucleoprotein	2,03					3,30	2,63	2,92
SCY1-like protein 1	-10,00	-11,73		-6,12			-5,22	-10,44

Small nuclear ribonucleoprotein D2	2,55			3,57			3,29	8,94	
Small subunit ribosomal protein S1				-2,06					
Small subunit ribosomal protein S12e		-2,24							
Small subunit ribosomal protein S15e	3,02					2,96	2,32		
Small subunit ribosomal protein S20				-2,55					
Small subunit ribosomal protein S26e	-3,39	-4,30		-3,29			-2,90	-3,08	
Small subunit ribosomal protein S27e	-4,35			-3,58					
Small subunit ribosomal protein S30e	2,22								
Small subunit ribosomal protein S9		-2,12		-2,24					
Small subunit ribosomal protein sae								2,03	
SnRNA-activating protein complex subunit 3	9,73	10,10	12,75	13,91	3,07		5,46	7,08	
S-phase kinase-associated protein 1				3,21		2,32		2,02	
Transcription elongation factor SPT6		-2,06							
Transcription factor SPEECHLESS	-5,12	-7,46		-4,40				-12,13	
Transcription factor VIP1						2,28			
Transcription initiation factor TFIIE subunit alpha						2,17			
Translation initiation factor 1	2,65								
Translation initiation factor 1			-2,04						
Translation initiation factor 2 subunit	2,36			2,52	2,99	2,45			
Translation initiation factor 2 subunit	2,31					2,04			
Translation initiation factor 2 subunit						-2,10			
Translation initiation factor 3 subunit K		2,66							
Translation initiation factor 4E	2,64			2,68		2,01			
Translation initiation factor 4G			2,59	2,11					
Translation initiation factor 5	2,08			2,23					
Translation initiation factor 5A								2,17	
Translation initiation factor 5A						2,09		3,37	
Translation initiation factor eif-2B subunit delta	-5,17	-7,63		-5,17					
Translation initiation factor IF-3				-2,09					
Transportin-1						-2,81			

	U3 small nucleolar ribonucleoprotein protein IMP4				-2,51				-2,16
	U3 small nucleolar RNA-associated protein 20						4,00		
	Upstream activation factor subunit UAF30		-2,47						
	UV excision repair protein RAD23					2,45	2,50		3,09
	YTH domain-containing family protein	14,13	7,07	14,16	12,33	6,33	3,68	12,60	18,93
	YTH domain-containing family protein				2,01				
Glycan Metabolism	Alpha-mannosidase				-2,14		-2,16		
	Alpha-mannosidase II	6,84			16,87				4,78
	Alpha-N-acetylglucosaminidase						2,18		
	Arabinosyltransferase				2,31				
	GPI inositol-deacylase			-2,10	-2,66		-2,47		
	Heparanase	-3,76	-4,52	-3,88	-2,80			-2,69	
	Mannosylglycoprotein endo-beta-mannosidase	4,12	3,08		2,46		2,33	2,92	
	Mannosyl-oligosaccharide alpha-1,2-mannosidase	2,51			2,24				
Lectins	Chitinase domain-containing protein	-3,49	-5,24		-2,85				-11,19
	Sorting nexin-1/2						-2,57		
Lipid Metabolism	11beta/17beta-hydroxysteroid dehydrogenase							4,14	
	12-oxophytodienoic acid reductase			-2,01					
	12-oxophytodienoic acid reductase						-3,73		
	12-oxophytodienoic acid reductase					-5,45	-8,64		
	12-oxophytodienoic acid reductase				-8,16		-9,47	-5,86	-7,06
	17beta-estradiol 17-dehydrogenase / very-long-chain 3-oxoacyl-CoA reductase				2,33				
	17beta-estradiol 17-dehydrogenase / very-long-chain 3-oxoacyl-CoA reductase				2,33				
	17beta-estradiol 17-dehydrogenase / very-long-chain 3-oxoacyl-CoA reductase			2,69					
	3-oxoacyl-[acyl-carrier protein	2,35							
	Acetyl-CoA acyltransferase 1				2,14				
	Acetyl-CoA C-acetyltransferase								2,17
	Acetyl-CoA carboxylase / biotin carboxylase 1						-2,10		
	Acyl-CoA oxidase	2,15	3,84	2,32					
	Acyl-CoA oxidase	2,15							

Acyl-CoA oxidase						-3,18		
Adrenodoxin-NADP+ reductase						4,00		
Alcohol-forming fatty acyl-CoA reductase	-2,25	-2,66		-2,42			-3,84	-4,12
Aldehyde dehydrogenase (NAD+)				2,03			2,02	
Aldehyde dehydrogenase (NAD+)							2,02	
Chloroplastic oxoene reductase							-2,34	
Chloroplastic oxoene reductase	2,16	2,15	2,39	2,79				3,16
Choline/ethanolamine kinase								-2,24
Choline-phosphate cytidylyltransferase	3,57	2,70					2,41	
Delta24-sterol reductase	-2,89	-3,10		-2,32		-2,05	-2,84	-3,42
Diacylglycerol diphosphate phosphatase / phosphatidate phosphatase						-2,04		-2,51
Enoyl-CoA hydratase		2,72		2,36				
Fatty acid alpha-dioxygenase		3,06						
Fatty acid omega-hydroxy dehydrogenase								-2,61
Fatty acyl-ACP thioesterase A						3,61		
Galactolipid galactosyltransferase				-2,91			-2,67	-3,73
Hydroperoxide lyase	-2,26	-2,22		-2,29			-2,44	-3,29
Leukotriene-A4 hydrolase				-2,12		-2,00		
Linoleate 9S-lipoxygenase	-4,18							
Linoleate 9S-lipoxygenase	-8,97	-8,26		-54,51	-13,60	-5,61		
Lipoxygenase	2,24	5,00		2,11				
Lipoxygenase		-2,97		-2,21				
Lipoxygenase	-2,96	-3,38		-2,78	-2,06	-2,14		-2,32
Long-chain acyl-CoA synthetase			-2,04					
Long-chain acyl-CoA synthetase		2,37				-3,24		
Long-chain acyl-CoA synthetase						-3,06	-3,20	-3,75
Long-chain acyl-CoA synthetase	-5,62	-9,03		-2,15			-8,58	-50,97
Long-chain acyl-CoA synthetase	3,00	3,00		2,97			2,48	2,58
Lysophosphatidic acid acyltransferase / lysophosphatidylinositol acyltransferase		-4,90	-4,49	-2,59				

	Peroxygenase	2,22			2,06			2,91	2,47
	Peroxygenase			-4,17					
	Peroxygenase	9,77	8,91	8,23	9,34	2,60		9,08	9,43
	Peroxygenase	4,13			3,01			11,54	4,77
	Peroxygenase	4,77	3,61	7,08	3,69	2,12		3,19	4,59
	Peroxygenase						2,14		
	Phospholipase D1/2	2,04							
	Phospholipase D1/2						-2,34		
	Plant 3beta-hydroxysteroid-4alpha-carboxylate 3-dehydrogenase	-12,00	-9,59						
	Sterol 14alpha-demethylase				2,14				
	Sterol 24-C-methyltransferase								-2,18
Membrane Trafficking	Arsenite/tail-anchored protein-transporting ATpase								-3,16
	Arsenite/tail-anchored protein-transporting ATpase				2,28				
	Clathrin coat assembly protein AP180								-2,91
	Coatomer subunit alpha								-2,25
	Coatomer subunit beta		2,91						3,02
	Coatomer subunit beta								
	Coatomer subunit delta	16,11	8,40	13,52	11,07	2,93	2,04		5,57
	E3 ubiquitin-protein ligase RNF115/126	2,61			2,27	2,73	2,48		5,70
	E3 ubiquitin-protein ligase ZNF598	2,92						2,50	2,67
	ELMO domain-containing protein	3,03							
	Endoplasmic reticulum junction formation protein lunapark	6,08	3,42	4,46	5,79	2,80	2,24	4,13	3,73
	Endoplasmic reticulum-Golgi intermediate compartment protein 3	6,88	5,22	6,04	9,11	5,08	2,96	4,27	3,43
	Endoplasmic reticulum-Golgi intermediate compartment protein 3							2,87	2,74
	ER membrane protein complex subunit 1								2,26
	ESCRT-II complex subunit VPS36							-2,12	-2,08
	Exocyst complex component 2 893		-5,07		-3,80			-2,25	-2,07
	Exocyst complex component 4		3,26						
	Exocyst complex component 7								-2,20
	Exocyst complex component 7	-3,10	-3,76		-2,22				-2,92

	Phosphatidylinositol-binding clathrin assembly protein								-2,91
	Protein transport protein SEC24	-2,20			-2,15		-2,04		-2,12
	Protein transport protein SEC61 subunit beta							3,12	5,54
	Rab5 GDP/GTP exchange factor	-2,38					-2,93		
	Ras-related C3 botulinum toxin substrate 1		-2,93						
	Ras-related protein Rab-11A						2,43		
	Ras-related protein Rab-11A	-4,06	-3,05						
	Ras-related protein Rab-2A								2,20
	Ras-related protein Rab-7A		2,09		2,32			2,35	3,60
	Receptor expression-enhancing protein 1/2/3/4		-2,18						
	TBC1 domain family member 15	2,04			2,45				
	Trafficking protein particle complex subunit 11						-24,70		
	Trafficking protein particle complex subunit 3	-2,59	-2,07						
	Vacuolar protein sorting-associated protein 13A/C	-2,06	-2,81					-2,16	-3,50
	Vacuolar protein sorting-associated protein 26A/B				2,27				
	Vacuolar protein sorting-associated protein 53		5,02						
	Vesicle transport through interaction with t-snares 1							2,07	3,85
	Vesicle transport through interaction with t-snares 1	2,50					3,08		
	Vesicle-associated membrane protein 7	5,47	3,75		5,50			7,71	5,26
	Vesicle-associated membrane protein 7		2,98		2,86		2,14		
	Vesicle-associated membrane protein 72								
	Vesicle-associated membrane protein 72		4,13		2,99				
Nucleotide Binding	ADP-ribosylation factor-like protein 8	-4,39	-5,88	-3,24	-3,56			-3,77	-5,27
	ADP-ribosylation factor-like protein 8				-3,86			2,62	2,18
	3'-phosphoadenosine 5'-phosphosulfate synthase	-2,36	-3,52		-2,06			-2,16	-3,42
	Adenosine kinase	6,93	5,34	6,97	10,47	3,55		9,64	8,62
	Adenosine kinase	-2,39							
	Adenosine kinase							9,64	
	Adenylate kinase	-4,40	-5,63		-4,73				-2,77

	Adenylate kinase									8,76
	Adenylate kinase		-2,29						-2,14	-2,63
	Adenylosuccinate lyase	-2,77	-3,79	-2,24					-2,01	-3,81
	Adenylosuccinate synthase	-4,02	-3,82		-4,00		-3,82			-3,74
	Apyrase				-4,85					
	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1/3									2,22
	UMP-CMP kinase	8,39	5,98	7,35	7,69	6,26	5,70	9,34		8,61
	Urate oxidase		3,11							
	Uridine kinase						-2,02	-2,50		-2,27
	Uridine nucleosidase		2,28							
	XTP/ditp diphosphohydrolase						2,59			
Protein Phosphatases	Nucleoredoxin									2,65
	Nucleoredoxin				2,85					
Secretion System	Signal recognition particle 43 kda protein		-2,15		-2,26					-2,50
	Signal recognition particle receptor subunit beta									2,93
	Signal recognition particle subunit SRP54									2,92
	Signal recognition particle subunit SRP54						3,50			
	Signal recognition particle subunit SRP68					2,85	2,26			
	Signal recognition particle subunit SRP72						-2,21			
Signalling And Transduction	14-3-3 protein epsilon									3,82
	Annexin D			2,62						2,07
	Annexin D						-2,13			
	Aquaporin PIP	-38,23	-13,19		-22,37		-10,03			
	Aquaporin TIP	-2,66					-2,87			
	ATP-binding cassette, subfamily G (WHITE), member 2	-2,63								
	ATP-binding cassette, subfamily G (WHITE), member 2							-4,71		
	ATP-binding cassette, subfamily G (WHITE), member 2, PDR						2,85			
	ATP-binding cassette, subfamily G (WHITE), member 2, PDR		3,61							
	ATP-binding cassette, subfamily G (WHITE), member 2, PDR		3,61							

BR-signaling kinase						-2,03			
Calcium binding protein 39		5,20							
Calcium permeable stress-gated cation channel		2,04		3,65		2,90			
Calmodulin									-2,31
Casein kinase 1									2,03
Casein kinase II subunit alpha								-2,51	-2,55
GTP-binding protein		2,73				-2,44			2,30
GTP-binding protein									-2,09
Low molecular weight protein-Tyrosine phosphatase								-2,21	-2,17
Mechanosensitive ion channel protein 1/2/3	5,40	3,56	3,91	3,98	2,17				
Mechanosensitive ion channel protein 1/2/3								-2,15	-2,88
Metal transporter CNNM	2,20	2,29		3,44				4,21	5,16
MFS transporter, PHS family, inorganic phosphate transporter	-2,71	-3,21		-2,25					
Mitochondrial pyruvate carrier 2									2,45
P-type Ca2+ transporter type 2C	-2,96								
P-type Ca2+ transporter type 2C	3,70	2,26		3,20					
Serine/Threonine-protein kinase SRK2				-2,18					
Serine/Threonine-protein phosphatase 2B regulatory subunit						3,32			
Solute carrier family 25 (mitochondrial adenine nucleotide translocator), member 4/5/6/31									2,31
Solute carrier family 25 (mitochondrial adenine nucleotide translocator), member 4/5/6/31						3,16			
Solute carrier family 25 (mitochondrial adenine nucleotide translocator), member 4/5/6/31									3,06
Solute carrier family 25 (mitochondrial adenine nucleotide translocator), member 4/5/6/31									3,06
Solute carrier family 25 (mitochondrial oxoglutarate transporter), member 11	-2,85								
Solute carrier family 25 (mitochondrial oxoglutarate transporter), member 11	-3,47							-4,54	-3,57
Solute carrier family 25 (mitochondrial uncoupling protein), member 8/9								-2,42	
Solute carrier family 35, member E1	-2,77	-2,18						-4,25	-4,40
Speckle-type POZ protein	-2,25	-2,66		-2,42				-3,84	-4,12
T-complex protein 1 subunit beta						-2,51			2,46
T-complex protein 1 subunit epsilon								-2,45	
Type II protein arginine methyltransferase	2,70			11,34					

	Ubiquinol-cytochrome c reductase cytochrome c1 subunit		3,74								
	Ubiquinol-cytochrome c reductase cytochrome c1 subunit		2,44								
	Ubiquitin carboxyl-terminal hydrolase 7	2,55									
	Ubiquitin carboxyl-terminal hydrolase 7							-2,69		-2,36	
	Ubiquitin carboxyl-terminal hydrolase L5							3,05			
	Voltage-dependent anion channel protein 2								2,00		
	Voltage-dependent anion channel protein 2									2,10	
	Voltage-dependent anion channel protein 2							2,10			
Terpenoid And Polyketides	(+)-neomenthol dehydrogenase		9,62								
	(+)-neomenthol dehydrogenase		9,62								
	(+)-neomenthol dehydrogenase		9,62								
	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase								-5,16		
	1-deoxy-D-xylulose-5-phosphate reductoisomerase				4,87						
	1-deoxy-D-xylulose-5-phosphate synthase		-9,40		-4,80				-5,68		
	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase									2,12	
	3''-deamino-3''-oxonicotianamine reductase	3,25	5,07	5,63	2,40				3,03	4,40	
	3''-deamino-3''-oxonicotianamine reductase		-2,40								
	3''-deamino-3''-oxonicotianamine reductase	-5,50	-3,75	-8,32	-6,70						
	3-dehydroquinate dehydratase / shikimate dehydrogenase	-3,05						-2,23	-2,32		
	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase				-2,18					-3,13	
	Abscisic-aldehyde oxidase	3,47	2,04		3,21						
	Abscisic-aldehyde oxidase	3,47	2,04		3,21						
	Abscisic-aldehyde oxidase	-2,62			-2,68						
	Abscisic-aldehyde oxidase	-6,34	-5,80		-4,65				-2,70	-4,04	
	All-trans-nonaprenyl-diphosphate synthase	-12,10	-16,18						-4,14	-4,13	
	Carotenoid 9,10(9',10')-cleavage dioxygenase 1							-3,02			
	Cis-zeatin O-glucosyltransferase						-3,31	-7,10			
	Diphosphomevalonate decarboxylase			-2,10							
	Geranyl diphosphate synthase							-2,74			

	Geranylgeranyl diphosphate synthase, type II						2,98	2,37	
	Geranylgeranyl diphosphate/geranylgeranyl-bacteriochlorophyllide a reductase		-3,19		-2,76				-3,22
	Isopentenyl phosphate kinase								-2,39
	Isopentenyl-diphosphate Delta-isomerase							2,12	
	NAD+-dependent farnesol dehydrogenase		3,22						
	Violaxanthin de-epoxidase	-2,32	-2,52	-2,18				-2,47	-2,92
	Zeaxanthin epoxidase		-2,79		-2,07				
	Zeta-carotene desaturase						-2,53	-2,08	
Unknown	Chromate reductase, NAD(P)H dehydrogenase (quinone)							2,90	2,57
	Diazepam-binding inhibitor (GABA receptor modulator, acyl-CoA-binding protein)							2,87	
	Diazepam-binding inhibitor (GABA receptor modulator, acyl-CoA-binding protein)						4,68		
	Phage shock protein A								-2,05
Xenobiotics Biodegradation	3-hydroxybutyrosineyl-CoA dehydrogenase				2,15				
	Cytochrome P450 family				3,48				2,33
	Cytochrome P450 family		5,76						
	Cytochrome P450 family 89 subfamily A	2,25	3,35				-5,42		
	Cytochrome P450 family 93 subfamily A	11,04	5,81					10,14	9,97
	Cytochrome P450 family 97 subfamily B polypeptide 3							-2,15	-2,88
	D-2-hydroxyglutarate dehydrogenase						-3,28		
	NADPH:quinone reductase				-3,04				