Transcriptomics studies under water-deficit stress - Towards genetic improvement of Bambara groundnut (*Vigna subterranea* (L.) Verdc.)

Faraz Khan

School of Biosciences

University of Nottingham

This dissertation is submitted for the degree of

Doctor of Philosophy

2016

I would like to dedicate this thesis to my loving parents who had always believed in me and supportive of my choices in life.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This dissertation contains fewer than 70,000 words including appendices, bibliography, footnotes, tables and equations and has less than 150 figures.

Faraz Khan

2016

Acknowledgements

First and foremost, I would like to sincerely thank my supervisor, Associate professor Sean Mayes for his relentless guidance, encouragement, support, constructive criticism throughout my study. His contribution to my study has been highly impactful. It was an extraordinary educational experience to be working with him, and I was much privileged to have such an opportunity. I would also like to thank my other supervisor Prof. Chungui Lu for his support and guidance during my study.

I gratefully acknowledge Crops For the Future (CFF) for covering my living expenses throughout the study and the University of Nottingham for awarding scholarship to fund this research.

I would also like to thank Dr Aryo Feldman, Razlin Azman Halimi, Dr Hui Hui, Dr Katie Mayes, Suresh Bonthala, Roberto Santos, and Dr Presidor Kendabie for their constant guidance and help throughout the study. A special thanks to Dr Wai Kuan for her help and support in identification of gene location in bambara groundnut.

A huge appreciation to Manuel Martinez Martinez and Steven Grundy for their help in growth room experiments, for that I am truly indebted.

I am also tremendously grateful of the help from Prof. Charlie Hodgman for his valuable advice and support in microarray analysis and i would also like to special thank Dr. Joanna Moreton for her extremely valuable work and guidance in RNA-sequencing analysis including Prof. Richard Emes, Dr. Tom Giles and Andrew Warry. I would also like to thank High-Throughput Genomics, Wellcome Trust Centre for Human Genetics for the completion of next-generation sequencing and their valuable inputs on the samples and protocols.

I am very grateful to Mark Meacham and Sue Flint for their help and support in growth room experiment. In addition the technical lessons for LiCOR instrument with the help from Jason Rayner and Ian Smillie are highly appreciated.

This study would not be completed without the help and guidance from Dr Ishan Ajmera, Dr Mohammed Cherkaoui Rbati, Ghita Sennaji, Sebastian Guerrero and Jiang Lu. I am greatly appreciative of their selfless efforts.

A special thanks to my friends with whom I have ventured through triumphs and failures during this past four years, for their help and support in and out of work.

Last but not least, to my mother, dad and sisters, who have been very loving, understanding, patient, and supportive throughout the period of my study.

Abstract

With the world population estimated to be nine billion by 2050, the need to exploit plant genetic diversity in order to increase and diversify global food supply, and minimise the over-reliance for food on a few staple crops is of the utmost importance to address food security challenges. Bambara groundnut (*Vigna subterranea* (L) Verdc.), is an underutilised legume indigenous to Africa, rich in carbohydrates, with reasonable amounts of protein. It is known to be drought tolerant, able to grow on marginal lands where other major crops cannot with minimal rainfall (<700 mm) and no chemical inputs. The present study aimed to investigate and evaluate transcriptomic changes in two bambara groundnut genotypes; DipC and TN (Tiga Nicuru), derived from landraces, in response to drought stress using microarray XSpecies and next generation RNA sequencing approaches by utilising data, resources and approaches derived from major crops and model plants. Crop improvement for abiotic stress tolerance and increasing/stabilising yield have been difficult to achieve due to the complex nature of these stresses, and the genotype x environment interaction (GxE). Using bambara groundnut as an exemplar species this study also highlights how a number of recent technologies and approaches used for major crop research, can be translated for use in the research of minor crops for a better understanding of the genetics governing drought traits.

To investigate the drought tolerance of bambara groundnut, microarray XSpecies and next generation RNA sequencing (RNA-seq) analysis was completed on leaf tissue from DipC and TN under drought and control (irrigation) conditions at different developmental stages (vegetative, reproductive and pod development). This

is the first drought experiment reported in bambara groundnut employing the RNA-seq approach. Both investigation of mild (microarray XSpecies) and relatively severe (RNA-seq) drought stress for the DipC and TN genotypes, adapted to similar environmental conditions, provided initial evidence that the two genotypes used different sets of genes to achieve drought response traits (including; ABA synthesis, hormone signaling, osmotic adjustment, accumulation of antioxidants, lignin synthesis, down-regulation of photosynthesis related genes, carbohydrate metabolism, cell-wall modification and transporters). Hence, both genotypes may have adapted in different ways to enable them to grow in the semi-arid conditions, suggesting that there may be more than a single way to achieve resilience in the face of drought stress. The key enzymes involved in metabolic pathways, such as carbohydrate metabolism, redox homeostasis, lipid metabolism, photosynthesis, generation of precursor metabolites/energy, and cell wall component biogenesis were affected by drought stress in both genotypes. XSpecies microarray experiment identified several differentially expressed genes (DEG) in each genotype and the four potential drought candidate genes (PAL1, Beta-fructofuranosidase, COMT, UBC-2) identified were validated utilising quantitative reverse transcriptase PCR (qRT-PCR).

In addition, both drought experiments (mild and severe) also showed that the two genotypes expressed a number of genes of what are classically considered to be 'drought-response' genes even under the control condition. These results suggest that high expression of drought-response genes even under control conditions in both genotypes may lead to greater root growth and other avoidance traits which prime the plant for future dry periods, hence preparing for drought conditions.

Morphological differences and the rapid reduction in photosynthesis, stomatal conductance and transpiration observed in both genotypes under drought stress provides a platform to link these physiological data with gene expression data. The observed

physiological responses (i.e reduction in stomatal conductance and photosynthesis) under drought stress were backed up by high expression of genes related to stomatal closure via ABA signaling and down-regulation of photosynthesis-associated genes.

A selection of genes chosen from microarray XSpecies and RNA-seq experiments were further used to identify their approximate chromosomal location in bambara groundnut using a cross-species approach. A total of 4 genes (HOX, AUX_IAA, acid phosphatase and dehydrin) were found to be near or within the confidence intervals of the QTLs underlying two drought traits (stomatal density/leaf area and CID). The initial results suggest that some of the locations of genes identified in XSpecies microarray and RNA-seq experiments could underlie QTL involved in controlling drought traits in bambara groundnut.

These data provide the basis for drought trait improvement in bambara groundnut, which will facilitate functional genomics studies. Analysis of this dataset have suggested that both genotypes are primed to respond to drought stress and have adapted in different ways to achieve drought tolerance. This will help in understanding the mechanisms underlying the ability of crops to produce viable yields under drought conditions. Future work should verify whether the identified genes are associated with the trait of interest.

PUBLICATIONS

1) F. Khan * [1], R.Azman [2], H.H. Chai [3], Sean Mayes [1, 2, 3], C. Lu [1] (2016). Genomic and Transcriptomic approaches towards the genetic improvement of an underutilised Crops: The case of; Bambara Groundnut. *African Crop Science Journal* (In press).

[1] School of Biosciences, University of Nottingham, UK campus, Nottingham, United Kingdom.

[2] Crops For the Future, JalanBroga, 43500 Semenyih, Selangor Darul Ehsan, Malaysia.

[3] Biotechnology Research Centre, School of Biosciences, University of Nottingham, Malaysia Campus, JalanBroga, 43500 Semenyih, Selangor Darul Ehsan, Malaysia.

*Corresponding author: Stxfk5@nottingham.ac.uk

2) Faraz Khan [1], Hui Hui Chai [3], Ishan Ajmera[1, 2], Charlie Hodgman[1, 4], Sean Mayes * [1, 2] , Chungui Lu [1](2016). Transcriptomic comparison of two Bambara groundnut landraces under drought stress. *BMC Genomics* (paper submitted).

[1]School of Biosciences, University of Nottingham, Sutton Bonington campus, United Kingdom, LE12 5RD.

[2]Crops For the Future, Jalan Broga, 43500 Semenyih, Selangor Darul Ehsan, Malaysia.

[3]Biotechnology Research Centre, School of Biosciences, University of Nottingham, Malaysia Campus, Jalan Broga, 43500 Semenyih, Selangor Darul Ehsan, Malaysia.

[4]Centre for Plant Integrative Biology, University of Nottingham, Sutton Bonington campus, LE12 5RD, UK.

*Corresponding author: sean.mayes@cffresearch.org

Contents

Supplementary data is provided on the attached disk.

List of Figures

List of Tables

Nomenclature

Acronyms / Abbreviations

ABA	Abscisic Acid
ANOVA	Analysis of Variance
BP	Biological process
CC	Cellular Component
CDF	Cell Description File
CEL	Cell Intensity Files
CPM	Counts per Million
DArT	Diversity Arrays Technology
DAS	Days After Sowing
DE	Differential expression
DW	Dry Weight
eQTL	Expression Quantitative Trait Analysis
FDR	False Discovery Rate
FW	Fresh Weight
gDNA	Genomic DNA

GEM	Gene Expression Marker
GFC	Gravimetric field capacity
MDS	Multidimensional scaling
MF	Molecular function
MM	Mismatch
NASC	Nottingham Arabidopsis Stock Centre
PAR	Photosynthetically Active Radiation
PCA	Principal Component Analysis
PM	Perfect Match
RBD	Randomized Block Design
RMA	Robust Multichip Average
ROS	Reactive Oxygen Species
TPM	Transcripts per million
TW	Turgid Weight

Chapter 1

Introduction

The present aim of the study is to evaluate the effect of drought and changes in gene expression of the two genotypes of bambara groundnut (DipC and TN) subjected to drought stress at different developmental stages of plant growth. In this study, bambara groundnut (*Vigna subterranea* (L.) Verdc) was used as an exemplar crop. In bambara groundnut, a combination of two transcriptomic approaches were employed. Firstly, a Cross-Species (XSpecies) microarray approach was applied, where the nucleic acid from bambara groundnut was cross-hybridised to the Soybean (*Glycine max*) Affymetrix GeneChip array. This approach was used to identify transcriptomic changes in two bambara groundnut genotypes DipC and TN (Tiga Nicuru), derived from a Botswanan and a Malian landrace, respectively, in response to drought. The aim of the study was to compare the transcriptome of the two genotypes of bambara groundnut and begin to understand the drought mechanisms underlying the ability of crops to produce viable yields under drought conditions. Secondly, Next Generation Sequencing (NGS) using the RNA-seq technology was applied with an aim to identify genes and expression patterns at different growth stages and time points in response to drought stress using the same two genotypes (DipC and TN) of bambara groundnut. Comparison of the two technologies were made as well (Microarray and RNA-seq). In addition, co-expression network analysis was performed on the transcriptomic data in order to identify potential gene regulators responsible for the expression of drought-

response genes. Chapter 1 introduces the background on bambara groundnut and an overview of the importance of bambara groundnut and potential contribution towards attaining food and nutritional security are described in details in chapter 2 section 2.1.1. This is followed by the description of the genotypes of bambara groundnut used for the study, the study of plant transcriptome and overview on co-expression network analysis. Lastly, project overview, aims, objectives and thesis outlines are all presented.

1.1 Bambara groundnut

1.1.1 Introduction

Bambara groundnut (*Vigna subterranea* (L.) Verdc) is an indigenous legume belonging to the family Fabaceae, subfamily Papilionoideae, that is widely grown by subsistence and small-scale farmers and cultivated mainly in sub-Saharan Africa [4]. In semi-arid Africa, it is considered to be the third most important after groundnut (*Arachis hypogaea*) and cowpea (*Vigna unguiculata*) [5]. It is a protein rich underutilised crop known to survive in areas of minimal water and capable of growing in poor soils [6], which makes bambara groundnut a potential crop for easing future global food and nutritional security issues.

It has been suggested that the centre of origin of bambara groundnut is in the regions between north eastern Nigeria and northern Cameroon, where its wild forms are found [7]. Bambara groundnut has been widely cultivated in tropical regions since the 17$^{\text{th}}$century, which includes Nigeria, Ghana and Haute Volta, Eastern Africa and Madagascar [8]. In addition, it is grown in some parts of South America, Oceania and Asia such as Sri Lanka, India, Malaysia, Indonesia and Philippines [9].

It has been reported that bambara groundnut consist of two botanical forms: wild

forms (var. *spontanea*) and domesticated forms (var. *subterranea*) [10]. Regions covering Nigeria to Sudan and Cameroon are the centre of origins for wild forms of bambara groundnut, whereas tropical areas, especially in sub-Saharan Africa are home for domesticated forms [10]. High genetic resemblance between wild and domesticated forms has been reported [11], which implies that the domesticated bambara groundnut is derived directly from the wild forms [12].

Bambara groundnut is a predominantly self-pollinating crop (cleistogamous) and has 11 pairs of chromosomes, 2n=2x=22 [13]. Bambara groundnut has a life cycle of between 110 to 150 days and the plant seeds takes around 7-10 days to germinate [14]. The flowering starts from 30-35 days after sowing and may continue till the end of crop life cycle [12]. Under a 12 hour photoperiod, bambara groundnut plants take 30-40 days to form mature pods after fertilisation [15]. The temperature for optimal growth of bambara groundnut is between 28 °C and 32 °C [16]. Bambara groundnut has a well-developed tap root and lateral roots under the soil and it grows up to a height of 30 cm-35 cm [6]. The roots form nodules in association with Rhizobia for nitrogen fixation [6]. It has trifoliate leaves with erect petiole grown from short, creeping, multi-branched lateral stems at ground level and lateral stem has numerous nodes and the distance (or the length of branch) from the base of the plant to the nearest node is always shorter than the more distant ones [6] (see figure 1.1). Due to the length of internodes, bambara groundnut landraces differ from each other in terms of growth habit, ranging through spreading, semi-bunched to bunch types [8].

The flowers of bambara groundnut are typically papilionaceous and are produced on long and hairy peduncles which elongate from nodes on the lateral stem [12]. For pod formation, peduncles elongate until their maximum length and bring the fertilised ovary into the soil or just above ground level after pollination and fertilisation [6, 15]. The size of the pods ranges from 1.5 cm to 3.7 cm in diameter depending on the

number of seeds inside the pod [6]. The pods are round, oval or spherical in shape with some pods containing only one seed and some containing two or more [11]. In terms of yield, it has been reported that when environmental factors such as the seasonal distribution of rainfall, day length and range of temperatures during the growing season are accounted for, the potential yields of bambara groundnut within its current areas of cultivation can be significantly increased without high levels of agronomic inputs. [5]

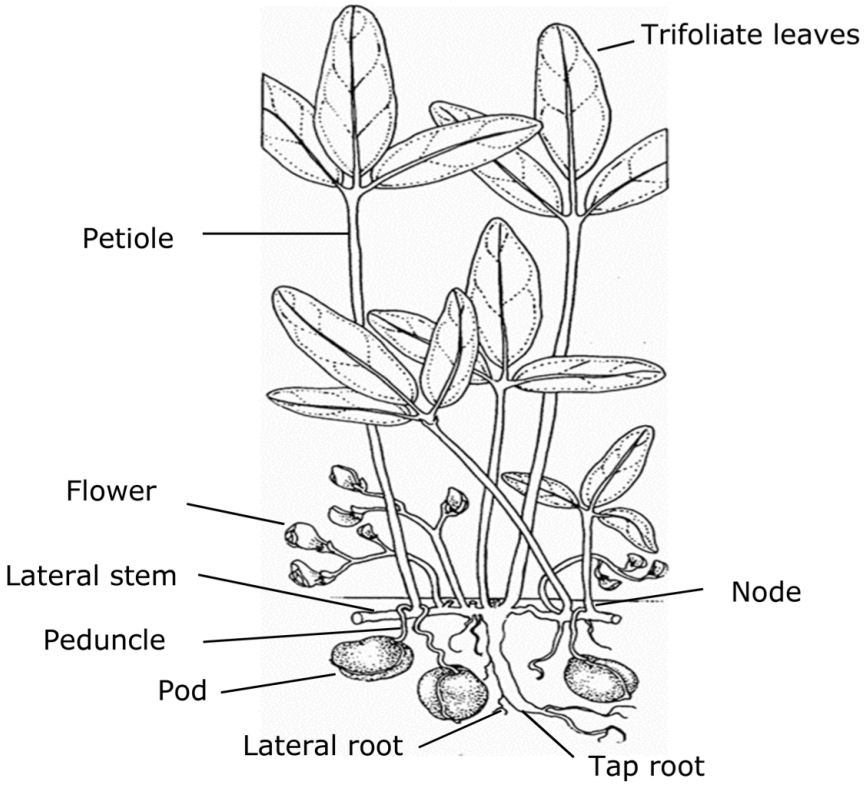

Figure 1.1: The morphology of bambara groundnut. [1]

1.1.2 Importance of bambara groundnut for food and nutritional security

It is predicted that the world population will increase to 9 billion by the year 2050 [17], hence the food supply must double in order to cope with the ever increasing world population coupled with the likely yield reducing effects of climate change [18]. Currently only 20 plant species comprise 90% of the world's calories [18]. The need for crop diversification is important in order to increase food supply, improve nutrition and avoid complete dependence on a limited number of plant which could be vulnerable to pests, diseases and have largely been bred for high input agriculture. That implies that the food basket should contain more underutilised crops, particularly those which are known to have resilience traits, with the ability to withstand drought, flooding, temperature extremes, and pests and diseases to a greater extent than current major crops [19]

Bambara groundnut is considered as a drought tolerant crop with reasonable protein content (18 to 22%), high carbohydrate (65%) and low levels of fats (6.5%) [20]. Chapter 2 section 2.1.1 gives a detailed description of bambara groundnut profiles. As an underutilised crop, there has been limited research done on bambara groundnut with only limited genomics resources currently in existence [12]. However, due to bambara groundnut containing desirable traits such as tolerance to various biotic and abiotic stresses and reasonable protein content it is a target crop for further research and a potential alternative crop for food production [12].

1.1.3 Bambara groundnut genotypes used for the study

The Institute of Tropical Agriculture (IITA) in Nigeria and Southern Africa Development Community (SADC) currently holds largest number of bambara groundnut

accessions. Approximately 2000 from Nigeria and 972 accessions from SADC [21]. In addition, genetic resources of bambara groundnut are widely conserved by indigenous farmers across sub-Saharan Africa. However, the germplasm for bambara groundnut has not been developed by conventional breeding yet and most of its germplasm exist in the form of landraces (as discussed in 2.1.2). Landraces have high genetic variation for breeding traits, which provide breeders with an opportunity to study genetic resources to improve yield, abiotic resistance and adaptability of crops to adverse environments.

In this study, the two genotypes derived from the landraces of bambara groundnut; DipC and Tiga Nicuru (TN) were used to study effect of drought stress in bambara groundnut. DipC and TN belongs to Botswana and Mali respectively. Both Botswana and Mali are known to be semi-arid, landlocked countries in the centre of Southern and West Africa, respectively. The annual rainfall in Botswana ranges from to 250 mm in the extreme southwest to 650mm in the extreme northeast, with a mean annual rainfall of 450 mm [22]. While for Mali, the annual rainfall varies across the country. Mali can be divided into into three climatic zones. The highest mean rainfall of between 700-1000 mm is observed in Sudanic in the South, 200-400 mm rainfall in the Sahelian in the central and little or zero rainfall in the Saharan in the North, with an average of annual rainfall of 440 mm across the country [23]. The temperature in Botswana ranges from 30°C to 40°C by late afternoon in the dry season (April to October) and 25°C to 30°C during the rainy season (November to March, [22]). While for Mali, it ranges from 16°C to 39°C with 4-5 months of rainy season (from April to October, [23]). As a result, both Botswana and Mali are effected by periods of drought, desertification and in some regions receives very minimal to negligible rainfall. The climatic conditions in their regions of origin (Botswana and Mali respectively), suggest that they are likely to be adapted to be more tolerant to drought than many other

landraces, while potentially having some variation, as they are morphologically and phenotypically distinct [12].

The contrast between the two genotypes for a number of traits such as days-to-maturity, stomatal conductance, 100-seed weight, leaf area, internode length, peduncle length, pod number per plant, and leaf carbon (Delta C13) isotope analysis, suggest that some of these mechanisms for adaption to drought could be non-identical in the two genotypes. For example, Delta C^{13} was associated with higher yield as observed in DipC, compared to TN [2, 12, 24]. DipC has a greater petiole length, leaf area and plant height and shorter internode length and peduncle length compared to TN [2]. DipC is classified as bunched type while Tiga Nicuru is categorised as a semi-spreading (see figure 1.2). Genetic diversity UPGMA (Unweighted Pair Group Method with Arithmetic Mean) analysis based on DArT (Diversity Arrays Technology) array markers signifies genetic differences between the two genotypes by allocating them in separate clusters (Figure 1.3) [25].

Figure 1.2: Plant architecture comparison between DipC (bunched type) and TN (semi-spreading) [2].

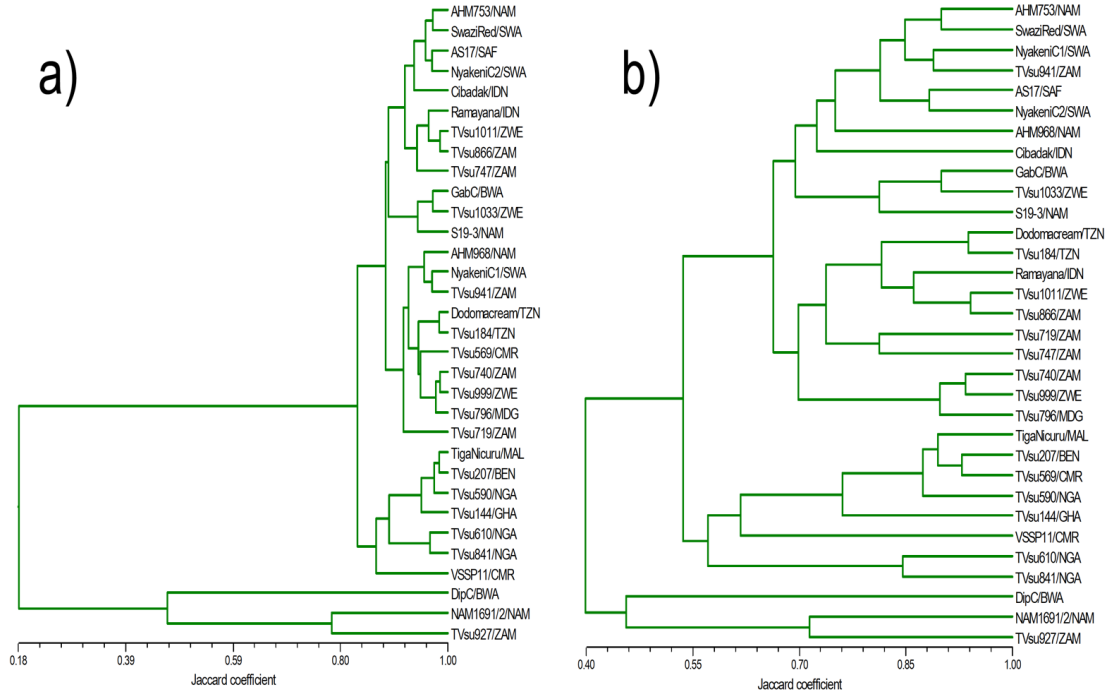

Figure 1.3: UPGMA dendrograms representing 32 bambara groundnut genotypes based on the similarity matrix of a) 76 polymorphic DArT markers (with redundancy) and b) 26 unique DArT markers from the initial PstI/AluI array[3].

Chai et al. [12, 24], reported that transgressive segregation was observed in the segregating F_5 population derived from the TN X DipC cross. The results showed that there were lines in the segregating population that performed better in terms of the ability to produce higher yields under drought conditions than the parental genotypes. Hence, evaluating the transcriptome of the two parental lines under drought stress could be a good indicator to investigate the molecular mechanism occurring in the two genotypes and its relationship to phenology and phenotype.

1.2 The study of plant transcriptome

Transcriptomic methods are used to study the expression pattern of genes under certain conditions, at specific points in development and in particular tissues [26]. Tran-

scriptomic technologies include Microarray chips (used in this thesis) and next generation sequencing (used in this thesis) [27, 28]. The transcriptome is the total set of transcripts produced from a sample. A transcriptome is bound to change with altered environmental conditions which reflects the expression of genes at certain time points and conditions for a particular tissue [29].

1.2.1 XSpecies Microarray

Microarrays have been considered to be an useful tool for studying gene expression under certain conditions which can aid understanding of biological systems and gene regulation at the transcription level [30]. Microarray has widely used in different transcriptomic studies because of the rapid production of data and the potentially complete coverage of a entire transcriptome on a single array, which exists for many species [29]. Microarrays consist of a defined array of millions of oligonucleotides that hybridise with complementary cRNA sequences, although the specific format is dependent on the system used (largely Affymetrix or Agilent) [29]. The oligonucleotides are printed or in situ synthesized for features/probes and the levels of hybridisation are detected and quantified using a fluorophore incorporated into copy RNA. There are two types of microarray which has been established: (1) spotted arrays which are produced by depositing and spotting the probes (cDNA, PCR products and oligonucleotides) onto the array surface and (2) oligonucleotide in situ arrays which are generated by direct synthesis of the probes onto the arrays, such as the Affymetrix GeneChip arrays which consist of short oligonucleotide sequences (25-mer probes) combined into probe-sets which are designed to interrogate gene models from the species they are designed to [30].

To study the transcriptome of species whose microarray chip is not available, one approach developed for the Affymetrix GeneChip platform is known as the XSpecies

(cross-species) microarray (http://affy.arabidopsis.info/xspecies/) that has been widely used for species without a microarray chip [31, 32]. In this approach pre-existing homologous sequences of the species of interest that are conserved within related phylogenetic groups are used to determine the putative sequences and identities of an unknown species. This is achieved by comparison of hybridisation strength in a cross-species hybridisation reaction, with at the genomic DNA level or the RNA level. This approach has been proven to be successful for studies in target species without a microarray chip such as Woodchuck (*Marmota monax*) [33], Chinese hamster (*Cricetulus griseus*) [34], Potato (*Solanum tuberosum*) [35], Banana (*Musa* spp.) [36] and Cowpea [37].

GeneChip arrays consist of up to 16 probe-pairs per probe-set, which is in contrast to other arrays which generally use single cDNA (many spotted arrays) or long oligonucleotides to assay a gene (such as Agilent Chips), with each probe-set in the GeneChip array being specific to a gene transcript. Each probe pair consists of a perfect match (PM) and a mismatch (MM) probe. PM probe have 25 nucleotides complementary to the design sequence while the MM probe is the same sequence as the PM probe except for a mismatch at the 13^{th} nucleotide in order to evaluate non-specific hybridisation [38].

In XSpecies, nucleic acid (DNA or RNA) of target species is extracted, followed by hybridisation of fluorescence-labeled or biotin-labeled nucleic acids onto microarrays designed for the other species [31]. After hybridisation the probe-sets that show good hybridisation are deemed to be complementary to the heterologous species genes and computational analysis involving the creation of a software mask is used, followed by the analysis of the pattern of hybridisation of samples to selected probes for gene expression studies. Often, an initial genomic DNA hybridisation of the heterologous species total genomic DNA is used to identify the probe-pairs which show good hy-

bridisation signal and these are retained for future analysis after the creation of a software mask for the GeneChip, leading to a custom Chip Description File (CDF) file.

Xspecies approaches have a number of limitations, including inefficient hybridisation of certain transcripts to the probes on the array due to sequence polymorphisms between two different species. This would be expected to lead to a decrease in the detection of transcript abundance, particularly as the signal strength is averaged across the individual oligonucleotides which compose the probe-set representing a specific gene model from the heterologous species [31]. One way to overcome this problem is the application of genomic DNA-based (gDNA-based) probe selection [31]. In this approach, genomic DNA of the target species is hybridised to the Affymetrix Soybean GeneChip array of another species and PM probes which show good hybridisation signal with the heterologous gDNA above a user or algorithm defined threshold are selected for subsequent transcriptome analysis, with the other probes masked out [31]. This probe selection approach was chosen to be the preferred approach over RNA based probe selection because RNA-based probe-selection uses probe masking to exclude probes which have low hybridising intensity values. However, it is very dependent on transcript abundance, so highly abundant transcripts can give high signal, even when they only hybridise very poorly to the target. The gDNA-based probe selection technique avoids bias due to transcript abundance [31]. This approach has been successful in studying *Brassica oleracea* to phosphorus (P) stress and showed 13-fold increase in the sensitivity of Arabidopsis ATH1 GeneChip when used to detect gene expression in Brassica RNA samples [31].

After the hybridisation a programme called Microarray Analysis Suite is commonly used to generate .CEL files through the scanning of the intensities for each probe. This is followed by differential expression analysis using software such as GeneSpring [39]

with Robust Multichip Average (RMA) normalisation algorithm [40]. Because of the nature of fluorescence-based detection-systems, a normalisation step is necessary to reduce sources of errors such as dye bias, scanning conditions, heat and light sensitivity, the efficiency of dye incorporation, differences in the cDNA hybridisation conditions and unequal quantities of starting RNA [31]. Individual probe-set are also replicated across the GeneChip, allowing the effects of differential hybridisation across the Chip to be accounted for. For the creation of the software mask after genomic DNA hybridisation, a parser script written in Perl was developed to generate probe-masking files which are used in the selection of probe pairs with gDNA hybridisation intensity greater than a defined threshold, leading to a custom CDF [31]. Once the CDF file is created, it is imported into GeneSpring to interpret RNA CEL files that are generated from the target species with defined threshold for differential expression analysis.

1.2.2 Next Generation Sequencing (NGS)

Recently a series of new technologies have emerged for gene expression profiling such as next generation technologies which are high-throughput methods for sequencing cDNA based upon a samples RNA content [41].

With the availability of Next Generation Sequencing (NGS) technologies, it is now possible to study the transcriptome of a species whose genome has not been sequenced yet [42]. Traditionally, model species are the main source of transcriptomic studies mainly because of the extensive expression and often their available genome information. However, in the current "-omics" era, a much greater variety of organisms can be studied at the genomic and transcriptomic level which will greatly aid in elucidating regulatory networks [43] and the understanding the genetic basis of complex traits [44] in non-model species (such as bambara groundnut).

RNA-seq is one important NGS approach which aids in studying the transcriptome

of a species of interest. A transcriptome consists of all expressed sequences, which is a reduced representation of the entire coding capacity of the genome [42]. RNA-seq can be used to study transcript structures (such as alternative splicing), allelic information (e.g. SNPs) and expression with high resolution and a large dynamic range [45]. RNA-seq is a powerful tool which has been the widely used mainly because: (1) of its ability to capture the expression of (ideally) all genes expressed under the specific experimental conditions (2) it does not require any prior genetic information, so is well suited for non-model organisms and (3) it has become cost-effective and affordable [42]. These advances greatly facilitate functional genomics research in non-model species, which are of substantial ecological or evolutionary importance.

To successfully assemble a transcriptome without the aid of a reference genome requires robust computational methods. This is achieved via *de novo* transcriptome assembly [41], which is used to assemble short RNA-Seq reads without a reference genome. It has been widely used for transcriptomic studies in organisms whose genome information is not available [46]. However, the process of assembling a transcriptome violates many of the assumptions of assemblers written for genomic DNA. For instance, uniform coverage and the 'one locus – one contig' paradigm are not valid for RNA: an accurate transcriptome assembler will produce one contig per distinct transcript (isoform) rather than per locus, and different transcripts will have different coverage, reflecting their different expression levels. Reconstruction of full-length transcripts from short reads accounts for considerable sequencing errors which poses substantial computational challenges [41]. For example, (1) some transcripts have low coverage, whereas others are highly expressed; (2) due to sequencing biases, read coverage may be uneven across the transcript's length; (3) reads with sequencing errors derived from a highly expressed transcript may be more abundant than correct reads from a lowly expressed transcript; (4) overlapping between transcripts encoded by adjacent loci

could lead to erroneous fusion to form a chimeric transcript; (5) due to alternative splicing, data structures need to accommodate multiple transcripts per locus and (6) the rise in ambiguity due to sequences that are repeated or are not unique between different genes. A method used for *de novo* transcriptome assembly should address each challenge in order to reconstruct robust full-length transcripts of variable sizes, expression levels and protein-coding capacity. [46].

There are several advantages of RNA-seq over microarray technology : (1) Unlike microarrays, RNA-seq doesn't require prior knowledge of transcript structures or genome sequences [47]; (2) it is far superior in detecting low abundance transcripts, differentiating biologically critical isoforms, and allowing the identification of genetic variants and (3) it allows broader dynamic range that leads to detection of more differentially expressed genes with higher fold-change [48]. However, despite these several advantages of RNA-seq, microarrays are still more common among researchers for transcript profiling largely due to the currently cheaper cost of microarray compared to RNA-seq and complexity of data storage and analysis [48].

1.2.3 Gene co-expression network analysis

The availability of large expression datasets has made it possible to study gene functions on a global scale. Genes that belongs to the same pathway often exhibit similar expression patterns under different conditions [49]. One approach applied in microarray and NGS technologies is to cluster the genes according to their expression patterns which can further be analysed in several ways, such as promoter analysis or gene ontology analysis [50]. However cluster analysis usually ignores the detailed relationships among genes which makes it difficult to get biological insight from the analysis [49]. Alternatively, gene co-expression network analysis has become quite popular in studying gene interactions and have become a rapidly developing area of study from which

many interesting results have been obtained [51–53].

A gene co-expression network is an undirected graph, where graph nodes are the genes, with edges connecting to each node representing a significant co-expression relationship [49]. Gene co-expression network analysis works on a value based (rank-based) network analysis that utilise the rank-transformed similarities [54]. Rank-based approaches works on the principle that two genes are connected only if the similarity between their expression profiles is above a certain threshold. Similarities can be measured by Spearman's correlation or other metrics [49]. Thresholding of the correlation preserves the continuous nature of the gene co-expression information, and leads to results that are highly robust [51]. This method works well for capturing strongly co-expressed genes [49]. Value based (rank-based) approaches have been applied to identify co-expression network in several studies [51, 53, 55, 56]. High performance web server such as DeGNserver works on the principle of rank based co-expression analysis to decipher genome-scale network [57]. The server makes use of a computer cluster to run a number of network inference algorithms and return the results to the user very quickly, thus facilitating genome-scale network reconstruction.

Several software tools have been designed for graph visualisation and labelling of nodes and edges with attributes. These include open source programs such as Cytoscape and Graphviz [58–60]. Cytoscape is generally considered to be better for the analysis of transcriptomic data as it allows direct integration and it supports user-created plugins [58]. In addition Cytoscape allows the user to analyse the network and can be used to investigate graphical features (shortest path length between two nodes and it also calculates vertex degrees that indicate the number of edges connected to a node. A graph (in the biological context) can be referred to as a pathway or network depending upon the number of nodes and their vertex degrees.

1.3 Project overview, aims and objectives

This project aims to study the transcriptome of two genotypes of bambara groundnut (DipC and Tiga Nicuru) under drought stress. Firstly this study will give a detailed description and hypothesis on how genomic and transcriptomic methodologies developed for major crops can be applied to bambara groundnut. Furthermore, in this present study a combination of XSpecies and RNA-seq approaches were employed to evaluate the bambara groundnut transcriptome under drought stress. Firstly, the two genotypes of bambara groundnut were subjected to a mild drought conditions in a controlled glasshouse. RNA from the two genotypes was cross hybridised with the Soybean Affymetrix GeneChip array to study expression profiling under drought stress at a later stage of plant development. Secondly, the two genotypes were subjected to different regimes of drought stress in growth chambers, allowing a more severe drought stress to be applied in a more controlled environment. From this study, RNA from the two genotypes was sampled at the vegetative and reproductive stages of plant development, allowing to study early gene expression responses under drought stress. In addition, morpho-physiological effects of drought stress on the two genotypes were studied and compared with peanut (*Arachis hypogaea*) in relation to drought tolerance. By keeping peanut as a comparator species, this study will help in developing an understanding of the nature of the drought tolerance level of bambara groundnut compared to a well-studied legume. Thirdly, an initial attempt is made to identify the chromosomal location of drought-related genes (identified in both genotypes from microarray and RNA-seq methodologies) by first finding the gene location in a species belonging to the same legume family (common bean (*Phaseolus vulgaris*)) where a full genome sequence exists and then overlaying the gene location onto bambara groundnut based on approximate positions identified in common bean and the map positions

in the controlled cross between TN and DipC. This allowed the gene position detected by mapping analysis to be compared with previously generated quantitative trait loci (QTL) data (Chai *et al*, paper submitted) from the F_5 segregating population derived from the cross between DipC and Tiga Nicuru [24].

The objectives/aims of the study are:

- To provide a framework on how genomic and transcriptomic methodologies developed for major crops can be applied to bambara groundnut. This will help in better understanding of the genetics governing important agronomic traits in bambara groundnut.

- To evaluate the effect of drought and changes in gene expression of the two genotypes of bambara groundnut (DipC and TN) subjected to drought stress at different developmental stages of plant growth. Identified genes and expression patterns will assist in understanding the mechanisms underlying drought tolerance in bambara groundnut which enables it to grow under semi-arid conditions.

- Comparing the transcriptome of the two genotypes (DipC and TN) of bambara groundnut under drought stress to identify what is common and how they differ.

- Comparing morpho-physiological effects of drought stress between bambara groundnut and peanut. This will help develop an understanding of the nature of the drought tolerance level of bambara groundnut compared to a well-studied legume.

- Localisation of drought-related genes (identified from microarray XSpecies and RNA-seq approaches) underlying QTL involved in controlling drought traits in bambara groundnut. Thus facilitating the marker-assisted selection process.

Thesis Outline

In the light of the project aims and objectives listed above, the work in this thesis is divided into 8 chapters:

Chapter 1: Introduction - Background of bambara groundnut and overview of the transcriptomic study.

Chapter 2: Literature Review - Detailed description on how genomic and transcriptomic methodologies developed for major crops can be applied to bambara groundnut and an outline of the effects of drought stress on bambara groundnut and other species.

Chapter 3: Materials, Methods and Data - Outlining the methods, materials, data sources and software tools used throughout this study.

Chapter 4: Bambara groundnut microarray XSpecies analysis - Analysis of the transcriptomic dataset to look for genes that are differentially expressed between irrigated (control) and drought stress conditions in two genotypes of bambara groundnut (DipC and TN) using XSpecies microarray technology.

Chapter 5: Morpho-physiological effects of drought stress on bambara groundnut at different developmental stages - Analysis of the effects of drought stress at different developmental stages on the two genotypes of bambara groundnut (DipC and TN) at morpho-physiological level and comparison of bambara groundnut with peanut (*Arachis hypogaea*) in relation to drought tolerance.

Chapter 6: Effects of drought stress on the transcriptome of bambara groundnut at different developmental stages - RNA-seq approach employed to the study the effect of drought stress on the two genotypes of bambara groundnut (DipC and TN) at different developmental stages

Chapter 7: Identification of cross species syntenic locations of drought-related genes - An attempt was made to identify the approximate chromosomal location of drought-related genes (identified in both genotypes (DipC and TN) from microarray and RNA-seq methodologies) by first finding a gene location in species belonging to the same legume family (common bean) and then overlay the gene loca-

tion onto a bambara groundnut genetic map based on approximate positions identified in common bean and lastly to identify whether there is any overlap with the QTLs already known.

Chapter 8: Conclusions - An overview of the findings and suggested future works.

Chapter 2

Literature Review

This chapter outlines how a number of recent technologies and approaches used for major crop research, can be translated into use in research of minor crops, using bambara groundnut as an exemplar species. Using drought tolerance as a trait of interest in this crop, this chapter will demonstrate how limitations can affect genomic approaches for understanding traits in bambara groundnut, and, how genomic and transcriptomic methodologies developed for major crops can be applied to underutilised crops for better understanding of the genetics governing important agronomic traits.

2.1 Introduction

The dependence of global food security on major crops is a major concern in the future for food supply and also for rural income, as yield gains from these major crops may not be enough to sustain the estimated nine billion people on the planet by 2050 [17]. Also, the danger presented by climate change, leading to increased drought, temperature, flooding, and salinisation, along with a predicted increase in pests and diseases, could drastically effect major crops growth and development. There is need to widen the exploitation of the available plant genetic diversity in order to increase food supply and avoid dependence on a limited number of plant species for global food and nutritional security. Underutilised crop could be a solution for more diversified agricultural systems, a rich source to explore novel trait values and additional food sources necessary to address food and nutritional security concerns [19, 61, 62].

With the availability of technologies, such as Next Generation Sequencing (NGS), it is possible to develop molecular markers for marker assisted selection (MAS) in underutilised crops [63]. This technology is able to generate significant sequence datasets, and allows in-depth comparisons to be made between underutilised crops and their major staple crop cousins [19]. For underutilised crops, their low and erratic yields can be due to the lack of genetic improvement and formal breeding programmes which may be hindering their wider cultivation and utilisation [64].

Several efforts have been made to conserve the germplasm of major crop species, but a more limited amount of effort has gone into conservation of underutilised crops such as bambara groundnut. The Consultative Group for International Agricultural Research (CGIAR) and Global Diversity Crops Trust are organisations whose focus is to develop a sustainability plan in order to ensure germplasm conservation and availability of underutilised crops, as a means to increase crop diversity [65]. Conserving the

genetic resources of underutilised crops exemplified by bambara groundnut, is essential as these crops are sources of livelihood in rural poor communities and its cultivation has a positive impact on farmer welfare [66]. However, conservation is only one aspect as without 'conservation into use', no impact is made on the lives of the farmers who could benefit. *Ex situ* and *in situ* germplasm conservation, with farmer participatory breeding and, identification of favourable traits, offer a productive solution to conserve and utilise the genetic resources of underutilised species. This approach can then be coupled with trait analysis to identify crops which can (i) survive in extreme conditions (e.g., drought and cold), (ii) have superior nutritional content, (iii) and have the potential to achieve high market value in order to increase their utilisation and consumption, contributing to the global food basket (iv) are acceptable to farmers.

This is the first review on bambara groundnut with a specific focus on the genomic/transcriptomic approaches available to address drought through genetic improvement. This review also highlights how some recent technologies and approaches used for major crop research, can be translated into use in research of minor crops, such as bambara groundnut. Previous reviews on genetic improvement of bambara groundnut using resources from major species has focused mainly on the overview breeding objectives and aims [**?**]. Other reviews published on bambara groundnut focus mainly on (i) conservation and improvement [6]; (ii) production [67]; (iii) developing the potential of the crop [5]; (iv) aspects of the commodity marketing of bambara groundnut [68]; (v) breeding approaches towards the genetic improvement of bambara groundnut and (vi) genetic diversity analysis of bambara groundnut [69**?**].

2.1.1 Bambara groundnut profiles

Bambara groundnut is probably one of the most drought-tolerant of the major grain legumes and may be found surviving successfully where annual rainfall is below 500

mm and has an optimum between 900–1000 mm per year [70], although it can grow in wetter conditions as long as the roots are not water-logged. It is cultivated mainly in sub-Saharan Africa, expanding to regions of West Africa, across Central Africa to East Africa and Southern Africa [4].

It is cultivated principally by farmers as a food security culture crop, because of its agronomic values and the ability to produce yield in soils considered insufficiently fertile for cultivation of other more favoured species such as common bean (*Phaseolus vulgaris*) and groundnut (*Arachis hypogaea*) [71], although there are markets and there is some early evidence that it could improve household incomes [66].

Bambara groundnut has reasonable protein content (18 to 22%), high carbohydrate (65%) and low levels of fats (6.5%), having a composition quite similar to chickpea (*Cicer arietinum*) and makes it a 'complete food' [20] (Table 2.1). The seed commands a relatively high market price in many African countries [72]. In countries such as Malawi, demand for bambara groundnut often exceeds supply [67]. Bambara groundnut is considered to be the third most valuable legume in Africa, after cowpea (*Vigna unguiculata* (L.) *Walp.*) and groundnut [73]. It is known to be deficient in sulphur-containing amino acids [5], but rich in lysine, leucine and glutamic acid, which makes a good complement to cereal-derived amino acids [20] (Table 2.2); hence in many African countries, bambara groundnut seeds are often milled and added to wheat flour and used to make a number of baked products [74]. Although, correlation between genotypes, seed colour, growing conditions and nutritional analysis has never been attempted. The seed is a useful ingredient in cooking as it can be eaten as a boiled or fried snack, and milled into flour [75]. Despite its 'balanced' macronutrient composition, bambara groundnut contains some anti-nutritional factors such as tannins and trypsin inhibitors [76]. A study conducted by Ijarotimi et al. [77], showed that fermentation has significant effects in decreasing the anti-nutritional factors; oxalate,

Table 2.1: **Macronutrient status of bambara groundnut in comparison with some more popular legumes**

Nutritional values	Bambara Groundnut	Soybean	Phaseolus Bean	Cowpea	Faba Bean	Chickpea
Calories	390	416	343	333	341	364
Protein (%)	21.8	36.5	23.8	23.6	26.1	19.3
Carbohydrate (%)	61.9	30.2	59.6	60	58.3	60.6
Fat (%)	6.6	19.9	2.1	0.8	5.7	6

Source: [68]

tannic acid, phytic and trypsin.

Table 2.2: **Amino acid content (mg/100 gm) of raw bambara groundnut**

Amino Acid	Raw Bambara Groundnut (mg/100 gm)
Lysine	2.8
Histidine	2.4
Arginine	4.9
Aspartic Acid	5.6
Threonine	2.6
Glutamic Acid	17
Glycine	3.3
Alanine	3.9
Cystine	0.7
Methionine	2.7
Isoleucine	3.9
Leucine	6.9
Tyrosine	3.4
Phenylalanine	4.8

**Values are mean of duplicate samples.* Source:[20]

2.1.2 Challenges towards research and development in bambara groundnut

Being classed as an underutilised crop, bambara groundnut faces several challenges towards its research and development. Bambara groundnut is still an underutilised crop mainly because (i) it currently has limited economic potential outside its areas of cultivation [5]; (ii) lack of appropriate processing techniques to overcome hard-to-cook effects [20]; (iii) absence of functioning value chains [68]; (iv) there is very little information and knowledge base on neglected plant species in terms genome information and germplasm collection [5]; (vi) biological issues such as photoperiod sensitivity to reproductive development and pod-filling, which effects the geographical range/time of planting and yield stability; and (vii) a lack of mechanisation for crop mechanisation (e.g., seed sowing) as well as machinery for post-harvest (e.g., pod-shelling). Furthermore, bambara groundnut is still grown as landraces and its yield can be unstable and unpredictable at different geographical regions. While being adapted to their current environment, landraces may not contain the optimal combination of traits [21]. Development of improved varieties of bambara groundnut, through controlled and coordinated multi-locational breeding programmes, is vital to harness the potential of the crop [**?**]. The International Institute for Tropical Agriculture (IITA) currently holds the largest ex situ collection, with 2055 accessions (as of January 2016) of bambara groundnut. This crop could be a potential exemplar for other underutilised crops.

2.1.3 Breeding/Molecular perspectives

Historically, genomic and molecular genetic analysis has been focused towards major species, but now with the advancement in high-throughput sequencing technologies,

such as NGS, and the reduction in their costs and also the availability of bioinformatic tools (Table 2.3), it is becoming possible to transfer information to crops and related-species. With the availability of genomic resources and the completion of reference genome sequences of legume crops, such as *Medicago truncatula* [78], common bean [79], soybean (*Glycine max*) [80] and cowpea [81], it is now possible to dissect information and transfer genomic and transcriptomic data to other legume crops such as bambara groundnut. For example, the USEARCH sequence analysis tool can be used for comparing sequences between closely related species [82], enabling advances in genetic marker development, location of orthologues and decoding of the genetic mechanism and pathways involved in drought tolerance in less studied crops, such as bambara groundnut through genomic and transcriptomic comparative analysis.

Table 2.3: **Bioinformatic tools and databases available for genomic/transcriptomic research**

Name	Description	URL/Source
Tools		
Genome Workbench	An integrated application for viewing and analyzing sequence data.	ncbi.nlm.nih.gov/tools/gbench/
MISA	Allows identification and localization of perfect microsatellites	pgrc.ipk-gatersleben.de/misa/
Trinity	Tool for RNA-seq de novo assembly	github.com/trinityrnaseq/trinityrnaseq/wiki
ArrayExpressHTS	R-based pipeline for RNA-seq data analysis	bioconductor.org/packages/devel/bioc/html/ArrayExpressHTS
USEARCH	Unique sequence analysis tool	drive5.com/usearch/
Assembly to Assembly Comparison (ATAC)	For comparative mapping between two genome assemblies, or between two different genomes	seqanswers.com/wiki/ATAC
Atlas-SNP2	SNP detection tool developed for RNA-seq platforms	sourceforge.net/p/atlas2/wiki/Atlas-SNP/
BLAST Ring Image Generator (BRIG)	Used for comparative analysis between large number of genomes	brig.sourceforge.net/
EdgeR	R package for differential expression analysis for RNA-seq data	bioconductor.org/packages/release/bioc/html/edgeR.html
solQTL	Tool for analysis and visualization of quantitative trait loci (QTL).	solgenomics.net/search/phenotypes/qtl
TriClust	Tool for cross-species analysis of gene regulation	baskent.edu.tr/~hogul/triclust/
Databases		
ArrayStar	Database that holds collection tools for microarray analysis	arraystar.com/
Bedtools	Database for comparing large set of genomic features	bedtools.readthedocs.io/en/latest/
Bionumerics	Integrated analysis of all major applications in Bioinformatics: 1D electrophoresis gels, all kinds of chromatographic and spectrometric profiles, phenotype characters, microarrays, and sequences.	applied-maths.com/bionumerics
Phytozome	Comparative platform for green plant genomics	phytozome.jgi.doe.gov/pz/portal.html

Ensembl Genomes	Genome-scale data for plants, bacteria, fungi, protists and invertebrate metazoa.	ensemblgenomes.org/
Legume Information System (LIS)	Genomics database for the legume family	legumeinfo.org/
Legume Federation	A consortium of scientists working to support robust agriculture	legumefederation.org/
LegumeIP	Integrative platform to study gene function and gene evolution in legumes	plantgrn.noble.org/LegumeIP/
Gene Expression Omnibus (GEO) BLAST	Tool for aligning a query sequence (nucleotide or protein) to GeneBank sequences included on microarray or serial analysis of gene expression (SAGE) platforms in the GEO database	ncbi.nlm.nih.gov/geo/

2.1.4 Drought stress in bambara groundnut

2.1.4.1 Drought stress tolerance mechanisms

Drought stress is one of the major abiotic stresses that inhibits proper plant growth and crop productivity. Drought stress is defined as stress that is caused by inadequate soil moisture to meet the needs of a particular crop at a particular time [83]. Drought tolerant plants respond to drought stress via a series of different mechanisms. They are divided into three groups namely (i) drought escape, (ii) drought avoidance, and (iii) drought tolerance. Drought escape is described as the ability of plants to complete their growth cycle and reach maturity before drought-stress develops to damaging levels [84]. Drought avoidance is demonstrated by crop species, which are able to maintain high water potential in the plant by minimising water loss and maximising water uptake under drought conditions, as seen in siratro (*Macroptiliumatro purpureum*), a tropical legume [85] and chickpea [86]. Mechanisms of avoidance in-

clude improved root traits, for greater extraction of soil moisture, decreased stomatal conductance, decreased radiation absorption and decreased leaf area for minimal water loss [87]. Drought tolerance allows plants to survive through water-use efficiency, i.e., performing all biological, molecular and cellular functions with minimal water. Such mechanisms are seen in a range of leguminous species, including mung bean (*Vigna radiata*) [88] and pigeon pea (*Cajanus cajan*) [89]. Plants with drought tolerance mechanisms are able to maintain their cell turgor through osmotic adjustment, which in turn will contribute to maintaining stomatal opening, leaf expansion and photosynthesis throughout the drought period [90].

Reduced water availability causes the production of abscisic acid (ABA) , the phyto-hormone which initiates stomatal closure and influences other aspects of plant growth and physiology. It is responsible for regulating a broad range of genes during drought. Sucrose Non-Fermenting1 *(SNF1)* protein kinase, Protein Phosphatase 2C (PP2C) family, *AREB* (ABA-responsive element)/ABF transcription factor are the key regulators of ABA signaling [91]. Other hormones, such as auxin and cytokinin are also affected [92]. Improving drought tolerance has been also linked to a reduction in shoot growth, while maintaining root growth, leading to altered partition between root and shoot. This process is achieved by cell-wall synthesis and remodeling. Formation of reactive oxygen species (ROS) and lignin peroxidases are key steps involved in cell wall thickening. Stomatal closure limits CO_2 uptake by leaves which leads to a reduction in photosynthesis. Severe drought also limits photosynthesis by reducing the activity of ribulose-1, 5-bisphosphate carboxylase/oxygenase (Rubisco), fructose-1, 6-bisphosphatase (FBPase), phosphoenolpyruvate carboxylase (PEPCase), pyruvate orthophosphate dikinase (PPDK) and NADP-malic enzyme (NADP-ME) [93]. Plant responses to drought affect vegetative growth by reducing leaf-area expansion and total dry matter (TDM), which in turn decreases light interception [94]. Under drought,

wheat (*T. dicoccoide*s) has shown a reduction in the number of grains, grain yield, shoot dry weight and harvest index [95]. In soybean, the loss of seed yield was reported to be maximal when drought appeared during anthesis and the early reproductive stages [96, 97].

Various drought-stress related genes have been identified in *Arabidopsis thaliana,* rice (*Oryza sativa*) and other plants [98], which can be classified into two main groups: (i) Effector proteins whose role is to alleviate the effect of the stress (such as water channels proteins, detoxification enzymes, LEA proteins, chaperones and osmoprotectants), (ii) Regulatory proteins which alter the expression or activity of effector genes and modify plant growth, such as the transcriptions factors *DREB2* and *AREB*, and protein kinases and phosphatases [99]. Several studies have reported differential expression of genes under drought stress in major crops (Table 2.4). Even though a large number of drought-related genes have been identified in plants, their stability of trait expression under different stress conditions is a major concern.

2.1.4.2 Molecular and physiological effects of drought stress on plants

Drought stress can cause cellular, physiological and morphological changes in the plant, for instance, a reduction in photosynthesis, leaf area and final yield in groundnut [100], cowpea [101] and chickpea [102]. In pea (*Pisum sativum*), germination and early seedling growth were reported to be influenced by drought [103]. Reduction in the number of grains, grain yield, shoot dry weight and harvest index were observed in wheat when it was subjected to drought stress [95]. Drought stress can affect crop growth at any developmental stage including, the vegetative, reproductive and grain filling stages at varying degrees depending on the species [104]. In soybean, the loss of seed yield was reported to be maximal when drought appeared during anthesis and the early reproductive stages [96, 97]. As water resources available for agriculture are

expected to decrease and becoming unpredictable due to climate change, the need to adopt and enhance drought-resistant in plants is essential to help to produce enough food for the ever increasing world population, and maintain environmental resilience in agriculture. For example, advanced lines BAT477 and SEA5 that are highly drought tolerant have been identified in common bean [105, 106]. Furthermore, [107] reported the introgression of wild emmer wheat, which is highly drought tolerant, into modern wheat cultivars in order to obtain drought related candidate genes for breeding purpose.

2.1.4.3 Effect of drought stress on bambara groundnut

In recent years, plant breeders have turned to landraces (i.e. locally adapted variants) for trait improvement in various crops, including barley (*Hordeum vulgare*) [108], sorghum (*Sorghum bicolor*) [109], sesame (*Sesamum indicum*) [110], and soybean [111]. An early attempt to investigate the use of landraces in drought tolerance has been carried out in wheat [112], though this did not delve into the specific genetics conferring the desirable traits. An alternative approach to identifying genes conferring drought avoidance and tolerance is to study species that are already very resilient in arid conditions. In this regard, bambara groundnut (*Vigna subterranea* (L) *Verdc.*) is a potential candidate. For bambara groundnut, several studies have been carried out to investigate the response to drought stress. Under drought stress, bambara groundnut landrace AS-17 showed paraheliotropic properties, in which the stressed plants had leaflet angles parallel to the incident radiation, leading to less transpiratory water loss due to the lower leaf temperature that resulted from decreased light interception [3]. From the results of Mabhaudhi et al. [113], bambara groundnut was observed to have drought escape mechanisms where, under drought stress, it had a shortened vegetative growth period, early flowering, reduced reproductive stage and early maturity in

order to minimise the adverse effect of drought on plant development. Higher root dry weight was reported when bambara groundnut landrace, Burkina (originally from Burkina Faso), was subjected to drought [14]. Denser and deeper root growth will allow the plant to utilise more soil moisture under drought stress. Stomatal closure plays an important role in regulating transpiration and improve plant water status over the drought stress period. Stomatal closure has been recognised as a universal response to drought stress in many species, such as rice [114], maize [115] and has been reported for bambara groundnut [90, 116]. Accumulation of proline was observed in bambara groundnut under drought stress which plays a vital role in osmotic adjustment [90]. Furthermore, Vurayai et al. [116] stated that reduced leaf area in drought-stressed bambara groundnut plants due to turgor reduction within expanding cells is common and is one of the earliest physiological responses to water stress. Bambara groundnut is more vulnerable to drought during the pod filling stage, followed by the flowering stage and then the vegetative stage, as plants stressed at the pod filling stage failed to fully recover their relative water content and chlorophyll fluorescence after irrigation was resumed [116].

Understanding and optimising the responses of bambara groundnut under drought is of central importance in order to identify the key features of the crop which need breeding attention. Drought experiments in bambara groundnut will help to identify novel drought-related genes which could be of great importance to understand biochemical and physiological behaviour of this plant during drought stress.

In an effort to generate drought transcriptomic data for this crop at relatively low cost, NGS technology can be used directly to develop molecular markers by generating transcriptomes (Expression Sequence Tags) which will then be used to identify candidate genes responsible for the crop's response to drought stress. Additional molecular markers can be discovered from translational genomics-based approaches including in-

vestigating known gene regulatory networks involved in drought stress response and tolerance in other species. For example, with the help of genomics, transcriptional regulatory networks of drought stress signals were identified in range of species [98], which will help in genomic study of less studied crops, such as bambara groundnut. Use of molecular markers for an agronomically important trait such as drought in underutilised crops, in this example bambara groundnut, will assist the integration of desirable alleles into specific genotypes that will contribute to improvement of breeding lines and the development of drought tolerant cultivars.

Table 2.4: **Stress-responsive genes contributing to drought tolerance in plants**

Functional Category	**Gene**	**Species**	**Parameters Evaluated**	**Mechanisms**	**Pathways**	**References**
AP2/ERF Transcription Factor	DREB1A (CBF)	*A. thaliana*	Activated expression of genes involved in stress tolerance (rd29A)	Signaling cascade and transcriptional regulation	ABA independent	[117, 118]
AP2/ERF Transcription Factor	OsDREB1A	*O.sativa*	Survivability	Signaling cascade and transcriptional regulation	ABA independent	[119]
AP2/ERF Transcription Factor	CBF1 (DREB 1B)	*Solanum lycopersicum*	Activated expression of genes, catalase 1 coupled with decreased accumulation of H2O2	Signaling cascade and transcriptional regulation	ABA independent	[120]
AP2/ERF Transcription Factor	HvCBF4	*H. vulgare*	Survivability	Signaling cascade and transcriptional regulation	ABA responsive	[121]
AP2/ERF Transcription Factor	OsDREB1F	*O. sativa*	Survivability	Signaling cascade and transcriptional regulation	ABA independent	[122]
AP2/ERF Transcription Factor	OsDREB1G, -2B	*O. sativa*	Survivability	Signaling cascade and transcriptional regulation	ABA independent	[123]
AP2/ERF Transcription Factor	OSDREB2A	*O. sativa*	Survivability	Signaling cascade and transcriptional regulation	ABA independent	[124]
AP2/ERF Transcription Factor	HARDY	*A. thaliana*	Survivability, WUE, photosynthesis	Signaling cascade and transcriptional regulation	ABA independent	[125]
AP2/ERF Transcription Factor	TaDREB2, -3	*T. aestivum*	Multiple	Signaling cascade and transcriptional regulation	ABA independent	[126]
AP2/ERF Transcription Factor	CBF4	*A. thaliana*	Activated expression of genes involved in stress tolerance	Drought avoidance	ABA responsive	[127]

bZIP Transcription Factor	OsbZIP23	*O. sativa*	Relative yield	Signaling cascade and transcriptional regulation	ABA responsive	[128]
bZIP Transcription Factor	OsbZIP46	*O. sativa*	Survivability, relative yield	Signaling cascade and transcriptional regulation	ABA responsive	[129]
bZIP Transcription Factor	OsbZIP72	*O. sativa*	Survivability	Signaling cascade and transcriptional regulation	ABA responsive	[130]
bZIP Transcription Factor	SlAREB1	*S. lycopersicum*	Multiple	Signaling cascade and transcriptional regulation	ABA responsive	[131]
bZIP Transcription Factor	ABF3/ABF4	*A. thaliana*	Reduced transpiration and better survival under drought stress. Growth arrest	Signaling cascade and transcriptional regulation	ABA responsive	[132]
NAC	SNAC1	*O. sativa*	Survivability, seed setting rate	Drought avoidance, signaling cascade and transcriptional regulation	ABA responsive	[133]
NAC	OsNAC9	*O. sativa*	Multiple	Drought avoidance, signaling cascade and transcriptional regulation	ABA responsive	[134]
NAC	OsNAC10	*O. sativa*	Multiple	Drought avoidance, signaling cascade and transcriptional regulation	ABA responsive	[135]
NAC	OsNAC5	*O. sativa*	Multiple	Drought avoidance, signaling cascade and transcriptional regulation	ABA responsive	[136]
NAC	OsNAC6	*O. sativa*	Survivability	Drought avoidance, signaling cascade and transcriptional regulation	ABA responsive	[137]
NAC	SNAC1	*O. sativa*	RWC, chlorophyll content	Drought avoidance, signaling cascade and transcriptional regulation	ABA responsive	[138]

NAC	TaNAC69	*T. aestivum*	Multiple	Drought avoidance, signaling cascade and transcriptional regulation	ABA responsive	[139]
Zinc Finger	DST		Survivability	Drought tolerance, signaling cascade and transcriptional regulation		[114]
Zinc Finger	ZFP252	*O. sativa*	Survivability	Drought tolerance, signaling cascade and transcriptional regulation		[140]
Zinc Finger	Zat10	*A. thaliana*	Yield, seed setting rate	Drought tolerance, signaling cascade and transcriptional regulation		[141]
Zinc Finger	OsMYB2	*O. sativa*	Survivability	Drought tolerance, signaling cascade and transcriptional regulation		[142]
Zinc Finger	TaPIMP1	*T. aestivum*	Water loss, proline content	Drought tolerance, signaling cascade and transcriptional regulation		[143]
Zinc Finger	StMYB1R-1	*S. tuberosum*	Multiple	Drought tolerance, signaling cascade and transcriptional regulation		[144]
Zinc Finger	OsWRKY11	*O. sativa*	Survivability	Drought tolerance, signaling cascade and transcriptional regulation		[145]
Zinc Finger	OsWRKY30	*O. sativa*	Survivability	Drought tolerance, signaling cascade and transcriptional regulation		[146]
Zinc Finger	ZPT2 - 3, CpMYB10	*Petunia*	Better survival rate during drought stress	Drought tolerance, signaling cascade and transcriptional regulation		[147]

Osmotic Adjustment	P5CS (Pyrroline-5-carboxylate synthase)	*O. sativa*	Increase in biomass accumulation	Drought tolerance		[148]
Osmotic Adjustment	SacB	*Beta vulgaris*	Better dry weight accumulation	Drought tolerance		[149]
Osmotic Adjustment	TPS (Trehalose-6-phophate synthetase)	*N. tabacum*	Delay in withering or enhanced moisture retention capacity	Drought tolerance		[150]
Osmotic Adjustment	IMT1 (myo-Inositol-O-methyltransferase)	*Tobacco*	Less inhibition in photosynthetic rate; better recovery from stress	Drought tolerance		[151]
Osmotic Adjustment	Trehalose-6-phophate synthetase	*Tobacco*	Increased leaf area, better photosynthetic activity and better RWC	Drought tolerance		[149]
Osmotic Adjustment	Trehalose-6-phophate synthetase	*O. sativa*	Better plant growth and less photooxidative damage	Drought tolerance		[152]
Osmotic Adjustment	TPS and TPP	*O. sativa*	Better growth performance and photosynthetic capacity	Drought tolerance		[153]
Osmotic Adjustment	AtPLC1	*A. thaliana*		Drought tolerance		[154]
Osmotic Adjustment	OsTPS1	*O. sativa*	Survivability	Drought tolerance		[155]
Osmotic Adjustment	mtlD	*E. coli*	Multiple	Drought tolerance		[156]
ROS Scavenging	OsSRO1c	*O. sativa*	Mutiple	Drought tolerance	ABA responsive	[157]
ROS Scavenging	MnSOD (superoxide dismutase)	*Alfalfa*	Better photosynthetic efficiency, yield and survival rate	Drought tolerance	ABA responsive	[158]

ROS Scavenging	MsALR (Aldose/aldehyde reductase)	*Alfalfa*	Decreased lipid peroxidation and better photosynthetic activity	Drought tolerance	ABA responsive	[159]
ROS Scavenging	AtALDH3 (Aldehyde dehydrogenase)	*A. thaliana*	decreased lipid peroxidation	Drought tolerance	ABA responsive	[160]
ROS Scavenging	Ascorbate peroxidase	*N. tabacum*	Better photosynthetic capacity under water stress	Drought tolerance	ABA responsive	[161]
Protein Kinase	OsCDPK	*O. sativa*	Enhanced levels of stress-responsive genes, rab16A, SalT, and wsi18. Stomatal movement.	Signaling cascade and transcriptional regulation	ABA responsive	[162]
Protein Kinase	DSM1	*O. sativa*	Plant growth.	Signaling cascade and transcriptional regulation	ABA responsive	[163]
Protein Kinase	OsSIK1	*O. sativa*	Survivability.	Signaling cascade and transcriptional regulation	ABA responsive	[164]
Protein Degradation	OsDIS1	*O. sativa*	Survivability			[165]
Protein Degradation	OsDSG1	*O. sativa*	Fresh weight			[166]
Protein Degradation	OsSDIR1	*O. sativa*	Survivability			[167]
Protein Degradation	OsRDCP1	*O. sativa*	Survivability			[168]
Protein Modification	SQS1	*O. sativa*	Survivability, relative yield			[169]
Nuclear Proteins	OsSKIPa	*O. sativa*	Survivability, yield	Drought tolerance. Transcriptional co-regulator		[170]
Nuclear Proteins	OsRIP18	*O. sativa*	Survivability	Drought tolerance. Transcriptional co-regulator		[171]

Metabolism Of Aba	DSM2	*O. sativa*	Survivability, seed setting rate			[172]
Metabolism Of Other Hormones	IPT	*A. tumefaciens*	Yield, biomass			[173]
Dehydrin/LEA	OsLEA3-1	*O. sativa*	Yield, seed setting rate	Drought tolerance	ABA responsive	[174]
Dehydrin/LEA	OsLEA3-2	*O. sativa*	Survivability, grains per spike	Drought tolerance	ABA responsive	[175]
Dehydrin/LEA	HVA1	*H. vulgare*	Plant growth, survivability, RWC	Drought tolerance	ABA responsive	[176]
Transporter	AtNHX1	*A. thaliana*	seed setting rate	Drought tolerance		[141]
Transporter	betA, TsVP	*E. coli*	Mutiple	Drought tolerance		[177]
Amino Acid Metabolism	OsOAT	*O. sativa*	Survivability, relative seed setting rate			[157]

2.2 Use of advanced genomics and transcriptomics for research and development in bambara groundnut

2.2.1 Advancement in crop genome sequencing and analysis

Significant progress in sequencing technologies have speeded up the time and lowered the cost per base pair, allowing a step change in access to crop genomes compared to the previous era of Sanger-based sequencing. Before the invention of next generation sequencing technologies, Sanger sequencing of bacterial artificial chromosome (BAC)-based physical maps was the main approach for genome sequencing of species, such as rice, maize and poplar (*Populus tremula*) [178]. Though physical maps of BACs provided a good template for completing gaps and correcting sequencing errors, the

genome coverage of physical maps was non-representative due to cloning bias and was relatively labour intensive. With the availability and advancement of next generation sequencing, it is possible to sequence large volumes of DNA faster, and with better genome coverage [179]. In 2010, the African Orphan Crop Consortium (AOCC) was launched with the aim to sequence 101 indigenous African crops. Bambara groundnut is one of the target species for sequencing, with the genome sequence generated from a variety from Zimbabwe (Mana), and re-sequencing will be performed on 100 genotypes developed over the years by multiple research organisations. The draft genome of bambara groundnut is expected to be published in the third quarter of 2016.

2.2.2 Molecular marker systems and breeding for drought resistance using *omic* technologies

2.2.2.1 Marker-assisted selection from major species to bambara groundnut

With the help of advanced genomic and transcriptomic data, breeders can have access to putative gene function, gene content, copy number variation between varieties, precise genomic positions and identification of both natural and induced variation in germplasm collections. In addition, promoter sequences allows epigenetic analysis and expression levels to be monitored in different tissues or environments and in specific genetic backgrounds using NGS and microarray technologies [180]. The significant reduction in cost and increased accessibility of *omic* technologies [181] has made genome-wide analysis of less studied crops possible. The availability of molecular markers and genetic linkage maps in many plant species, such as *Medicago truncatula* [182], common bean [183], soybean [184] and cowpea [185] have made it possible to dissect complex traits into individual quantitative trait loci (QTL), with sequencing and annotation of large genomics DNA fragments. Marker-assisted breeding approaches

will help in identifying important agronomic traits corresponding to various biotic and abiotic stresses [186]. Sequence data derived from *Medicago truncatula*, soybean, cowpea and common bean can be used to develop cross-species simple sequence repeat (SSR) markers. For example, *Medicago truncatula* based SSRs provided genetic markers for linkage mapping in alfalfa (*Medicago sativa*) [187] and more distantly related crop legumes [188, 189]. This approach can be applied in bambara groundnut, where a closely related legume crop sequence data, such as common bean, mung bean or soybean can help in identifying molecular markers for the traits of interest. Furthermore, with the help of genome sequencing of major legume crops targeting genes for the specific traits of interest in crop relatives, such as drought resistance, is possible. In this approach, phenotypes of interest in crop relatives are mapped and characterised against major crops using the available genomic resources (Fig.2.1) [78]. This strategy should be applicable to most closely related legume crops and can be applied in bambara groundnut. It was successful in *Medicago truncatula*, where phenotypic mapping was performed in *Medicago truncatula* ranging from disease resistance genes to QTL for morphology. Resistance against *Colletotrichum trifolii* (anthracnose) [190, 191], *Phoma medicagnis* (black steam and leaf spot) [192], and *Erysiphe pisi* (powdery mildew) [190] have all been mapped in *Medicago truncatula*. The cloned *RCT1* gene found in *Medicago trunculata* for resistance to anthracnose has been shown to function in alfalfa [191].

With the help of genomic and transcriptomic analysis, sequence data derived from major species will help in the development of new molecular markers for drought resistance and gene discovery in underutilised crops, leading to crop improvement. Though MAS has been implemented in many crop breeding programmes, it still at a primitive stage for many minor and underutilised crops. Perhaps one of the biggest challenges many underutilised crops face is the absence of structured genetic resources that allow

a powerful dissection of the genetic control of complex traits. Many traits are governed by multiple genes and it is more difficult to understand the underlying genetic control of these quantitative traits, without structured (and preferably immortal) populations and genetic stocks. Additionally, imprecise localisation of the QTLs and instability of QTL between experiments and environments adds to the complexity of this approach for underutilised species [193].

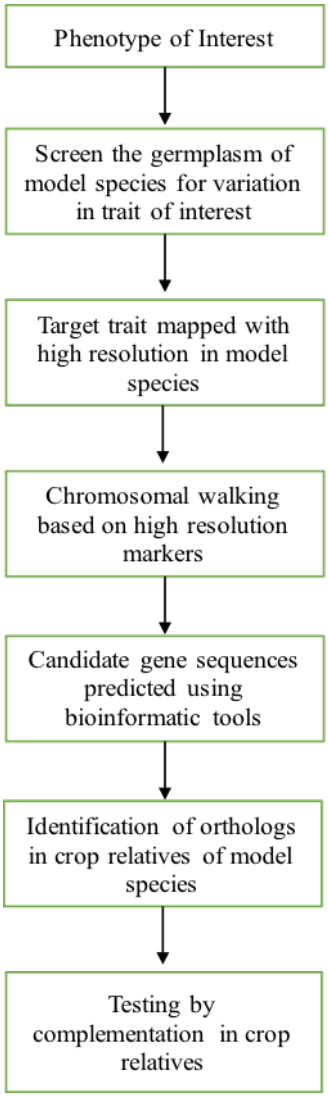

Figure 2.1: **Gene discovery in underutlised species from model species**. Key steps listed, where a model species was used to find gene of interest based on genome location. [78]

2.2.2.2 Genomic-assisted breeding for drought resistance

The availability of advanced expression analysis techniques such as NGS, microarrays, real-time PCR, transcriptomics, proteomics and metabolomics platforms have made it possible to carry out extensive gene expression analysis to identify and characterise candidate genes for drought tolerance [194] (Table 2.4). Breeding for drought resistance is challenging as it is a complex trait controlled by many genes, there is limited knowledge of the inheritance mechanisms and the effect of drought is different for every genotype [195]. Knowledge of the relative values of the alternative alleles at all loci segregating in a population could allow breeders to design a genotype in silico [196]. Marker-assisted backcrossing approaches and marker-assisted recurrent selection have been implemented in legume crops such as chickpea to improve the crop's drought resistance [197]. Breeding approaches, such as marker-assisted backcrossing approaches and marker-assisted recurrent selection have recently been complemented by a new approach called genomic selection (GS), which predicts the breeding values (i.e genomics assisted breeding values; GEBVs) of lines in the next generation based on historical genotyping and phenotyping data [198, 199]. GS has begun to be incorporated into breeding programs for crops, although many questions remain in terms of which crops are most suited and what some of the key parameters for successful deployment are [200]. With genome coverage provided by GS, NGS will help in providing estimates of gene expression levels and determination of epigenetic states of genes [180].

To date, there has been no broad screening of the bambara groundnut germplasm under drought stress. Though the crop is known to be drought-tolerant from the physiology point of view, drought tolerance should be linked to comparatively higher growth rates and productivity under water stress conditions. In bambara groundnut, populations were analysed for specific genetic traits of interest which includes genetic mapping of photoperiod response in bi-parental populations after development of sin-

gle genotype parental lines [201] and mapping of phenotypic traits associated with domestication syndrome in bambara groundnut and plant morphology [15, 202, 203]. The genotypes obtained from crosses in bambara groundnut are important resources to study and optimise for specific traits through breeding programmes [**?**]. Genomic resources including SSR markers in bambara groundnut were developed and will be employed in identification of QTLs for specific physiological traits relevant for drought tolerance [204]. Results from Beena et al. [204], showed that there was a significant reduction in the physiological (transpiration, photosynthesis and chlorophyll content) and morphological (leaf area, total biomass) traits under water limiting conditions. The availability of SSR [204, 205] and DArT [206] markers of bambara groundnut have made it possible to carry out diversity analysis, mapping of QTLs for various agronomics traits under drought and their use in marker-assisted breeding. Random amplified polymorphic DNA (RAPD) and fluorescence based amplified fragment length polymorphism (AFLP) have been developed for several landraces of bambara groundnut and the study revealed high levels of polymorphism among landraces [207]. Genetic linkage maps of bambara groundnut were constructed by combining microsatellite and DArT markers from a 'narrow' and 'wide' cross between bambara groundnut landraces Tiga Nicuru x DipC and DipC x VSSP11 to identify marker-trait linkages and to develop the crop through marker-assisted selection by selecting marker allele that is linked to a trait of interest [208, 209]. The genetic map created from the Tiga Necaru x DipC cross was based on an intra-sub-specific cross exploits variation within the domesticated landraces gene pool, using an F_3 population of the cross between DipC and Tiga Necaru [12, 24, 202, 209]. This cross showed variation for agronomic traits of breeding interest. This map will be useful for comparative genomic analysis between the mapping populations in this crop and also between bambara groundnut and other related legume crops [202]. Based on the results from

the 'narrow' genetic cross (F_3 generation) and 'wide' genetic cross (F_2 generation) for Tiga Necaru x DipC and DipC x VSSP11, respectively, a candidate marker bgPabg-596774 was identified for the following traits; pod number, node number, pod weight, seed number, seed weight and biomass dry weight which could be used for MAS [202]. Development of the population (Tiga Necaru x DipC and DipC x VSSP11) into full Recombinant Inbred Lines (RILs) (at F_6 generation currently) or the development of Near Isogenic Lines (NILs) could allow the evaluation of the effects of these QTL alone and also the development of heterozygous plants for the QTL region, allowing large-scale fine mapping programme [202].

Due to the limited understanding of the drought response mechanisms that are active in different germplasm within bambara groundnut, implementing molecular breeding for drought resistance or selection of candidate genes for gene editing is challenging. Furthermore, high temperature is usually co-incidental with drought stress. Therefore, a single gene effect to increase drought resistance is possible, but the combined stress may require a multi-gene transformation strategy that combines several major functional or regulatory genes or a series of genes in a signalling cascade contributing to drought resistance seems promising for improving long term drought resistance in plants [210]. In addition, combining traditional breeding (such as cross and/or recurrent backcrossing of wild relatives and elite cultivars) will help in building the desired traits for abiotic resistance [210] in bambara groundnut.

2.2.3 Comparative genomics and transcriptomics from major species to bambara groundnut

Advances in crop genomics, transcriptomics, molecular and bioinformatic tools have given us an opportunity to understand plant biology in a more unified way and also help transfer information from a major species to minor species [211]. Strong conven-

tional breeding practice is required for successful genomic investment on underutilised crops [193]. Exploiting biotechnological tools and translation research from major species to underutilised crops will lead to better results in breeding processes using two main methods (1) translation of technologies, such as marker system approaches based on next generation sequencing (2) translation of actual genetic-trait information from related species based on locational or network analysis [**?**].

2.2.3.1 Comparative genomics

Comparative genomics has been widely used in modern day research as it can provide important information about species whose genomes have not been sequenced by comparing with a known species [212]. Rapid advancement in crop genomics has provided a chance to conduct detailed functional and structural comparisons of genes involved in various biological processes among major crops and other plant species. Therefore, comparative genomics using bioinformatic tools can provide an opportunity for transferring important information from major species to bambara groundnut. Some examples that illustrate comparative genomics in legumes are; (1) Use of barrel medic to map-based clone the *RCT* gene that confers resistance to multiple races of *anthracnose* (*Colletotrichum trifolii*) in alfalfa [191], (2) Floral regulatory genes identified in *Arabidopsis thaliana* were used to find genes in common bean effecting determinacy [213] and (3) Identification of the gene underlying Mendel's I locus, responsible for the trait corresponding to yellow or green colour of seed in grass meadow fescue [214]. A candidate gene that plays a critical role in cholorophyll catabolism during plant senescence was found in rice and later its orthologous gene was fine mapped in pea (*Pisum sativum*). Thus traits, genes, tools and species were combined to link this trait and the underlying gene in several models and crops [214]. In addition, minor crops can provide good models for a trait and trait analysis absent in major crops [193].

For example, if alleles contributing to drought tolerance can be found in bambara groundnut, the underlying physiological mechanism and the genes responsible could be useful and can be utilised by MAS to search for alleles within the species of interest or the trait might be transferred through direct gene transfer into major crops.

2.2.3.2 Comparative transcriptomics

Comparative transcriptomics has been tested in bambara groundnut. Comparisons between the bambara groundnut leaf transcriptome and other species has been carried out to identify appropriate cross-species orthologues and gene models for the crop [215]. The results showed that soybean had the highest transcript sequence similarity to bambara groundnut than any other species used in the analysis (other species were *Medicago truncatula, Vitis vinifera*, *Populas trichocarpa*, *Ricinis communis*, *Arabidopsis lyrata*, *Vigna radiata*) and could potentially be used as a gene model for gene expression profiling in bambara groundnut [215], although the tetraploid nature of soybean adds complications.

Bambara groundnut DNA was hybridised to *Arabidopsis* ATH1 and *Medicago truncatula* Affymetrix GeneChips for high and low stem number respectively as there is no Affymetrix GeneChip available for bambara groundnut [24, 203]. The cross species microarray approach coupled with genetical genomics has been applied on bambara groundnut using the soybean Gene Chip array. The drought experiment conducted by [12] used leaf RNA from an F5 segregating population derived from a controlled cross of between DipC and Tiga Nicuru cross-hybridised to the soybean GeneChip from Affymetrix. The results identified 1531 good quality gene expression markers (GEMs) on the basis of differences in the hybridisation signal strength. An expression based genetic map was constructed using 165 GEMs. Significant QTLs were detected using the GEM map for various morphological traits (including internode length, pe-

duncle length, pod number per plant, pod weight per plant, seed number per plant, seed weight per plant, 100-seed weight, shoot dry weight). An XSpecies microarray experiment was conducted in order to identify and detect genes and gene modules associated with low temperature stress responses in bambara groundnut. This found 375 and 659 differentially expressed genes ($p<0.01$) under the sub-optimal (23°C) and very sub-optimal (18°C) temperatures, respectively. Further, 52 out of top 100 differentially expressed genes were validated using NGS technology generated from the same samples used to generate cross-species microarray data. The results showed 50% similarity between the XSpecies Microarray approach and NGS technology. The identified gene modules could be useful in breeding for low-temperature stress tolerant bambara groundnut varieties [216]. These approaches have the potential to identify polymorphisms between individuals for gene expression analysis and mutation discovery [217], which will accelerate the generation of markers for specific traits in minor crop species [203].

2.3 Concluding remarks

Recent developments in genomics and transcriptomics have opened up opportunities to develop data sets for several underutilised species which could facilitate crop improvement. Genomics has led us to gather a wealth of information from the identification of genetic variation, epigenetic states of genes and the potential to measure gene expression with high precision and accuracy. This will not only benefit breeding but also facilitates systematic comparison of gene functions across sequenced genomes which will directly benefit crop improvement. Projects, such as Encyclopedia of DNA elements (ENCODE) (although yet to be implemented in plants) will build a foundation for extracting knowledge of gene function and variation, thus generating new data for the prediction of phenotype from genotype [180]. Knowledge gained from integration

of gene function into networks, such as controlling flowering time in response to day length and over-wintering will pave the way for crop improvement. These networks have been identified in *Arabidopsis thaliana* and rice, with allelic variation strongly influencing networks outputs. Processes, such as gene duplication and footprints of domestication can be mapped to networks such as flowering time [218, 219]. Improved precision of predicting the phenotype from genotype is possible with the use of 'systems breeding' approaches which use diverse genomic information leading to food security and crop improvement [180].

The improvement in genomics and transcriptomics will help in identifying target genes that underlie key agronomic traits related to drought. Molecular markers will be developed using the information gained from the trait of interest which will be later on used for breeding applications. Finding gene targets that are related to various biotic and abiotic stresses will be productive towards the aim of crop improvement, as plant growth is severely effected by stresses, such as drought, cold and salt especially in marginal physical and economic environments. Molecular analysis of bambara groundnut germplasm using advanced genomic tools will help in the discovery of genes for key agronomic traits. Functional genomic tools, physical maps and the availability of high-throughput and cost-effective genotyping platforms will all contribute towards crop improvement. There are various challenges that have to be looked upon before applying genomics to underutilised crops. Most of the underutilised crops lack large-scale collections of germplasm, although local communities and small-scale farmers do have extensive knowledge that help in the search for genes that are vital for crop improvement. Restriction on flow of germplasm due to intellectual property rights is also one of the major challenges [220].

The potential value of underutilised crops (and bambara groundnut) is the part they can play in minimising Africa's challenges of rural development, hunger, malnu-

trition and gender inequality. Bambara groundnut is resilient and reliable crop that thrives in unsuitable areas which could potentially be unsuitable for peanut, maize, or even sorghum [73]. It is mainly reported to be grown by women, therefore, offers a reliable way of empowering women financially, hence improving the lives of their families [1]. As a legume, it has the ability to fix atmospheric nitrogen either as an intercrop or rotational crop, thus minimising the use of chemical fertilisers. Being a cheap source of soil nitrogen, it can help resource-poor farmers to achieve some added nitrogen for the growth of the main crop.

The Green Revolution that partially succeeded in Africa, but has led to a focus on major crops which, in some cases, are been grown in the wrong places and under the wrong agricultural systems for them to ever be truly productive. Investigating the many underutilised crops which exist as a component of climate resilient, low input, agriculture is one way to mitigate the risk of total crop failure. The advent of next generation sequencing has opened up the possibilities for minor crops, allowing both within species analysis and comparative analysis to related species. Access to germplasm is still a significant issue, but the tools to begin a more context and nutritionally focused agricultural revolution are coming into place.

Chapter 3

Materials, Methods and Data

This chapter outlines the materials, methods, software tools and datasets used in this project.

3.1 Bambara groundnut microarray XSpecies analysis

3.1.1 Plant materials

The F_5 segregating population derived from the cross between two genotypes of bambara groundnut; Tiga Nicuru (maternal) and DipC (paternal) lines and their two parental lines were evaluated in this drought experiment to determine the segregation patterns and traits inheritance related to drought in the two parents of the cross. 65 F_5 individual lines and also the parental lines of the genotypes DipC and Tiga Nicuru were planted in both drought and irrigated control plots. Drought responsive gene expression profiling and gene co-expression network analysis was performed on the two parental lines. [24] studied the effect of drought stress on the morpho-physiological characteristics for the parental lines and the F_5 segregating population derived from TN x DipC cross. This study deals with the transcriptomic analysis of the two parental lines.

3.1.2 Site description, experimental design and crop management

This glasshouse experiment was conducted by Hui Hui Chai [24] and the author performed the XSpecies microarray data analysis, co-expression network analysis, primer designing and qRT-PCR validation of the selected genes.

The experiment was conducted between late June 2012 and late November 2012. A 12-hour photoperiod was created using an automated blackout system (Cambridge Glasshouses, Newport, UK) with day and night temperatures set at 28°C and 23°C respectively. Trickle tape irrigation with PVC micro-porous tubing was placed beside

each plant row. The DipC and TN plants were irrigated at 0600 hours and 1800 hours for 20 minutes with a measured flow rate of 1L/hour per tube, with each tube 5m in length. Two independent soil pits (5 m x 5 m x 1 m) containing sandy loam soil were used in the glasshouses, with each pit isolated by a Butyl liner and concrete pit structure for drought and irrigated plots separately. The PR2 water profile probe (Delta-T devices, UK) was used to measure the soil moisture content. A randomised block design (RBD) with three blocks for each soil pit was implemented for this experiment. Three replicate plants for the control (continuously irrigated) and four replicates for drought treatment plot were used. Three seeds were sown per replicate at a depth of 3-4 cm with a spacing of 25cm x 25cm between each plant position and later thinned to one plant per replicate at 20 days after sowing (DAS). Figure 3.1 shows the treatment regime. The irrigation system for the drought treatment plot was turned off at 50 DAS and resumed at 92 DAS for the drought treatment plot (in total a six weeks treatment after 100% flowering). Normal irrigation continued for the irrigated plot throughout. During drought, an average 50% reduction in stomatal conductance was observed. Leaves from irrigated and drought-treated plants were collected at 92 DAS, while those from recovered plants were also collected at 107 DAS. Labelled aluminum foil was used to wrap the harvested leaf, which was then transferred into liquid nitrogen to flash freeze the samples. All samples were stored in a -80°C freezer before RNA extraction. DNA extraction from the two parental genotypes was done using the DNA extraction Qiagen kit.

Throughout the growing season, *Phytoseilus persimilis*, a biological control agent, was used against red spider mite (*Tetranychus urticae*) and applied every two weeks. In addition, chemicals such as Savona (soap) against Aphids and Thiovit (sulphur) against mildew or fungal infections were applied as needed.

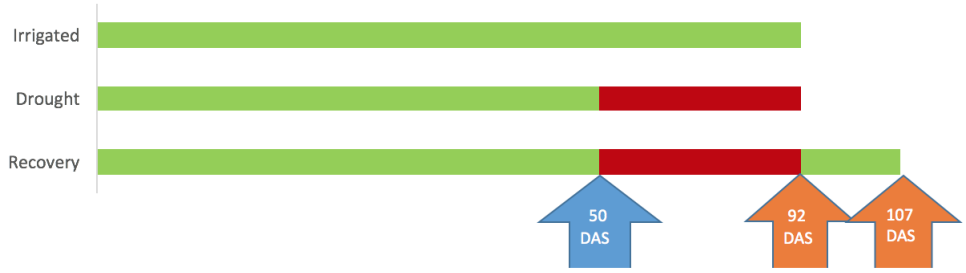

Figure 3.1: **Treatment Regimes.** Horizontal bars represent the timelines for the three treatments. The green and red regions correspond respectively to periods of irrigation and drought, with the numbers in the arrows referring to the start/end of the different periods to the days on which leaves were collected for transcriptomics. DAS = Days After Sowing. Leaves from irrigated and drought-treated plants were collected at 92 DAS, and leaves from the recovered plants at 107 DAS.

3.1.3 Environmental factor measurements

Environmental factor measurements were taken by Hui Hui Chai, University of Nottingham. Environmental factors within the glasshouse were controlled and monitored using an automated record system (Cambridge Glasshouses, Newport, UK) placed in the glasshouse to maintain a consistent environment for the growing of bambara groundnut in the glasshouse. The conditions, such as photosynthetically active radiation (PAR), humidity and temperature were recorded every eight minutes throughout the experiment. For soil moisture measurements, three PR2 profile tubes (Delta-T devices, UK) were inserted into each soil pit across the diagonal from the irrigation source towards the end of the trickle tape, at least 1m apart from each other. Three PR2 readings, which are displayed in the unit of %Vol (volumetric water content as a percentage), were taken twice a week at 1000 am starting from 16 DAS until 133 DAS at soil depths of 300mm, 400mm, 600mm and 1000mm.

3.1.4 RNA and DNA extraction

RNA and DNA extraction was performed by Hui Hui Chai, University of Nottingham. Leaves from irrigated and drought-treated plants were collected at 92 DAS. Also, three replicates of the two genotypes from the drought treatment plot that were subjected to a 2-week water recovery were harvested 107 DAS, with the aim to compare plant expression during drought stress and after recovery. Labelled aluminum foil was used to wrap the harvested leaf, which was then transferred into liquid nitrogen. All samples were stored in a -80°C before RNA extraction. In addition to RNA samples, DNA was extracted from two genotypes as well. Extraction of genomic DNA (gDNA) from the two genotypes was performed based on DNA extraction Qiagen kit. RNA was extracted using the RNeasy Qiagen kit (Qiagen, UK) according to the manufacturer's instruction. DNA was eliminated using DNase. 80µl of DNase I incubation mix, containing 10µl DNase I stock solution and 70µl buffer RDD, was added and incubated at room temperature for 15 min. Nanodrop and gel electrophoresis were performed to check the quality and quantity of RNA as RNA samples requires 100ng/µl and 10µl for microarray analysis. For samples with high concentration (>400ng/µl), a 2-fold dilution was done. To make sure the samples were free from active RNAse, 0.63µl of 40 U/µl RNasin (Promega, UK) was added for every 25µl of RNA sample.

3.1.5 Quantification of nucleic acid

Quantification of nucleic acid was performed by Hui Hui Chai, University of Nottingham. The concentration and quality of nucleic acid was examined using spectral absorbance ratios and electrophoretically on an agarose gel. Spectral absorbance ratios (A260/280) of DNA and RNA (ng μl^{-1}) were determined using the Nanodrop ND1000 spectrophotometer (Thermo Scientific, USA) using the ND-1000 V 3.7.0 software. The

pedestal of Nanodrop was first cleaned with 2 µl sterile water, followed by loading 2 µl samples onto the pedestal for measurement. A ratio of ~1.8 was generally accepted for DNA as being of good quality whereas a ratio of ~2.0 was required for RNA. Furthermore, samples were tested on a Agilent bioanalyser for integrity before preparation for the microarray by the Nottingham Arabidopsis Stock Centre (NASC). RNA samples with 2 µl each were loaded into the PCR tubes and sent to Plant Sciences, The University of Nottingham, Sutton Bonington Campus, UK for Agilent analysis.. The size of the 18S peak and 28S peaks were then calculated, a ratio of 2 is ideal as the 28S/18S ratio is one of the key indicators of RNA quality (Table 3.2)

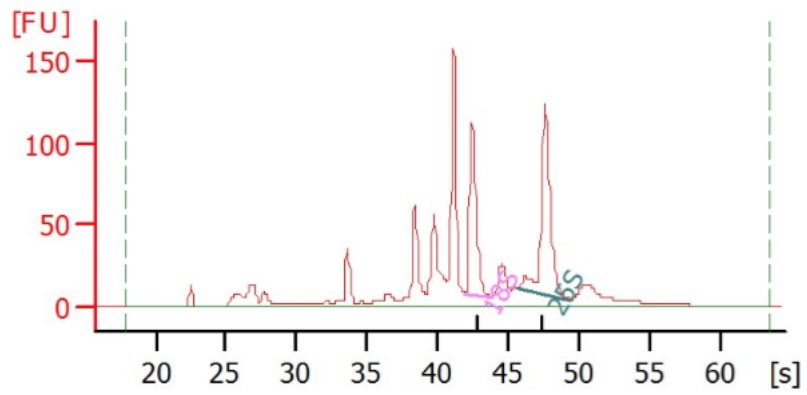

Figure 3.2: **Agilent analysis of high quality RNA using Qiagen commercial kit was presented**. X-axis: Runtime (s); Y-axis: Fluorescence units

3.1.6 Primer Design for qPCR validation

PCR forward and reverse primers were designed using Primer-BLAST [221] for the selected drought induced genes from this study. The primers were designed in three steps. Firstly, the target gene sequence for which the primers needed to be designed was downloaded from the soybean (*Glycine max*) database (http://soykb.org/). Secondly, soybean specific target gene sequence was blasted against the bambara groundnut transcriptome generated from RNA-seq data for a low-temperature stress experi-

ment [216], by creating a BLAST database. Thirdly, the target gene sequence obtained from bambara groundnut BLAST database was entered and searched through BLAST database on NCBI website (http://www.ncbi.nlm.nih.gov/) to confirm the target gene. Once the gene sequence was generated from BLAST database, it was utilised to produce primers with appropriate primer size, GC content, melting temperature (Tm) using Primer-BLAST.

3.1.7 Polymerase chain reaction (PCR)

PCR was performed to check the quality of all the primers designed from the XSpecies microarray study. PCR analysis was performed using the 7000 Sequence Detection System (Applied Biosystems). Cycling parameters were set as: 95ºC for 10 min, 40 cycles of denaturing at 95ºC for 30 s, annealing at 60ºC/58ºC for 30s, and extension at 72°C for 30 seconds. 20 µl PCR master mixes was set up as below:

Table 3.1: **PCR mix for 20 µl reactions for each pairs of primers**

Master mix recipe for PCR	
PCR components	**Volume (µl)**
25 mM MgCl2	1 µl
10mM dNTPs	0.4 µl
Taq Polymerase (Invitrogen)	0.4 µl
cDNA template	2 µl
10x MasterMix	2 µl
Forward primer	0.4 µl
Reverse primer	0.4 µl
Nuclease free water	13.4 µl
Total	20µl

3.1.8 Gel electrophoresis

To make a gel, agarose (Bioline, UK) was dissolved in 0.5X TBE buffer and heated in microwave with occasional swirling until a clear solution was observed. After cooling, either SYBR® Safe or Ethidium bromide (0.5 µl; 10 mg ml^{-1} stock; per 50 ml gel) was added and the gel was poured into an appropriate 54 gel cast tray. DNA, RNA and PCR products were quantitated and/or checked by running them respectively on a 1% (w/v), 1.5% (w/v) and 2% (w/v) stained agarose gel at 80 V for 60 min. When PCR products were subjected for analysis, 2-log ladder was also loaded alongside with the samples in order to identify the band size. The gel was then visualised under UV light using the Gel Doc 2000 Gel Documentation System and associated Quantify One 1-D Analysis Software (Biorad, California, US).

3.1.9 cRNA and genomic DNA Affymetrix labelling and XSpecies hybridisation

The cross-species (Xspecies) approach involved cross-hybridisation of nucleic acids of target species onto the microarray derived from closely related species. In this approach, cRNAs from a species of interest are taken and hybridised to a closely related species commercial microarray chip. In this case, the Glycine max GeneChip was used as it belongs to the same legume Family as bambara groundnut [222]. A volume of 10 µl of 50 ng $µl^{-1}$ and 100 ng $µl^{-1}$ was prepared for DNA and RNA samples respectively to cross-hybridise onto a microarray. Prior to XSpecies analysis, a preliminary quality check was carried out for the samples using Agilent 2100 Bioanalyzer. RNA extracted from bambara groundnut under irrigated, drought and water recovery conditions were reverse transcribed to synthesize double stranded cDNA. After the purification of double-stranded cDNA products, they were transcribed in vitro to generate

Biotinylated complementary RNAs (cRNAs) followed by purification and fragmentation. The purified and fragmented cRNAs were then hybridised to Affymetrix Soybean GeneChip array. The scanned arrays produces .CEL raw data files which were loaded onto Genespring GX version 13.1 to do further analysis. Extracted DNA was labelled and hybridised to the Affymetrix Soybean TEST3 array which resulted in generation of gDNA cell-intensity files (.CEL files) after scanning. To identify probe pairs that hybridise efficiently to the gDNA, a series of user defined threshold values were used to analyse the effect on exclusion of probe-sets as the threshold is increased. The perfect match (PM) probes which hybridises efficiently to gDNA were selected for interpreting GeneChip arrays challenged with RNA from the species of interest [31]. All DipC and TN samples under drought, irrigation and after water recovery were sent for microarray analysis and were cross-hybridised onto the Soybean GeneChip array. XSpecies hybridisation was conducted in The Nottingham Arabidopsis Stock Centre (NASC) International Affymetrix service, The University of Nottingham, Sutton Bonington Campus, UK.

3.1.10 Probe Selection

The soybean array contains >37,500 probe sets each containing 11 probe pairs per probe set. In addition, the soybean array also contains transcripts to study nematode and fungal transcripts containing 15,800 *Phytophthora sojae* and 7,500 *Heterodera glycines* transcripts. The oligo probes are synthesised in situ and are complementary to each corresponding sequence. 11 pairs of oligo probes were used to measure the level of transcription of each sequence represented. Probe pairs from the gDNA hybridisation files (CEL) were selected using a .CEL file parser script (http://affymetrix.arabidopsis.info/xspecies/), which generates a probe mask, a cell description file (CDF), which was imported into Genespring GX version 13.1 (Agilent

Genomics) for further analysis. gDNA intensity thresholds of 0 to 1000 were used to generate a custom CDF file that will be used to analyse RNA samples [40]. After hybridisation, a probe pair was retained if the perfect match hybridisation value exceeded a series of user-defined threshold (0 to 1000) using a CEL file parser [31], which also generated a CDF file for each threshold value.

Customised CDF files generated from the hybridisation were used to filter out raw expression intensity values (RNA.CEL files) for low- or non-specific hybridisation and then normalised using the RMA (Robust Multichip Average) algorithm available in Genespring. Differentially expressed genes (DEGs) were calculated for each threshold value using a T-test (corrected by Benjamin Hochberg FDR multiple testing). Probe-sets of corrected p-value <=0.05 and fold change of > 2 was considered to be differentially expressed (either up or down regulated). Probe-pair retention was evaluated at a series of thresholds (0 to 1000). Principal Component Analysis (PCA) was also carried out in GeneSpring and Bioconductor package "prcomp" (https://stat.ethz.ch/R-manual/R-devel/library/stats/html/prcomp.html).

3.1.11 Differentially expressed genes identification (DEGs) and Normalisation

Before the differential expression analysis, assessment of correlation between the samples under different treatments (drought, irrigation, recovery) was done for both genotypes using principal component analysis (PCA). In a microarray experiment, it is not possible to make a visual inspection of the relationship between genes as the expression of thousands of genes is measured across different conditions such as treatments and time points. Principal component analysis reduces the dimensionality of the experiment which makes it easier to identify the relationship between samples.

Initially to identify differential gene expression, RNA.CEL files were loaded into

Genespring. RMA normalisation was performed on the RNA samples to reduce any chance of gDNA being hybridised to Mismatch (MM) probes. Greater sensitivity and specificity of RMA makes it an ideal method for detecting differential expression levels and fit to models [223]. The RNA hybridisation files (CEL) were loaded in Genespring applying the customised CDF generated from the gDNA hybridisation. Genes that had a FDR corrected p-value <=0.05 and a fold change of > 2 from a t-test (corrected by Benjamin Hochberg FDR multiple testing) were selected as DEGs between treatments (e.g. Drought, Irrigation and Recovery).

3.1.12 Identification of over-represented Gene Ontology terms

Prior to the co-expression network analysis, GO-term overrepresentation of the gene-sets obtained from the hybridisation of Soybean Affymetrix GeneChip array with DipC and TN was performed to investigate whether gene-sets associated with particular biological functions as represented by Gene Ontology (GO) annotations, are statistically overrepresented in the identified gene groups using the Benjamini & Hochberg FDR corrected hypergeometric test. Go-term overrepresentation of all the gene-sets in Soybean GeneChip array was also performed to compare between DipC, TN and Soybean datasets. The gene ontology tool, BINGO, was used for GO-term over-representation [224]. BINGO maps the predominant functional themes of the tested gene set on the GO hierarchy, and with the help of cytoscape's version 3.4.0 visualization environment to produce an intuitive and customizable visual representation of the results.

3.1.13 Construction of the co-expression network

Co-expression network analysis was carried out using DeGNserver [57] and cytoscape 3.4 [58]. Separate networks were generated for each genotype. The input probe-sets were restricted to those that were differentially expressed between every pair of

treatments (Drought, Irrigation and Recovery) and RMA-normalised values were used across all samples. Links were assigned between pairs of nodes (i.e. probe-sets) when their Spearman's Rank correlation was 0.9. This threshold value was used because high number of genes were predicted at threshold 0.9 without any loss of information. Once the co-expression network was created, all the up and down regulated genes under drought stress for both DipC and TN were chosen to visualise and identify relevant co-expression network. The co-expression network was imported into cytoscape for visual representation and network analysis. For each landrace, another input file was made which, for each probe-set, defined the parent (DipC or TN), the direction of differential expression caused by drought (up or down), and role in relation to drought tolerance. This aided interpretation of the combined network derived from both landraces

3.1.14 Expression validation of differentially expressed genes using real-time qPCR

Four genes which are potential candidate drought associated genes with an expression level of > 2-fold change and FDR corrected p-value <=0.05 from the differential expression analysis were chosen for qPCR validation. The actin-11 from the bambara groundnut transcriptome sequence was used as a control gene. Actin-11 is known to be one of the most stable reference genes for gene expression normalisation in soybean and rice [225, 226]. PCR forward and reverse primers were designed using Primer-BLAST [221] for the chosen differentially expressed genes (Details on how the primers were designed are discussed in 3.1.6). PCR was performed to check the quality of all the primers designed for the four drought-related genes (see 3.1.7 for details).

First Strand cDNA Synthesis Kit, quantitative RT-PCR kits (Invitrogen) and the RNA samples used for microarray profiling were used for qRT-PCR analysis. First strand cDNA synthesis for all the RNA samples was prepared using SuperScript III

First-Strand Synthesis kit. The first-strand cDNA was prepared for analysis by qPCR using PerfeCta SYBR Green SuperMix on a LightCycler 480 Real-Time PCR system (Roche), containing 2X reaction buffer (with optimized concentrations of MgCl2), dNTPs (dATP, dCTP, dGTP, dTTP), AccuStart Tag DNA Polymerase, SYBR Green 1 dye and stabilizers. The synthesized cDNA is cleaned from remaining RNA using the enzyme mix included in the kit (*E.coli* RNase H). qPCR components was prepared for 10 µl reaction (Table 3.2). Melting curve analysis was performed. Sample cycle threshold (Ct) were standardized for each template based on the actin-11 gene control primer behaviour. The 2(-Delta Delta C(T)) method was used to analyze the relative changes in gene expression from qRT-PCR experiment. To validate if the right PCR product was generated in the expression studies, the desired fragment of intact cDNA for all genes was sent for sequencing after the gel extraction using QIAquick Gel Extraction Kit.

Table 3.2: **qPCR components for 10 µl reactions**

qPCR components	**Volume (µl)**	**Concentration**
PerfeCta SBYR Green	5µl	1X
Forward Primer	0.25µl	250nm
Reverse Primer	0.25µl	250nm
Rnase free water	2µl	
cDNA	2.5µl	
Total	10µl	

3.2 Morpho-physiological effects of drought stress on bambara Groundnut at different developmental stages

3.2.1 Experimental site and plant materials

The drought stress experiment was conducted in a growth room at Sutton Bonington Campus, the University of Nottingham, UK (Fig. 3.3). The growth room will be set to a 12 hour day length (12 hours light, 12 hours dark cycle within each 24 hour period) with a constant temperature of 23°C (night) - 28°C (day) for the duration of the experiment. Temperature and humidity were also be constantly monitored. Two genotypes derived from the landraces of bambara groundnut; DipC and Tiga Nicuru were used in this study. Peanut (*Arachis hypogaea*) was used for the comparative study with bambara groundnut in relation to drought tolerance. Peanut accessions were provided by USDA (http://www.usda.gov/wps/portal/usda/usdahome). Peanut accession GRIF 302 from India was used for this study.

Figure 3.3: Growth room at the Sutton Bonington campus, University of Nottingham, UK.

3.2.2 Experimental design and crop management

Experiments were set up using a Complete Randomised Block Design (CRBD), the first experiment will keep normal (control) irrigation (75% field capacity) conditions, and second imposed drought conditions at 50% - 25% field capacity. The experiment was divided into three developmental stages - vegetative, reproductive and pod development. The drought was imposed at 19, 33 and 61 DAS for vegetative, reproductive and pod development stage, respectively. Each developmental stage was established as a separate experiment. Three biological replicates per treatment (i.e. drought and irrigated) for DipC, TN and Peanut were used at each stage for this study. Pots with the volume of 10 litre were filled with soil (50/50 JI3 + sand mix), watered to 75% field capacity and seeds subsequently planted at 1 cm depth. Two plants per pot were grown, one each for non-destructive and destructive analysis. Plants were weighed

every day and watered accordingly to maintain required field capacity. For vegetative and reproductive stages, the morpho-physiological measurements were taken from 19 DAS till 33 DAS and from 33 DAS till 47 DAS at an interval of 7 days over a period of two weeks, respectively. For the pod development stage, similar measurements were taken from 61 DAS until 103 DAS at a 14 days intervals over a span of six weeks (see Table 3.3). Yield measurements were taken at the time of harvest. To ensure that there is no systematic bias for the measurements, all plants were at the same developmental stage and had the same field capacity at the time of measurements. Leaf sampling was done at the start and end period of the drought treatment at each stage (drought and control plants) for DipC and TN, for the extraction of RNA. Leaf samples from vegetative and reproductive stages were send for RNA-seq analysis.

Throughout the growing season, plants were carefully monitored in order to protect plants against thrips attack and other pests and diseases.

Table 3.3: **Morpho-physiological treatment regimes**

Developmental stages	**Measurements at DAS***	**Drought weeks**	**Field Capacity (%)**
Vegetative stage (7-33 Days)	19 DAS*	Week 0	75%
	26 DAS*	Week 1	50%
	33 DAS*	Week 2	25%
Reproductive stage (33-61 Days)	33 DAS*	Week 0	75%
	40 DAS*	Week 1	50%
	47 DAS*	Week 2	25%
Pod development stage (61-150 Days)	61 DAS*	Week 0	75%
	75 DAS*	Week 2	50%
	89 DAS*	Week 4	50%
	103 DAS*	Week 6	25%

**DAS = Days after sowing*

3.2.3 Field Capacity

Field capacity of the soil was determined following the gravimetric field capacity test using the method of Mabhaudhi et al. [113]. Three small drained pots (representing three reps) were used. Each pot was filled with soil. Thereafter, water was added to the pots until saturation was achieved. Pots were then left to drain for 12 hours and thereafter mass of soil was measured hourly until a constant mass was reached. At this point it was assumed that the soil was now at field capacity. Following this, the soil was taken out, put it in labelled brown paper bags and the wet mass of the soil determined. Thereafter, brown bags with soil in them were put to dry in an oven set at 80°C for 72 hours after which dry mass of the soil was measured. Gravimetric field capacity (GFC) was then calculated as follows:

$$\theta m = \left(\left(\theta w - \theta d\right)/\theta d\right) \times 100\%$$

where: θm = gravimetric field water capacity,

θw = wet mass of soil, and

θd = dry mass of soil.

3.2.4 Morphological measurements

A range of morphological measurements were recorded for both irrigated (control) and drought pots during this study. The measurements were taken at all three stages; vegetative, reproductive and pod development. For the vegetative stage, the first measurements started at 19 DAS until 33 DAS (two weeks), for stage 2 it was started from 33 DAS until 47 DAS (two weeks). For pod development stage measurements started from 61 DAS until 103 DAS over a span of six weeks. A list of the measurements taken are listed in Table 3.4.

Table 3.4: **List of morphological measurements examined and their brief description**

Morphological measurements	**Method**
Leaf number/plant (LN)	Recorded every week from the start of the drought
Leaf area/plant (cm^2) (LA)	Destructive leaf area measurements using leaf area meter LI-3000C
Plant height/plant (cm) (PH)	Recorded every week from the start of the drought. Measured from the base of the plant at ground level to the highest point of the terminal leaflet
Flower number/plant (FN)	Counted each day from the first day of flowering for the duration of the study.

3.2.5 Physiological measurements

A range of physiological measurements were taken for both irrigated and drought pots during this study. The measurements were done at all three stages; vegetative, reproductive and pod development. For the vegetative stage, the first measurements started at 19 DAS until 33 DAS for two weeks, for reproductive stage two weeks measurements started from 33 DAS until 47 DAS. For pod development stage measurements started from 61 DAS until 103 DAS over a span of six weeks.

3.2.5.1 Gas Exchange

Gas exchange measurements were done using the LI-6400XT photosynthesis system (Li-Cor Biosciences, Nebraska, USA). The system measures the differences of CO_2 and H_2O in the air stream flowing through the leaf cuvette to assess the rate of photosynthetic carbon assimilation and transpiration, respectively [227]. Light, CO_2 concentration, temperature and humidity are controlled manually or automatically. Gas

exchange measurements are used to estimate various parameters of photosynthesis. That includes CO_2 assimilation, stomatal conductance and transpiration as assessed by leaf gas exchanges. The middle leaflet of three fully expanded leaves, per plant, per replicate, were measured between 9.00 am to 12.00 pm. Measurements were done at set values of CO_2 concentration, leaf temperature, humidity and Photosynthetic active radiation (PAR). Reference CO_2 concentration was set to 400 µmol CO_2 m^{-2} s^{-1} while sample humidity was maintained between 45 and 55 percent by controlling either the sample or reference water mole fraction and/or flow of air into the leaf chamber. Leaf temperature was set at 28°C. The value of PAR for measuring gas exchange was set at 1000 (µmol quanta per m^2 leaf per s).

3.2.5.2 Relative water content (%) (RWC)

Relative water content was calculated by the method of Clausen et al. [228]. One middle leaflet of three fully expanded leaves was chosen randomly and harvested from each plant per replicate. Three leaf discs (13 mm diameter) were punched from the leaflet and then placed on a pre-weighed weighing boat to obtain the fresh weight (FW). The leaf discs were placed in petri dish containing distilled water and left overnight under a light source to allow discs to fully hydrate to their turgid weight (TW). Next morning the leaf discs were dried with tissue paper and TW was obtained. The leaf discs were placed in an oven at 80°C for 48 hours to allow dry weight (DW) to be measured. Their RWC was calculated as:

$$RWC\,(\%) = (FW - DW)\,/\,(TW - DW) \times 100$$

Where: FW = fresh weight of leaves; TW = turgid weight of leaves; DW = dry weight of leaves.

In several species, however, changes in dry mass are minimal even after an imbi-

bition period of 24 hours [229]. In order to minimise the imbibition time, leaf disks, instead of whole leaves, have been used, in order to increase the surface area for water intake[228].

3.2.5.3 Leaf Chlorophyll

Leaf Chlorophyll SPAD-2 plus chlorophyll meter was used to measure the chlorophyll content of the leaves. Measurements were taken on three youngest fully expanded leaves on each pot and then averaged to represent the chlorophyll reading. SPAD measurements were done in the morning from 9 am.

3.2.6 Harvest data

A range of harvest data were taken for both irrigated and drought pots at 150 DAS, at the time of harvesting from plants of all 3 developmental stages. List of all the data taken are listed in Table 3.5.

Table 3.5: **List of harvest data taken and their brief description**

Yield measurements	**Method**
Pod. No/plant (PN)	Counted at harvest. Number of pods with more than one seed was also determined
Pod weight/plant (gm) (PW)	Weight of dried pods was recorded after maintaining the harvest pods for three weeks at 37°C
Shoot dry weight/plant (gm) (SDW)	Weight of above ground material after drying for 48 hours at 80°C.
Seed Number/plant (SN)	Counted after removing the shells of all the pods
Seed Weight/plant (gm) (SW)	Weight of seeds was recorded after maintaining the seeds for three weeks at 37°C
Harvest Index (HI)	Seed weight (gm)/Total above ground biomass yield

3.2.7 Statistical analysis of physiology, morphology and harvest data

Analysis of variance (ANOVA) was performed on all the physiology, morphology and harvest data using GenStat 18thedition (VSN International, 2016) to determine the statistical significance between drought and control conditions at different time points and developmental stages for different traits. T-test were also performed to determine statistical significance for pairwise comparison between treatments at each time point and developmental stage. Differences were considered statistically significant when p-value < 0.05.

3.3 Effect of drought stress on the transcriptome of bambara groundnut at different developmental stages

3.3.1 RNA extraction

Total RNA was isolated from leaf tissues sampled at different time points at vegetative and reproductive growth stages mentioned in 3.2.2 using TRIzol reagent (Life Technologies). DNA contamination was treated by DNase I (Qiagen) and column clean-up using the RNeasy minikit (QIAGEN). RNA was extracted at 2 time-points for vegetative and reproductive stage respectively. The timing of RNA extractions at both stages are listed in Table 3.6. For the vegetative stage, 3 leaves from each of 3 biological replicates for controls from DipC and TN were harvested at 19 DAS, followed by 3 leaves each for 3 control and drought replicates at 33 DAS. For the reproductive stage, 3 leaves from each of 3 control replicates were harvested at 33 DAS, followed

by 3 leaves each for the 3 control and drought replicates at 47 DAS. All samples were stored in a -80°C freezer before RNA extraction. Nanodrop and gel electrophoresis were performed to check the quality and quantity of RNA. 1µg of total RNA was sent for RNA-seq from each sample, except DipC Drought 3 - 47 DAS where only 742ng were available.

Table 3.6: **RNA sampling regime**

Stage	**Field capactity (%)**	**DipC**		**TN**	
		Control	**Drought**	**Control**	**Drought**
Vegetative stage					
19 DAS*	75%	3L/Rep*		3L/Rep*	
33 DAS*	25%	3L/Rep*	3L/Rep*	3L/Rep*	3L/Rep*
Reproductive stage					
33 DAS*	75%	3L/Rep*		3L/Rep*	
47 DAS*	25%	3L/Rep*	3L/Rep*	3L/Rep*	3L/Rep*

**3L/Rep = 3 leaves/Replicate, *DAS = Days after sowing*

3.3.2 Quantification of nucleic acid

The concentration and quality of nucleic acid was examined using spectral absorbance ratios and electrophoretically on an agarose gel. Spectral absorbance ratios (A260/280) of DNA and RNA (ng μl^{-1}) were determined using the Nanodrop ND1000 spectrophotometer (Thermo Scientific, USA) using the ND-1000 V 3.7.0 software. The pedestal of Nanodrop was first cleaned with 2 µl sterile water, followed by loading 2 µl samples onto the pedestal for measurement. A ratio of ~1.8 was generally accepted for DNA of good quality whereas ratio of ~2.0 was required for RNA. Furthermore, samples were tested on a Agilent bioanalyser for integrity before preparation for the RNA-seq analysis. RNA samples with 2 µl each were loaded into the PCR tubes and sent to Plant Sciences, The University of Nottingham, Sutton Bonington Campus, UK for Agilent analysis. The size of the 18S peak and 28S peaks were then calculated, a ratio

of 2 is ideal as 28S/18S ratio is one of the key indicators of RNA quality.

3.3.3 RNA-seq library preparation and transcriptome sequencing

The extracted RNA samples were sent to High Throughput Genomics, Oxford Genomics Centre, Oxford, UK for the the library preparation. Library preparation was done in the following step: (i) Preparation of rRNA-depleted or poly(A)-enriched RNA (ii) Fragmentation of RNA, (iii) cDNA synthesis, (iv) Adaptor ligation and (v) Clean up.

Firstly, total RNA quantity and integrity were assessed, using Quant-IT RiboGreen RNA Assay Kit (Invitrogen, Carlsbad, CA, USA) and Agilent Tapestation 2200 R6K. All samples were normalised to 1µg total RNA, except DipC Drought 3 - 47 DAS where only 742ng were available, and used to prepare libraries using Illumina TruSeq Stranded mRNA Library Prep Kit (cat # RS-122-2101) with minor modifications. The polyA containing mRNA molecules were isolated using poly-T oligo attached magnetic beads during the RNA purification step (following the kit's instructions). The following custom primers (25 µM each) were used for the PCR enrichment step: (i) Multiplex PCR primer 1.0 (5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTAC ACGACGCTCTTCCGATCT-3'), (ii) Index Primer (5'-CAAGCAGAAGACGGCATA CGAGAT[INDEX]CAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'). Indices were eight base tags internally [230]. Amplified libraries were analysed for size distribution using the Agilent Tapestation 2200 D1000. Libraries were quantified using Picogreen and relative volumes were pooled accordingly. Sequencing was performed as 75bp paired end reads on a HiSeq4000 according to Illumina specifications.

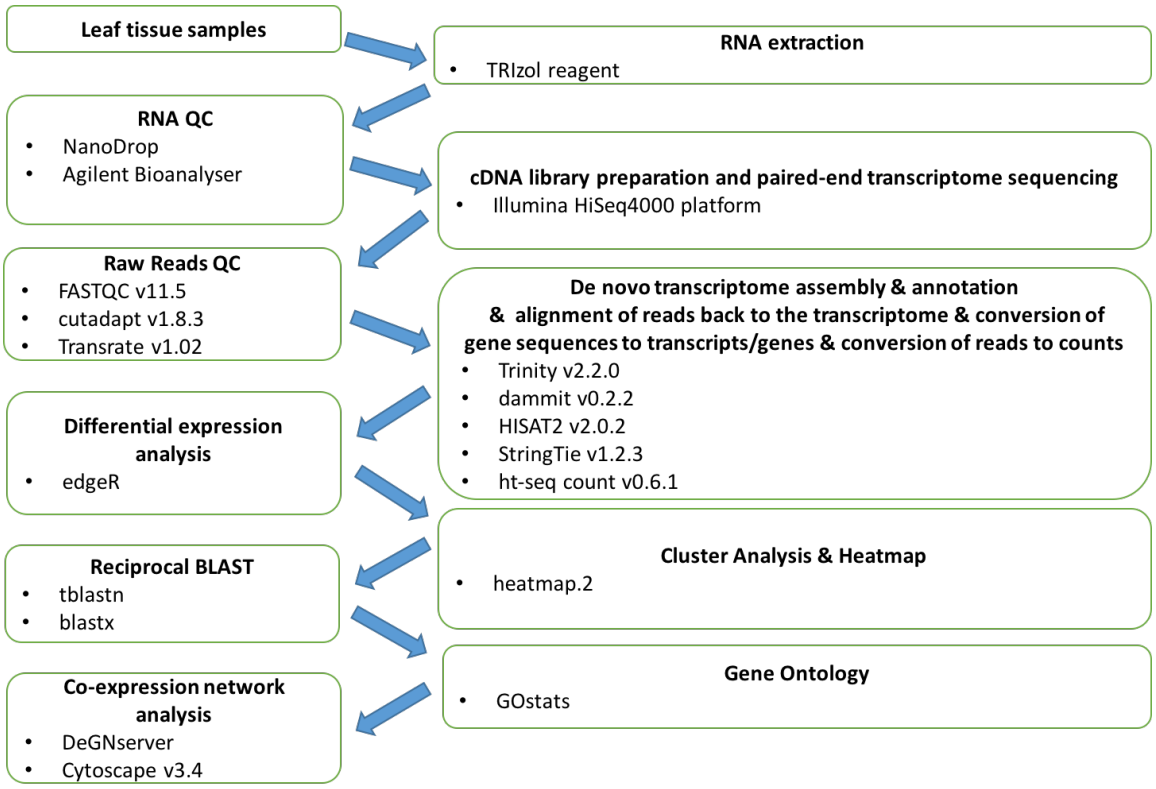

Figure 3.4: **Flow diagram of RNA sequencing analysis**. RNA was extracted from leaf tissue samples of bambara groundnut using TRIzol reagent, and its quality was assessed with NanoDrop and Agilent Bioanalyser. cDNA library preparation and paired-end transcriptome sequencing was performed on the Illumina Hi-Seq Platform, producing paired 75bp reads. Adapters were removed from the raw reads using Cutadapt, followed by QC with FastQC and Transrate. Trinity was used to do the transcriptome assembly. Tool "dammit" was used to annotate the assembled sequences. Trimmed sequences/reads were mapped back to the assembled transcriptome using HISAT2 for expression estimation followed by StringTie to assemble the read alignments into potential transcripts/contigs for each sample which are consistent across the multiple RNA-seq samples. Read counts were generated for these sequences using htseq-count and the output was used for differential expression analysis with edgeR. A heat map was generated using heatmap.2 to obtain a global view of gene expression across samples. To identify GO term for bambara groundnut genes firstly a reciprocal blast search was used to identify putative bambara groundnut orthologs of *Arabidopsis thaliana* and *Glycine max* genes, followed by gene ontology analysis using GOstats. DeGNserver was used for co-expression network analysis followed by network visualisation using cytoscape.

3.3.4 De novo transcriptome assembly, quality checking and annotation

De novo transcriptome assembly and annotation was done by Jo Moreton, ADAC, University of Nottingham, UK. Before transcriptome assembly, all the data files (FASTQ) were quality checked using FASTQC (version 0.11.5) (bioinformatics.babraham.ac.uk). Adaptor trimming was done using cutadapt (version 1.8.3) [231] and PRINSEQ (Standalone lite version) was used to investigate the degree of sequence duplication in a dataset [232]. *De novo* transcriptome assembly was done using Trinity (version 2.2.0) [46] poor quality transcripts were removed by Transrate (version 1.0.2) (bintray.com/blahah/generic). Additionally, transcripts with Transcripts per million (TPM) < 0.5 were removed. TPM values, are normalised for both gene length and sequencing depth so that the sum of all TPM's in each sample are the same allowing more direct comparisons. For transcriptome assembly all the reads were combined into a single input in order to obtain a consolidated set of contigs (transcripts) which will aid in comparing transcriptome across all samples. The FASTA files were combined to create four groups: (i) Control Vegetative (CV, 12 elements), (ii) Control Reproductive (CR, 12 elements), (iii) Drought Vegetative (DV, 6 elements), (iv) Drought Reproductive (DR, 6 elements). To determine a consensus transcriptome from multiple samples, cd-hit-est (CD-HIT version 4.6) (http://weizhongli-lab.org/cd-hit/) was used to combine the four fasta files. A Perl script was used to filter for transcripts present in a minimum of 2/3 in each cluster: (i) At least four sequences in a cluster were required for Drought Reproductive (DR) and Drought Vegetative (DV). (ii) At least eight sequences in a cluster were required for Control Reproductive (CR) and Control Vegetative (CV). These four filtered files were concatenated and then cd-hit-est to generate a non-redundant consensus gene set. The tool "dammit" (version 0.2.2)

(https://pypi.python.org/pypi/dammit/0.3) was used to annotate the assembled sequences.

3.3.5 Read alignment and assembly

Trimmed sequences/reads were mapped back to the assembled transcriptome using HISAT2 (version 2.0.4) (ccb.jhu.edu) for expression estimation. StringTie (version 1.2.3) [233] was then used to assemble the read alignments into potential transcripts/contigs for each sample. This was followed by the StringTie "merge" step to assemble the sequences into non-redundant transcripts which are consistent across the multiple RNA-seq samples. The StringTie merge step produces an annotation (GTF) file which identified 54,225 potential "gene" sequences in total across all samples. Read counts were generated for these 54,225 sequences using htseq-count (version 0.6.1) [234]. Only the primary alignments were used for the read counts. A combined count table for all samples was generated. Absent genes (no read counts in all samples) were removed because this reduces the number of tests for differential expression and therefore the false discovery rate detection.

3.3.6 Normalisation, differential expression analysis

The Bioconductor package "edgeR" was used to find differentially expressed (DE) genes for a number of comparisons [235]. The code to find DE genes was generated by Jo Moreton, ADAC, University of Nottingham, UK which was edited (according to objective) and run by the author. Raw count data was imported into edgeR. A two-step normalisation method was applied on the count data; (i) First, to remove genes with low counts, genes must have at least one count per million (CPM) across all samples (ii) Secondly, trimmed mean of M values (TMM) normalisation was done to re-compute the library sizes for estimating relative RNA levels. The two steps refer to different

aspects of normalisation. CPM normalisation accounts for library size differences between samples, and produces normalised values that can be compared on an absolute scale (e.g., for filtering). TMM method scales reads by the weighted log-fold-change values of a reference sample with genes that have extreme log-fold-changes (M values) and extreme absolute expression levels (A values) removed from the data. TMM normalisation accounts for composition bias, and computes normalisation factors for comparing between libraries on a relative scale. Both steps (CPM and TMM) are needed in the analysis pipeline as CPM normalisation doesn't account for composition bias, and TMM normalisation doesn't produce normalised values [235, 236].

The total number of Differentially expressed (DE) genes at 5% false discovery rate (FDR < 0.05) and log fold change (logFC) ≥ 1 were found. The script used for Differential expression (DE) can be found in Appendix 1. EdgeR uses empirical Bayes estimation and exact test based on the negative binomial distribution with model-based scale normalisation of sequence data, to identify statistically robust DEG, and this is reliable even for data of small sample sizes [235]. Quantile-adjusted conditional maximum likelihood (qCML) method was used to estimate common and tagwise dispersions. Once negative binomial models were fitted, gene-lists were normalised and dispersion estimates were obtained, exact-test based on negative binomial distribution was performed to identify differential expression analysis between groups.

To determine the relative similarities or differences between the samples, two plots were generated, multidimensional scaling (MDS) and MA plot. An MDS plot measures the similarity of the samples and present this analysis in 2-dimensions. Whereas, MA plot shows the relationship between concentration and fold-change across the genes. The differentially expressed genes are colored red and the non-differentially expressed are colored black. Final images were exported as a PNG and PDF file.

3.3.7 Cluster analysis

Hierarchical cluster analysis was performed on normalised CPM values using Bioconductor package heatmap.2 (https://stat.ethz.ch/R-manual/R-devel/library/stats/html/heatmap.html) on differentially expressed genes. Criteria were set to remove all gene expression values below log fold change of < 1. Gene expression values were log transformed and in order to remove statistical noise, these were filtered to remove genes with insufficient read depth or intensity. Heat maps and dendograms were also generated using heatmap.2. Pixel settings were changed to maximise the x- and y-axis, and the final image was exported as a PNG and PDF file. Command line for cluster analysis can found in Appendix 2.

3.3.8 Gene Ontology (GO) Analysis

Gene ontology analysis was done Jo Moreton, ADAC, University of Nottingham, UK. The Bioconductor package "GOstats" (Bioconductor version 3.3) was used to test for over-representation of GO terms using a hypergeometric test (hyperGTest) [237]. First a reciprocal blast search (against *Glycine max* and *Arabidopsis thaliana*) was run to get the information required for GOstats. Secondly, EnsemblPlants BioMart (http://plants.ensembl.org/biomart) was used to get the GO to gene mappings (i) Arabidopsis thaliana genes (TAIR10 (2010-09-TAIR10)), (ii) Glycine max genes (V1.0 (JGI-Glyma-1.1)). Test for over-representation of GO terms was done on the differentially expressed genes at various conditions. GOstats was run separately for the *Arabidopsis thaliana* and *Glycine Max* GO. Differentially expressed gene list reports of GO term enrichments were generated in the categories biological processes (BP), molecular function (MF), and cellular component (CC).

3.3.9 Reciprocal BLAST

Reciprocal blast was used to identify putative bambara groundnut orthologs of *Arabidopsis thaliana* and soybean genes. For the *Arabidopsis thaliana* and soybean complete set proteins were blasted against the transcripts (nucleotide) of bambara groundnut. To compare the *Arabidopsis thaliana* and soybean query protein sequences against the bambara groundnut nucleotide sequence, tblastn was performed. For comparison of reciprocal nucleotide to protein sequence blastx was performed. In each case the best hit was checked to identify reciprocal best hits (not E value threshold was used) as the evidence of a reciprocal best hit is better evidence than an arbitrary cutoff.

Reciprocal blast against soybean resulted in identification of 15056 bambara groundnut orthologs, while 10370 orthologs were identified for bambara groundnut when blasted against *Arabidopsis thaliana.*

3.3.10 Construction of the co-expression network

Co-expression network analysis was carried out using DeGNserver [57] and cytoscape 3.4 [58]. Separate networks were generated for vegetative and reproductive growth stages in both DipC and TN. The input gene-sets were restricted to those that were differentially expressed between every pair of treatments at each time points. TMM-normalised values were used for all samples. Links were assigned between pairs of nodes (i.e. gene-sets) when their Spearman's Rank correlation was 0.95. This threshold value was used because high number of genes were predicted at threshold 0.95 without any loss of information. Once the co-expression network was created, all the up and down regulated genes under drought stress at different time points and growth stages for both DipC and TN were chosen to visualise and identify relevant the co-expression network. The co-expression network was imported into cytoscape for visual representation and

network analysis. Furthermore, the separate networks created for each genotype were merged in each growth stage (vegetative and reproductive) to identify cases where co-expression between pairs of nodes is detected in both genotypes. In addition, a separate merged network was created to identify cases where co-expression between pairs of nodes is detected in both growth stages.

3.4 Identification of cross species syntenic locations of drought-related genes

The putative syntenic blocks of Gene Expression Marker (GEM) linkage markers were identified by mapping the sequence tag associated with the DArT Seq markers used to construct the genetic maps in bambara groundnut (generally 64 nucleotides) to the common bean genome using CLC Genomic Workbench v7.5 (Qiagen) default settings except with 'ignore' for 'non-specific match. The mapping results were inspected manually to ensure good corresponding locations with the neighboring markers. Similar approaches were applied to the genes of interest selected from the microarray and RNA-seq results, discussed in Chapter 4 and 6, except the 'length fraction' was set to 0.3 with manual checking to ensure good alignment at the exons regions.

3.5 Software tools and databases

Several software/tools and databases were used throughout the study. Description on the application of each tools and databases in the study are presented in Table 3.7 and 3.8.

Table 3.7: **List of all the tools used during the study**

Name	Application in thesis	Source
GeneSpring	Transcriptome analysis and visualisation of cross species microarray data of bambara groundnut	www.genomics.agilent.com
Cytoscape	Visualizing of co-expression networks	http://www.cytoscape.org/
PlantPan	For detecting transcription factor binding sites	http://plantpan2.itps.ncku.edu.tw/
Multialin	Multiple sequence alignment	multalin.toulouse.inra.fr/multalin/
EdgeR	For differential expression analysis for RNA-seq data	https://bioconductor.org/packages/edgeR
DeGNserver	Co-expression network analysis of cross-species microarray data of bambara groundnut	http://plantgrn.noble.org/DeGNServer/
Primer-BLAST	To find primers for genes of interest from cross-species microarray work	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
BINGO	To calculate overrepresented GO terms in cross-species micrarray work	http://apps.cytoscape.org/apps/bingo
HISAT2	Read alignment of RNA-seq data	https://ccb.jhu.edu/software/hisat2/index.shtml
PRINSEQ	Quality control and RNA-seq data preprocessing	http://prinseq.sourceforge.net/
Transrate	For de-novo transcriptome assembly quality analysis	http://hibberdlab.com/transrate/
StringTie	Assembly of RNA-Seq alignments into potential transcripts	https://ccb.jhu.edu/software/stringtie/
Trinity	De-novo assembly/alignment of RNA-seq data	https://trinityrnaseq.github.io/
GOstats	Testing on GO terms in bambara groundnut RNA-seq data using resources from arabidopsis and soybean.	https://bioconductor.org/packages/release/bioc/html/GOstats.html
HTseq-count	Transcript counting of RNA-seq data	www-huber.embl.de/HTSeq/
Cutadpat	Trim adapters from high-throughput sequencing reads of RNA-seq data	https://pypi.python.org/pypi/cutadapt
CLC	Identifying physical location of each drought-related gene in common bean and mapping to the GEM linkage map	http://www.clcbio.com/

Table 3.8: **List of all the databases used during the study**

Databases	**Application in thesis**	**Source**
BioGrid (build 3.4.140 216)	To create putative bambara groundnut network	thebiogrid.org/
BioMart	Used to get the gene ontology to gene mappings in bambara groundnut from arabidopsis and soybean.	www.biomart.org/
SoyKB	To find gene sequence and upstream regions of gene of interest.	soykb.org/
EnsemblPlants	To look for homologues for genes of interest and identify gene location	plants.ensembl.org/
LegumeIP	To study gene function	plantgrn.noble.org/LegumeIP/
NCBI	BLAST was performed to find regions of similarity between biological sequences	www.ncbi.nlm.nih.gov/

Chapter 4

Bambara groundnut microarray XSpecies analysis

4.1 Introduction

The ability to grow crops under water limited conditions is a significant factor in relation to global food security. Drought is one of the major abiotic stresses that inhibits plant growth and can reduce crop productivity. Hence, drought tolerance is a key target in ensuring global food supply. Details on different drought mechanisms and how drought effect bambara groundnut and other crops in relation to physiology, morphology and transcriptional response has been discussed in 2.1.4. Bambara groundnut is an underutilised crop grown by subsistence farmers in Africa as it is known to grow well in regions of water deficit (see 2.1.1 for details).

This study focuses on the analysis of the transcriptomic changes in two bambara groundnut genotypes; DipC and TN (Tiga Nicuru), derived from different landraces in response to drought. The contrast between the two parental lines for a number of traits such as days-to-maturity, stomatal conductance, 100-seed weight, leaf area, internode length, peduncle length, pod number per plant, and leaf carbon (Delta C13) isotope analysis, suggest that some of these mechanisms for adaption to drought could be non-identical in the two genotypes. Chai et al. [12, 24], reported that transgressive

segregation was observed in the segregating F_5 population derived from the TN X DipC cross. The results showed that there were lines in the segregating population that performed better in terms of the ability to produce higher yields under drought conditions than the parental genotypes. In addition, significant quantitative trait loci (QTL) were mapped for stomatal conductance and leaf carbon (Delta C^{13}) isotope analysis (CID) in the drought treatment F_5 population, but not in the irrigated population, suggesting the drought response between the two parental genotypes differs for these traits which are segregating in the F_5 population [238]. Hence, evaluating the transcriptome of the two parental lines under drought stress could be a good indicator to investigate the molecular mechanism occurring in the two genotypes and its relationship to phenology and phenotype.

A cross-species hybridisation approach based on the Soybean Affymetrix GeneChip array has been employed. The study aims to identify genes and expression patterns which enable bambara groundnut to grow under semi-arid conditions, compare the transcriptome of the genotypes of bambara groundnut to identify what is common and how they differ, investigate if the cross-species hybridisation of bambara groundnut with the Affymetrix Glycine-max GeneChip will work and understanding the drought mechanisms underlying the ability of crops to produce viable yields under drought conditions.

4.2 Results

4.2.1 Probe selection based on gDNA

Genomic DNA (gDNA) of both genotypes was hybridised to the Affymetrix Soybean GeneChip array to study global genome hybridisation for probe selection (see 3.1.10). The numbers of retained probe-pairs and probe-sets is shown in Table 4.1. With the increase in threshold values, the number of probe pair retained in the probe mask file started decreasing rapidly (Fig. 4.1), while the number of probe-sets (genes) decreases at a slower rate. This suggests that even at higher gDNA hybridisation thresholds, at least some of the gene-designed oligonucleotides are cross-hybridising for many of the probe-sets and that the cross-species array approach is a reasonable approach for bambara groundnut transcriptomics.

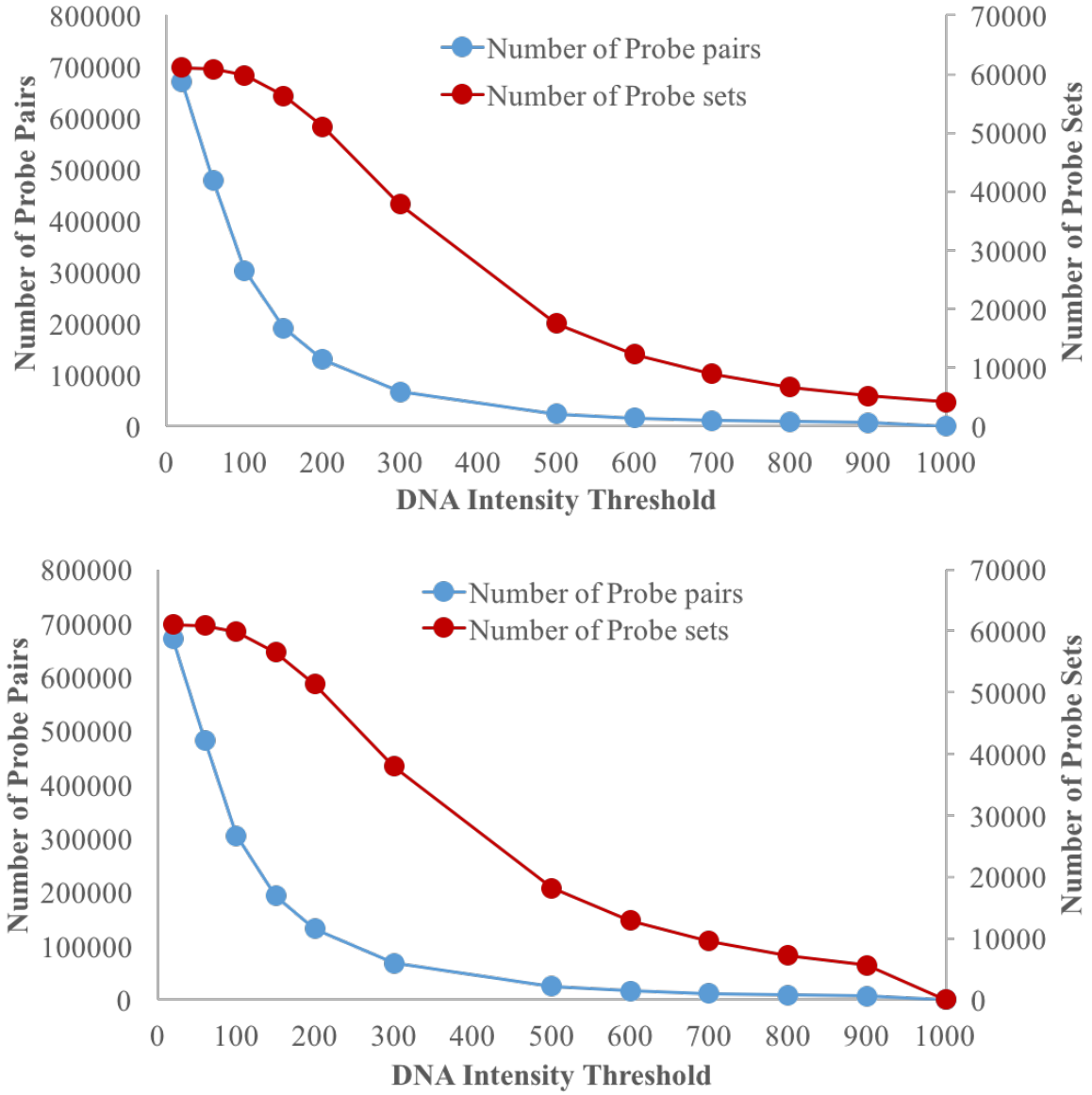

Figure 4.1: **Effect of intensity thresholds**. Number of probe pairs (blue line) and probe sets (magenta line) retained for DipC (top) and Tiga Nicuru (bottom) respectively at different gDNA intensity thresholds.

The number of retained probe-sets and probe-pairs on Soybean chip for both the DipC and TN gDNA hybridisation were determined corresponding to each threshold value (Table 4.1). A custom CDF file of threshold 100 was chosen for differential expression analysis in both genotypes for further analysis as it allowed good sensi-

tivity to detect maximum number of differentially-expressed transcripts (Table 4.1). Furthermore, both genotypes were found to be highly similar in terms of probe-sets detected at this threshold. A total of 59,533 probe-sets were common to both genotypes at threshold of 100, while 249 and 302 probe-sets were specific to DipC and TN respectively. Thus, suggesting high sequence similarity (>99%) at this level of sequence resolution.

Table 4.1: **Retained probe-sets and probe-pairs at different threshold value.**

Threshold value	Number of probe sets (Soybean chip hyb. to DipC gDNA)	Number of probe sets (Soybean chip hyb. to TN gDNA)	Number of probe pairs (Soybean chip hyb. To DipC gDNA)	Number of probe pairs (Soybean chip hyb. To TN gDNA)	Number of DEGs in DipC	Number of DEGs in DipC
20	61072	61072	670388	670388	6165	6165
60	60877	60895	479538	482352	6927	6814
100	**59782**	**59835**	**302834**	**304708**	**7183**	**7159**
150	56266	56511	190570	193522	7036	7159
200	51071	51319	129806	132521	6638	6731
300	37813	38000	66907	68106	5275	5345
500	17469	18176	23464	24693	2784	2911
600	12258	12930	15701	16650	2089	2170
700	8896	9566	11193	12061	1574	1673
800	6687	7208	8415	9070	1195	1291
900	5140	5657	6559	7140	958	1057
1000	4085	4482	5304	5733	802	877

4.3 Principal Component Analysis

The PCA plot (Fig. 4.2) shows that, under irrigation, the two genotypes appear to have similar transcriptomes. The first two Principal Components account for 25.45% and 17.11% of the variance respectively, suggesting that it is due to a range of hybridisation/expression differences between the chips. Recovery after drought, however, caused the most variation and suggests that the recovery transcriptome does not return

to the irrigated state (control). The DipC drought treatment sample 'D.DipC.Rep2' could be a potential outlier and this needs to be borne in mind in further analysis. The 3D PCA plots of genotype-specific data showed good separation of the three treatments (conditions) and better PCA scores (see Appendix 3).

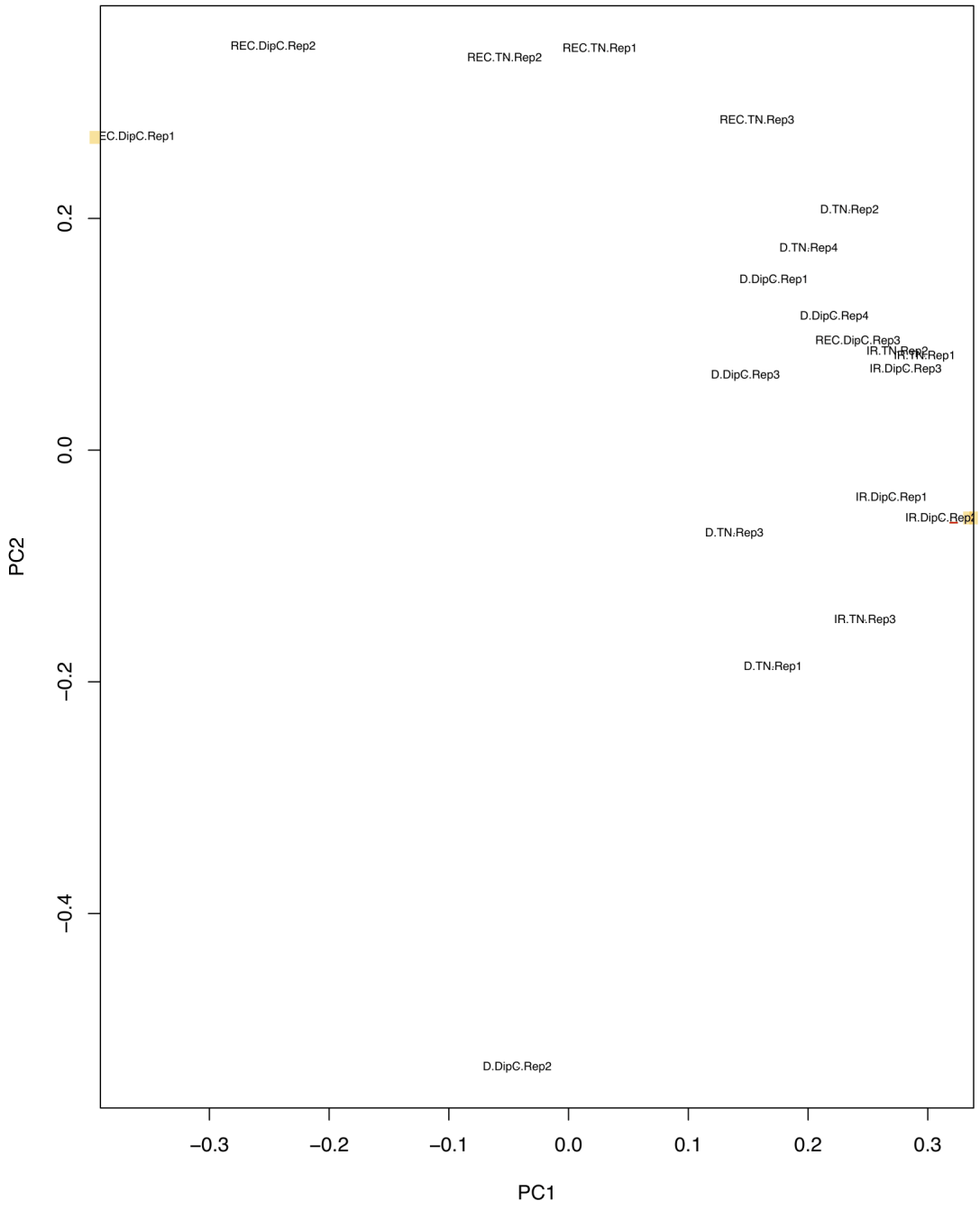

Figure 4.2: **PCA plot of the expression data from the microarrays**. The PC1 and PC2 values for each chip has been placed on a scatter plot. Each chip result is defined by a three-part character string consisting of the treatment, landrace and replicate number. IR, D and REC refer to irrigation, drought and recovery; the landraces are DipC and TN; and Rep1-4 refers to the specific biological replicate. Note, Irrigation and Recovery treatments have only 3 replicates while drought treatment has 4.

4.4 Gene expression under irrigated conditions

It is pertinent to consider the state of the genotype transcriptomes before any drought treatment has taken place. However, owing to the high background noise in microarray studies, it is unclear what intensity level defines a gene as being transcribed. Figure 4.3 shows that the ranked intensity values follow a roughly sigmoidal curve. The point of inflection (at which the declining gradient is at its shallowest) covers the top two-thirds of probe sets, and corresponds to an RMA (Robust Multi-array Average) value of 0.97. This may be a stringent cut-off, given that an RMA value of 1 corresponds to the average across all probe sets on the array, but it ensures that there were few, if any, false positives. This left 39855 probe sets for DipC and 39890 for TN. There are 26496 probe sets in common between the two genotypes, suggesting differences in general transcriptional behavior between the two genotypes.

Each genotype had a little over 90 probe sets with functional annotations related to ABA signaling and drought responses (see Appendix 4 and 5) of which 60 were common to both (Appendix 6). These include homologues of much of the ABA synthesis and response network, the *DREB1* transcription factor, Early-Response to Dehydration proteins 3, 4, 8, 14-16 and 18, four osmoprotectant genes, two drought-response genes influencing photosynthesis, and 21 other probe sets corresponding to drought-related proteins of unknown function. Appendix 7 lists the genes differentially expressed between the two irrigated genotypes, but at this stage nothing stands out as remarkable.

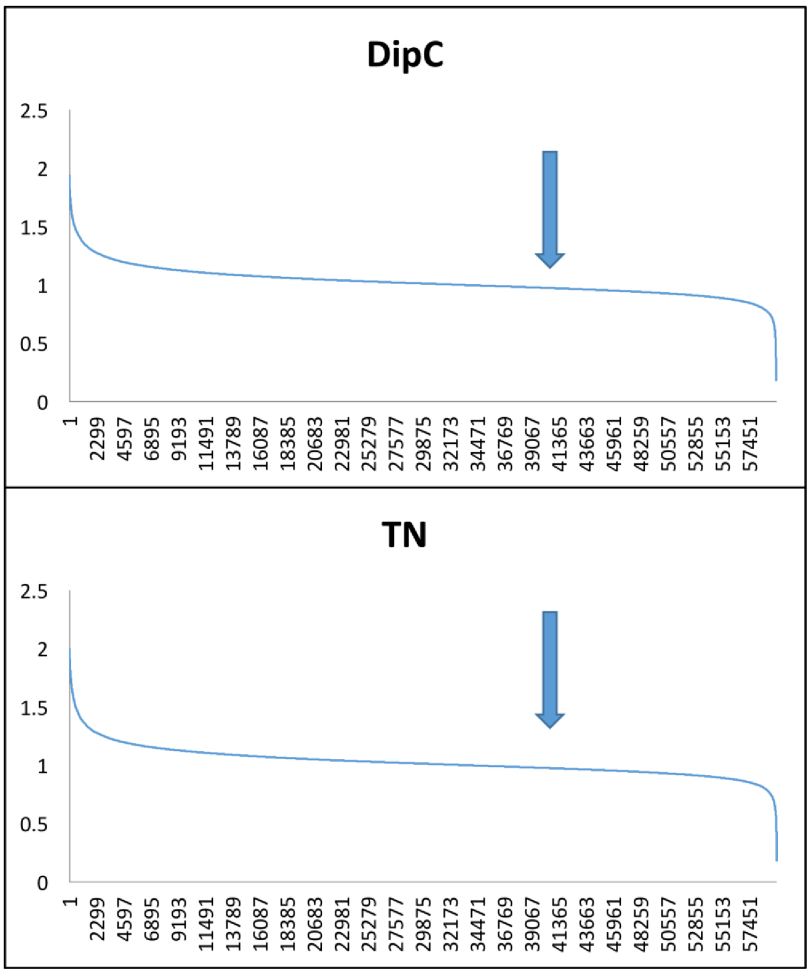

Figure 4.3: **Ranked mean intensities of irrigated genotype samples.** The mean RMA values for the irrigated probe-sets have been ranked highest to lowest and then (with the exception of 102 probe sets with values >2) plotted as shown. The arrows mark the position of the top 66% of probe-sets, and correspond to RMA values of 0.97. For DipC, 59782 probe sets are reduced to 39855. For TN, 59835 probe sets are reduced to 39890.

4.5 Identification of differentially expressed genes

For DipC and TN, the numbers of genes differentially expressed (DE) as a result of drought and recovery treatments as detected by the cross-species microarray approach

are shown in Table 4.2, with the full lists of probe sets and functional annotations on Appendix 8-11. The method used to identify DE genes are described in 3.1.11 The top up and down regulated genes in DipC and TN are shown in Table 4.4 and 4.5 respectively. The numbers for DipC were consistently higher than for TN, and drought caused more down- than up-regulation while recovery had the reverse effect.

Recovery led to many more differentially expressed genes (486 and 391) than drought (189 and 81). There were six possible system effects that can be gleaned from these data (see Fig.4.4). The upregulated genes under drought that return to normal on recovery and down regulated genes returned to normal expression at recovery are the strictly drought-responsive genes (~75% in both genotypes), while those that significantly changed and did not return to pretreatment levels (~25%) correspond to a drought-induced state change. The latter may be due to epigenetic effects, such as a change in the methylation state of gene-regulatory regions. The larger numbers of differentially expressed genes from drought to recovery may be accounted for by aging and other highly variable factors (see Fig.4.2), such as soil conditions in each pit.

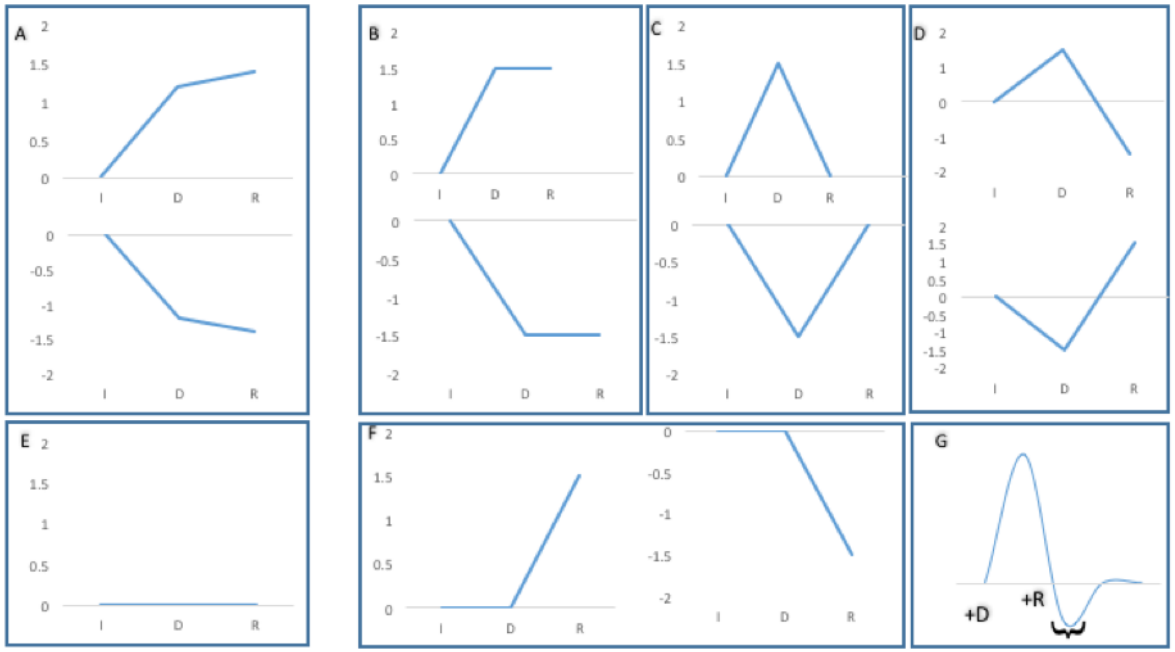

Figure 4.4: **Categories of system response in gene expression..** The panels A-F show the alternative ways in which gene expression may be affected by drought, recovery and ageing; A-D and F include examples of both significant activation and repression. The vertical axes refer to increase (+) and decrease (-) in expression levels with respect to. the irrigated samples. The categories are A: drought and age perturbation, B: drought-induced state change (i.e. the level of gene expression has been permanently altered by the drought treatment but not by subsequent aging), C: Drought response, D: Drought response with recovery overshoot, E: constitutive (i.e. no effect), and F: age perturbation. Panel G is a cartoon of the change in levels of gene expression over time in response to drought induction (+D) and recovery (+R). The level may overshoot and fall below untreated levels at early times after the start of recovery, denoted by the curly bracket.

Table 4.2: **Differentially-expressed gene numbers**

	Drought versus Irrigation				Drought versus Recovery	
	UP-regulated under drought	Down-regulated under recovery	Down-regulated under drought	UP-regulated under recovery	UP-regulated	Down-regulated
DipC	80	68	109	94	340	146
Tiga Nicuru	28	22	53	42	294	97

The fold changes of the up-regulated genes under drought stress in both genotypes are relatively small (mostly $<$ 4-fold), suggesting this drought treatment was not perceived as severe. Furthermore, there were only nine differentially expressed genes in common to both genotypes (see Table 4.3). The only common up-regulated gene was Beta-fructofuranosidase, which hydrolyses sucrose to provide more glucose, hence playing a potential role in osmoprotection. In contrast, half of the common down-regulated genes were related to transcription and also play roles in stomatal regulation. Excluding the potential outlier 'D.DipC.Rep2' had little effect upon the common gene analysis (see Appendix 12), so it has been included in subsequent analyses.

Mostly, the up-regulated gene under drought stress in DipC relate to secondary metabolism of cell-wall components, while the TN genes include transcription-related factors, most notably a CONSTANS-like gene. Furthermore, GO term over-representation analysis for both DipC and TN showed an emphasis on various metabolic processes, related to cellular amino acids and their derivatives, secondary metabolites and carbohydrates (Table 4.6). Hence, despite having highly similar probe-sets, there is a very different transcriptional response to drought stress by each genotype. Microarray data has a limited dynamic range, even when within species, so it is important to validate a small set of microarray observations. Hence, validation through qRT-PCR was performed.

Table 4.3: **Overlapping up and down regulated genes**

Gene Name	**P-value**	**Fold change**	**Gene Description**	**References**
UP-regulated gene				
Beta-fructofuranosidase	8.90E-04	3.1939	Catalyzes the hydrolysis of sucrose. A rise in monosaccharide content caused by the Beta-fructofuranosidase can compensate for the decline in photosynthetic carbon assimilation indicated by the decrease in net photosynthesis	[239, 240]
Down-regulated genes				
MEE59 (maternal effect embryo arrest 59)	8.94E-04	8.5802	Embryo development ending in seed dormancy	
Calcineurin-like phosphoesterase family protein (CPPED1)	6.72E-04	5.8575	Plays inhibitory role in glucose uptake. Down-regulation of CPPED1 improves glucose metabolism.	[241]
Putative lysine-specific demethylase JMJD5 Jumonji/Zinc-finger-class domain containing protein	0.0031	4.9714	Plays role in a histone demethylation mechanism that is conserved from yeast to human. Down-regulation may lead to an increase in methylated histones and hence general down-regulation of transcription	[242]
MYB-like transcription factor	0.0244	4.1039	Arabidopsis homolog is known to regulate stomatal opening, flower development, and plays role in circadian rhythm	[243]
F-box family protein (FBL14)	0.0016	3.7447	Functions in signal transduction and regulation of cell cycle	
BRH1 (BRASSINOSTEROID-RESPONSIVE RING-H2)	0.0077	2.8992	BRH1 is known to influence stomatal density	[244]
Bundle-sheath defective protein 2 family/bsd2 family	0.0031	2.4419	Protein required for post-translational regulation of Rubisco large subunit (rbcL). This may be one mechanism by which photosynthesis is slowed down.	[245]
Mitochondrial substrate carrier family protein	0.03	2.4353	Involved in energy transfer	

Table 4.4: **Top up-regulated genes under drought in DipC and TN**

Gene Name	**P-value**	**Fold change**	**Gene Description**	**References**
UP-regulated genes in DipC				
PAL1 (Phenylalanine ammonia-lyase 1)	0.018	3.9018	Key enzyme involved in the biosynthesis of isoprenoid antioxidative and polyphenol compounds such as lignin and is involved in defense mechanism.	[246]
ATEP3/AtchitIV	0.0013	3.8454	Encodes an EP3 chitinase that is stimulated under abiotic stress.	[247]
TXR1(Thaxtomin A resistant 1)/ATPAM16	6.87E-05	3.7188	TXR1 is a component of a dispensable transport mechanism. Involved in negative regulation of defense responses by reducing reactive oxygen species (ROS).	[248]
Acetyl-CoA C-acyltransferase, putative / 3-ketoacyl-CoA thiolase	0.0018	3.5548	Functions in Jasmonic acid synthesis which plays a role in plant response to mechanical and abiotic stress.	[249]
UBC-2 (ubiquitin-conjugating enzyme 2)	0.0044	3.407	Ubiquitination plays a part in increasing rate of the protein breakdown. Arabidopsis plants overexpressing UBC-2 were more tolerant to drought stress compared to the control plants	[250]
Rho GDP dissociation inhibitor 2	0.0013	3.348	Involves in the regulation of Rho protein and small GTPase mediated signal transduction.	[251]
Histidine amino acid transporter (LHT1)	0.0013	3.2566	Amino acid transmembrane transporter involved in apoplastic transport of amino acids in leaves.	[252]
COMT (3-Caffeic acid o methyltransferase)	0.0063	3.2347	Involved in lignin biosynthesis. High activation of lignifying enzymes was found in drought-stressed white clover (Trifolium repens L.), which lead to reduced forage growth.	[253]
Glycine decarboxylase complex H	0.0057	3.1136	Functions in photo respiratory carbon recovery. Carbon dioxide is found to be low in plants subjected drought stress due to the closing of stomata in order to prevent water loss.	[254]
Up-regulated genes in TN				
Clp amino terminal domain-containing protein, putative	0.0353	3.7787	Protein and ATP binding.	
CONSTANS-LIKE 1	0.0259	3.2941	Transcription factor regulating flower development and response to light stimulus.	[255]
DRB3 (DSRNA-BINDING PROTEIN 3)	0.0209	2.9844	Assists in miRNA-targeted RNA degradation.	[256]
SIGE (SIGMA FACTOR E)	0.0322	2.8084	Responds to affects of abiotic stresses. Phosphorylation of major sigma factor SIG1 in Arabidopsis thaliana inhibits the transcription of the psaA gene, which encodes photosystem-I (PS-I). This disturbs photosynthetic activity	[257, 258]

Reticulon family protein	0.0298	2.7725	Playing role in promoting membrane curvature
Cytochrome c oxidase family protein	0.0259	2.7279	Essential for the assembly of functional cytochrome oxidase protein.
DNA-binding S1FA family protein	0.0491	2.7174	Binds to the negative promoter element S1F
DNA photolyase	0.0321	2.667	DNA repair enzyme
Zinc knuckle (CCHC-type) family protein	0.0405	2.5674	Zinc ion binding
Monosaccaride transporter	0.0259	2.5474	Plays a role in long-distance sugar partitioning or sub-cellular sugar distribution
Nodulin MtN3 family protein	0.0259	2.3766	Key role in the establishment of symbiosis
Serine acetyltransferase, N-terminal	0.04	2.302	Catalyzes the formation of a cysteine precursor

Table 4.5: **Top down-regulated genes under drought in DipC and TN**

Gene name	**P-value**	**Fold Change**	**Gene description**	**References**
Down-regulated genes in DipC				
Dihydroxyacetone kinase	0.0031	6.4898	Glycerone kinase activity	
Phosphoglucomutase, putative / glucose phosphomutase, putative	0.0077	6.4718	Involved in controlling photosynthetic carbon flow and plays essential role starch synthesis. Down regulation of photosynthesis-related gene will lead to significant reduction in plant growth.	[259]
Auxin-induced protein 22D AUXX-IAA	0.0036	4.6273	Involved in stress defense response. Many AUXX-IAA genes were found to be down-regulated in Sorghum bicolor under drought conditions.	[260]
CP12-1, putative	0.0146	4.3904	Involved in calvin cycle, therefore linked to photosynthesis. Most drastic down-regulated genes which were photosynthesis-related was observed in barley (Hordeum vulgare L.)	[261]
PHS2 (ALPHA-GLUCAN PHOSPHORYLASE 2).	0.0146	4.3757	Encodes a cytosolic alpha-glucan phosphorylase.	
APRR5 (PSEUDO-RESPONSE REGULATOR 5), Pseudo ARR-B family	0.001	4.1458	Linked to cytokinin-mediated regulation	
Thiamine biosynthesis family protein	0.0025	4.1323	Catalyzes the activation of small proteins, such as ubiquitin or ubiquitin-like proteins.	
Zinc finger (C3HC4-type RING finger)	0.0071	3.6115	Mediate ubiquitin-conjugating enzyme (UBC-2) dependent ubiquitation.	[262]
WRKY40	0.0331	3.1049	Regulator of ABA signaling. It inhibits the expression of ABA-responsive genes ABF4, AB14, AB15, DREB1A, MYB2 and RAB18.	[263]
Down-regulated genes in TN				
AGL83 (AGAMOUS-LIKE 83)	0.02599	4.3746	DNA-binding transcription factor	
CRR23 (chlororespiratory reduction 23)	0.0259	3.6257	A subunit of the chloroplast NAD(P)H dehydrogenase complex, involved in PS-I cyclic electron transport. Located on the thylakoid membrane. Mutant has impaired NAD(P)H dehydrogenase activity. Part of drought repressing photosynthesis.	[264]
MYB30 (MYB DOMAIN PROTEIN 30)	0.0321	3.2508	Acts as a positive regulator of hypersensitive cell death and salicylic acid synthesis. Involved in the regulation of abscisic acid (ABA) signaling.	[265]

Photosystem II family protein, putative	0.0292	3.1588	Linked to photosynthesis. Down-regulation of photosynthesis-related genes during drought stress was observed in maize (Zea mays), which in turn leads to significant reduction in plant growth.	[266]
Phosphoesterase	0.0472	3.1362	Hydrolase activity, acting on ester bonds	
Zing-finger (C3HC4-type)	0.045	2.9477	Mediate ubiquitin-conjugating enzyme (UBC-2) dependent ubiquitation.	[262]
NHX2 (Sodium proton exchanger 2)	0.04	2.7422	Involved in antiporter activity. Also involved in potassium ion homoeostasis and regulation of stomatal closure. Involved in the accumulation of K+ that drives the rapid stomatal opening. Down-regulation of genes related to stomatal regulation has been observed in soybean, which appears to be a part of drought response, leading to a reduction in the amount of stomata in leaves.	[267, 268]
Inositol 1,3,4-trisphosphate 5/6-kinase	0.0351	2.0903	Part of IP3 signal transduction pathway	[269]

Table 4.6: **GO-term overrepresentation of all the gene-sets in Soybean GeneChip array to compare DipC, TN and Soybean datasets**

GO-ID	Description	FDR corrected p-value
Soybean		
6519	cellular amino acid and derivative metabolic process	6.68E-09
44281	small molecule metabolic process	6.68E-09
5975	carbohydrate metabolic process	1.13E-03
6091	generation of precursor metabolites and energy	2.97E-03
44237	cellular metabolic process	2.19E-02
9058	biosynthetic process	4.83E-02
DipC		
6091	generation of precursor metabolites and energy	1.04E-03
6519	cellular amino acid and derivative metabolic process	1.04E-03
44281	small molecule metabolic process	1.04E-03
19748	secondary metabolic process	1.98E-02
TN		
5975	carbohydrate metabolic process	1.78E-03
6519	cellular amino acid and derivative metabolic process	1.78E-03
44281	small molecule metabolic process	1.78E-03
6091	generation of precursor metabolites and energy	1.23E-02
19748	secondary metabolic process	4.98E-02

4.6 Confirmation of candidate drought-related genes by real-time qRT-PCR

qRT-PCR analysis was performed using the methods described in 3.1.14. Four differentially expressed genes (*PAL1*, Beta-fructofuranosidase, *COMT* and *UBC-2*) were chosen for further analysis as they showed high levels of expression under drought

([239, 246, 250, 253]) (Table 4.3 and 4.4). Figure 4.5 shows the results of q-PCR analysis. The transcript levels of Beta-fructofuranosidase, *COMT* and *UBC-2* confirmed the expression trends seen in the microarray data. *PAL1* showed the expected increase in DipC, but also an increased in TN was observed, which was not observed in the microarray results. The reason for this is unclear.

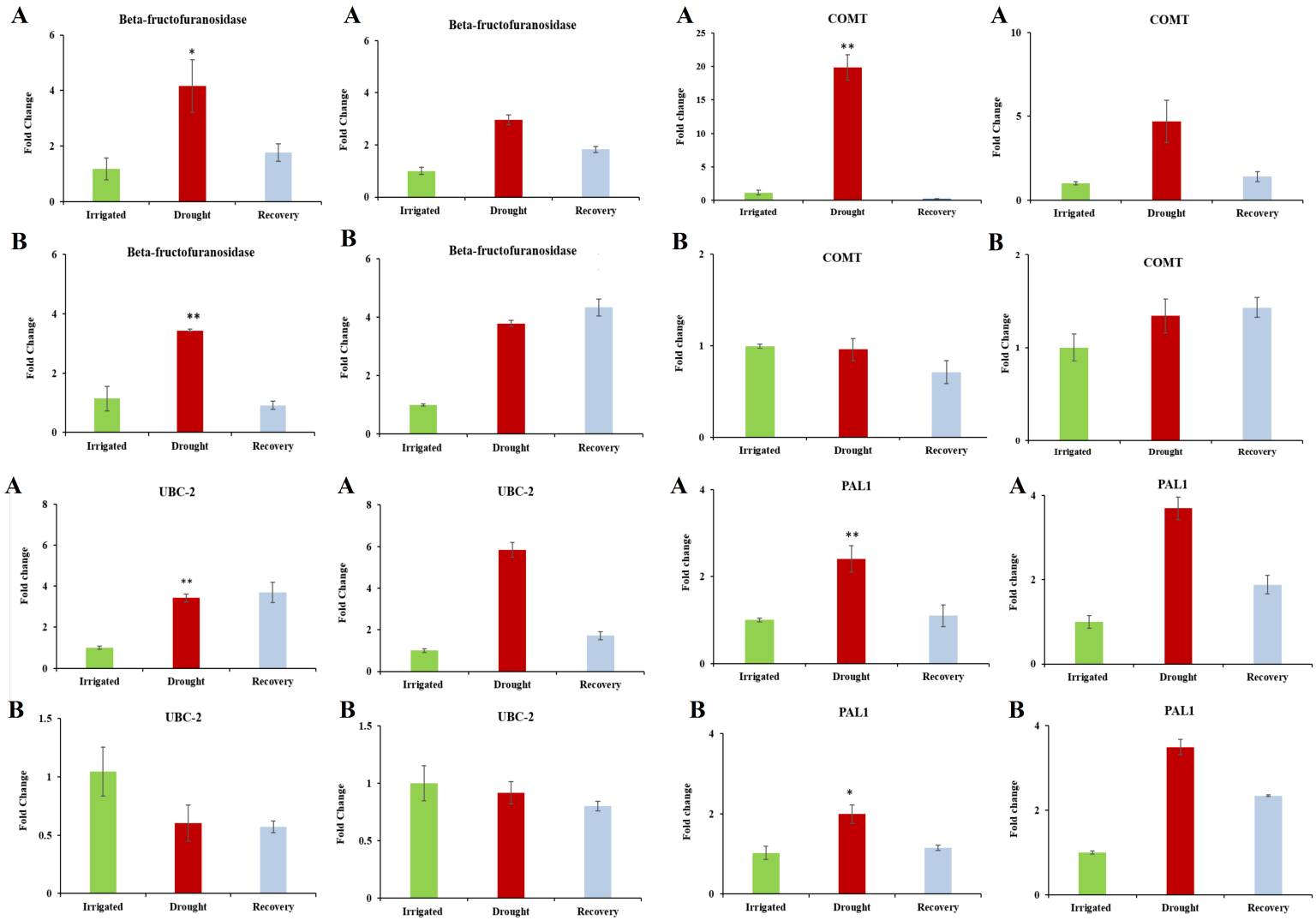

Figure 4.5: **Comparison of qPCR and microarray intensity values**. Rows [A] and [B] respectively refer to results for DipC and TN. The left and right-hand pairs of columns correspond to the qPCR and microarray values for DipC and TN respectively. The gene under study is named at the top of each panel. In order, the investigated genes are Beta-fructofuranosidase, *COMT*, *UBC-2* and *PAL1*. qPCR and Microarray values are shown as fold changes with respect to the irrigated (control) treatment. Error bars denote the standard error. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired T-test between groups.

4.7 Transcription factors associated with drought stress

The DE genes from both genotypes identified various transcription(-related) factors (TFs) from co-expression network analysis mentioned in 3.1.13. Common to both genotypes are the downregulation of *BRH1*, a *MYB*, *MEE59* and *JMJD5*. The latter is a histone demethylase and, hence, plays role at the epigenetic level. Its downregulation could result in indirect repression of multiple genes. On top of the common genes, DipC shows upregulation of 2 TFs (*WRKY51* and a *bHLH* TF) and downregulation of 4 others (*ATAUX2-11*, *WRKY40*, a C2H2 Zn-finger and 3 probe-sets for GIGANTEA). TN, on the other hand, shows upregulation of genes for CONSTANS-like 1, S1FA DNA-binding, and a double-strand RNA-binding protein (which can aid microRNA-mediated RNA degradation). The downregulated TFs in TN are *MYB60* and a second *MEE59*.

Co-expression networks were built for DipC and TN separately (see Appendix 13 and 14), and the drought-specific network of each were merged. This resulted in more TFs being included, which are features of recovery treatment. By looking at the number of links that each node has in the genotype-specific and merged networks, it is possible to rank the potential importance of the different TFs (see Table 4.7). The DipC TFs had higher numbers of links than TN, reflecting the higher number of differentially-expressed genes. In the case of DipC, *WRKY40* stands out as being the TF with the most co-expressed genes, with *ATAUX2-11*, *PRR7* and a Zinc-finger protein (GmaAffx.33796.3.S1_at) also looking relevant. For TN, however, the TFs are not linked so highly, with CONSTANS-like 1 and *MYB60* showing the greatest involvement. For this genotype the differentially-expressed TFs in common with DipC

seem almost as important.

Table 4.7: **Vertex degrees of differentially expressed Transcription factors**

	DipC				**TN**		
Probe-set	**Name**	**V°* (whole)**	**V° *(drought)**	**Probe-set**	**Name**	**V° *(whole)**	**V° *(drought)**
Gma.16733.1.S1_at	WRKY40	68	17	GmaAffx.45249.1.S1_at	CONSTANS-like 1	16	3
Gma.6670.1.S1_at	PRR7	49	7	GmaAffx.84566.2.S1_at	MYB60	8	3
GmaAffx.33796.3.S1_at	Zinc-finger like C2H2	45	7	GmaAffx.86517.1.S1_at	AGL83	6	1
GmaAffx.92679.1.S1_s_at	ATAUX2-11	41	9	Gma.1576.1.S1_at	Zinc-finger C3HC4	5	1
GmaAffx.35309.1.S1_s_at	GRF2	35	6				
GmaAffx.65059.1.S1_at	bHLH	32	7				
GmaAffx.90399.1.S1_at	C3HC4 Zinc-finger	31	9				
Gma.15774.1.S1_at	Zinc-finger C3HC4	26	3				
GmaAffx.53180.1.S1_at	PRR7	25	9				
GmaAffx.80492.1.S1_at	PRR5	9	2				
GmaAffx.73009.2.S1_at	WRKY51	7	5				
Common TFs							
GmaAffx.60283.1.S1_at	BRH1	42	6				
GmaAffx.9286.1.S1_s_at	MYB	27	4				
Gma.17248.1.A1_at	JMJD5	26	3				
GmaAffx.10162.1.S1_at	MEE59	13	3				

**(V° refers to the number of links of each TF node, in either the whole landrace-specific network, or merged drought-specific network)*

4.8 Discussion

Landraces are a potentially valuable resource for finding genes conferring useful traits. Bambara groundnut is an underutilised African legume whose landraces are adapted

to semi-arid conditions. We have developed single genotypes developed from landraces for analysis. There have been several drought studies carried out on bambara groundnut, but the molecular mechanisms of how it responds and adapts to drought is still under investigation. This study has carried out transcriptomic comparison in two genotypes of bambara groundnut; DipC and TN, in an attempt to identify potential genes conferring advantageous traits.

Cross-species hybridisation to the soybean microarray has been shown to be informative for investigating the bambara groundnut transcriptome as good gene (probe-set) retention was observed at high gDNA hybridisation thresholds. In support of the results, Bonthala et al. [216], reported high correlation between cross-species microarray data and RNA-sequencing approaches for detecting differentially expressed genes under a cold temperature stress experiment in bambara groundnut.

Probe-sets retained by the mask after genomic hybridisation are almost identical (>99%), suggesting that, at this level of resolution, the two genotypes are highly similar at the sequence level. Four known drought-associated genes, seen to be differentially expressed in these data, were subjected to qPCR, and supported the notion that the observed trends in the microarray data are valid.

The 26,496 probe sets common between the two genotypes, under irrigated conditions, (with a RMA cut-off of 0.97), include some sixty drought- and ABA-related genes. The latter include genes for producing osmoprotectants. They might provide two components of the drought avoidance capability of these genotypes, by retaining normal cell functioning when water access becomes limiting. Clearly, if the plant has already activated part of the drought response, it could have multiple effects. The presence of osmoprotectants might draw in even more water than otherwise might be the case, and there will be a greater proportion of biomass devoted to root growth, resulting in even deeper roots that are better able to survive drought later on. Bam-

bara groundnut is known to allocate a greater fraction of its dry weight to roots than shoots irrespective of soil moisture status [270]. This strategy may have clear advantages when water subsequently becomes limiting, suggesting adaptation to harsh environments and a decision to prioritise survival. In addition, as bambara groundnut is grown in harsh environments and has not undergone intensive breeding for yield and above ground biomass, this suggests it still allocates more effort to develop root architecture to handle drought when it happens. Moreover, Nayamudeza [270], also stated that the fraction of total dry weight allocated to roots in bambara groundnut is greater than groundnut. In addition, relatively higher expression of drought-related genes in both genotypes under irrigated conditions (including *ABI1* (ABA Insensitive 1), *ABF1* (ABRE binding factor 1), *ERD4* (Early responsive to dehydration 4), *RD19* (Response to dehydration 19), compared to other species such as Soybean [271], suggest that bambara groundnut could at least be partially in a ready state for drought, even in the absence of the drought stress. However, further research is needed to validate the hypothesis.

Given that 59782 and 59835 probe-sets were used to evaluate the transcriptome changes after probe-masking in DipC and TN, respectively, there were only very small numbers of genes significantly differentially expressed (189 in DipC and 81 in TN) under drought treatment. It could be speculated that the slow and progressive drought treatment might not cause significant shock to the plants.

The up-regulated genes in both genotypes were subdivided into ~75% drought responsive (with expression levels returning to normal after recovery) and ~25% drought perturbed (where the expression levels remained altered). In the case of down-regulated genes, 80-85% of expression levels returned to normal. The drought-perturbed expression levels might be caused by changes at the chromatin level, through DNA methylation or histone modification, and it is therefore interesting to note that a protein-lysine

demethylase is repressed by drought.

The above observations show that the two genotypes appear to be very similar in terms of genotype (validating the comparability of the transcriptome data compared using the microarray) while having differences in general transcriptional behaviour in irrigated conditions and in response to drought stress. However, when the sets of differentially expressed genes are compared, there is almost no overlap. Out of 189 and 91 genes differentially expressed in DipC and TN, respectively, only 9 were common between the two genotypes, suggesting that some of the mechanisms for adaption to drought is substantially in the two genotypes. Of these, Beta-fructofuranosidase contributes to osmoprotection, a *MYB* gene is associated with stomatal opening, *BRH1* affects stomatal density, and *bsd2* affects photosynthesis, while *JMJD5* plays an epigenetic role as mentioned above. Figure 4.6 illustrates how two genotypes have adapted to achieve the drought response traits (transcriptional and hormone signaling to affect cell-wall modification, lignin synthesis, photosynthesis, transporters, osmoprotection, oxidative stress) through largely different sets of effector genes.

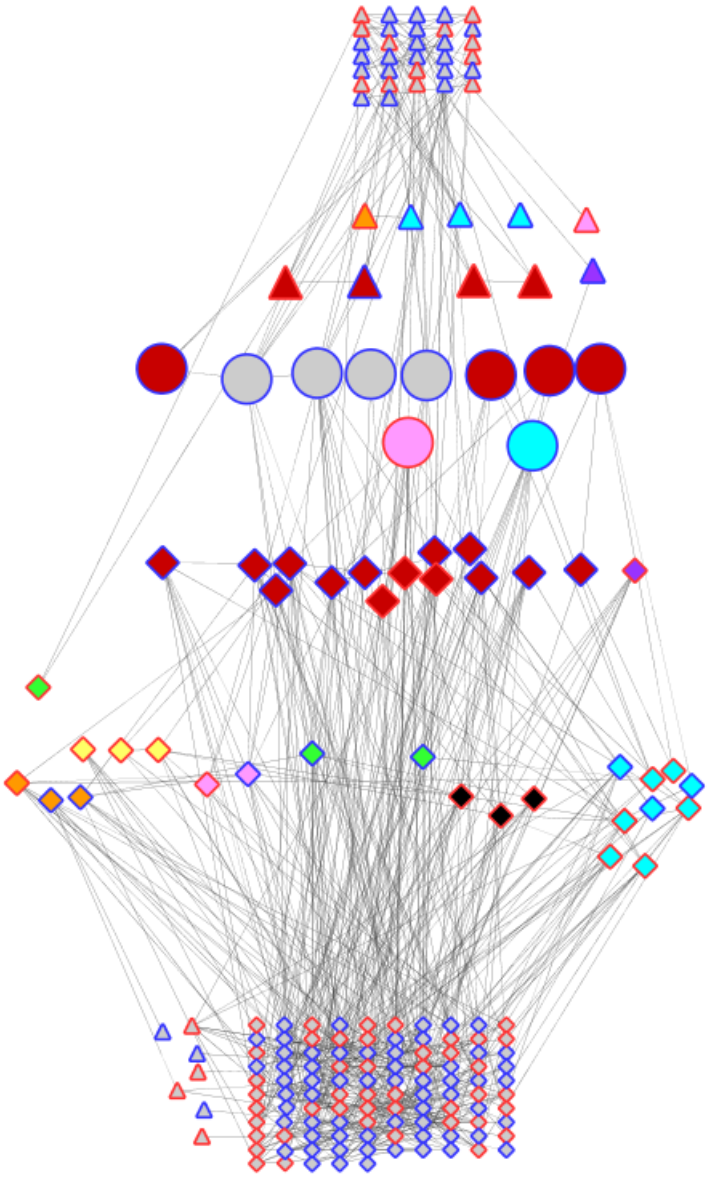

Figure 4.6: **Comparison of genotype co-expression networks**. Cytoscape has been used to layout the merged drought-responsive network of co-expressed probe sets. Node shapes are triangles, diamonds circles and circles respectively for differential expression of probe sets of TN, DipC, both (i.e. common) and both but affecting stomata. They have been coloured according to their activity in relation to drought response: red (transcription), orange (cell wall), yellow (lignin synthesis), green (photosynthesis), blue (transporters), indigo (hormone signaling), pink (osmoprotection), black (oxidative stress) and grey (others). Node borders have been coloured red and blue to denote up- and down- regulation under stress. Nodes have been arranged in 7 horizontal bands with probe sets in common in the middle flanked by TFs and hormone-signalling genes, other genes that play various roles response to drought, and others. Nodes have been linked by the criteria of the co-expression analysis. The figure was created by Charlie Hodgman, University of Nottingham, UK.

Several transcription factors that seem likely to play a role in the bambara groundnut drought response. Common to both genotypes are *BRH1* and a *MYB* which are known to affect stomata, and *JMJD5*. DipC shows a bigger response, with changes to *WRKY40* being of particular interest. It is a well-known member of plant drought-response networks [272] and is the most highly linked TF node in the coexpression networks. For DipC, the network also reveals the importance of *PRR7*, a core circadian clock component known to play a complex role in abiotic stresses [273]. *ATAUX2-11* and a *C2H2* Zinc-finger protein. It is somewhat surprising that TN does not show a >2-fold change in the expression of *WRKY40*, but it does appear to have roles for CONSTANS-like 1 (another clock-related gene associated with flowering that may be associated with abiotic stress [274], *MYB60* which affects stomatal closure [275], and *AGL-83*, a *MADS-Box* protein of uncertain role.

4.9 Conclusion

Understanding the mechanisms underlying the ability of crops to produce viable yields under drought conditions is a priority for global food security. This study has examined the transcriptomic response to drought and recovery in two genotypes derived from landraces of bambara groundnut, in an attempt to investigate the molecular mechanisms occurring in the two landraces. In addition, this study also tested if the cross-species hybridisation to the soybean microarray is suitable for investigating the bambara groundnut transcriptome. It was shown that many potential drought-responsive genes are expressed even under irrigated conditions in both landraces, suggesting that bambara groundnut could at least be partially in a ready state for drought, even in the absence of the drought stress. In terms of differential expression, there were only a very small number of genes differentially expressed under drought treatment in both landraces, suggesting that the slow and progressive drought treatment might not

cause significant shock to the plants. Although the transcription factors and drought-response genes were largely different between the two landraces, they may achieve the same effect in terms of survival under drought conditions. The DipC genotype showed differential expression of some well known drought-related transcriptions factors (especially *WRKY40*), while TN showed differential expression of CONSTANS-LIKE 1 and *MYB60* instead. Cross-species hybridisation to the soybean microarray has been shown to be informative for investigating the bambara groundnut transcriptome as good gene retention was observed at high gDNA hybridisation thresholds.

Chapter 5

Morpho-physiological effects of drought stress on bambara groundnut at different developmental stages

5.1 Introduction

Drought stress in one of the major environmental constraint to plant productivity. To cope with drought stress, plants respond to drought stress via three different mechanisms. These include (i) drought escape, (ii) drought avoidance, and (iii) drought tolerance. Drought escape is described as the ability of plants to complete their growth cycle and reach maturity before drought-stress develops to damaging levels [84]. Drought avoidance is demonstrated by crop species, which are able to maintain high water potential in the plant by minimising water loss and maximising water uptake under drought conditions. Mechanisms of avoidance include improved root traits, for greater extraction of soil moisture, decreased stomatal conductance, decreased radiation absorption and decreased leaf area for minimal water loss [87]. Drought tolerance allows plants to survive through water-use efficiency, i.e., performing all bi-

ological, molecular and cellular functions with minimal water. Plants with drought tolerance mechanisms are able to maintain their cell turgor through osmotic adjustment, which in turn will contribute to maintaining stomatal opening, leaf expansion and photosynthesis throughout the drought period [90].

The measurements widely used to investigate the effects of drought stress on plants include stomatal conductance, photosynthesis, transpiration, leaf area, relative water content (RWC) and chlorophyll content [116]. A reduction in stomatal conductance due to stomatal closure (via ABA signaling) was observed after the drought stress was imposed, which decreases the CO_2 influx, thus leading to a reduction in photosynthesis [276]. Subramaniam et al. [277], reported that photosynthesis and transpiration rate decreased progressively with increasing duration of duration of drought stress, indicating that plants under mild stress were postponing tissue dehydration. The effects of long duration of soil water deficit on canopy assimilation is the reduction in leaf area. Drought stress reduces leaf area by slowing leaf expansion and reducing in the supply of carbohydrates [276]. Additionally, drought stress can lead to a reduction in morphological traits such as plant height and leaf number, thus reducing water loss which could assist the plant to avoid drought.

Bambara groundnut as discussed in 2.1.1 is reputed to be a drought tolerant crop able to survive in regions of minimal water. Collinson et al. [90], stated that bambara groundnut achieves its drought tolerance through physiological responses such as osmotic adjustment and stomatal closure. Stomatal closure limits CO_2 uptake by leaves which leads to a reduction in photosynthesis. Physiological and morphological response of bambara groundnut under drought stress has been studied a number of times. Some bambara groundnut landraces show paraheliotropic properties, leading to less transpiratory water loss due to the lower leaf temperature that results from decreased light interception [3]. From the results of Mabhaudhi et al. [113], bambara

groundnut was observed to have drought escape mechanisms where, under drought stress, it had a shortened vegetative growth period, earlier flowering, reduced reproductive stage length and early maturity in order to minimise the adverse effect of drought on plant development. Higher root dry weight was reported when the bambara groundnut landrace, Burkina (originally from dry regions of Burkina Faso), was subjected to drought [14]. Denser and deeper root growth may allow the plant to utilise more soil moisture under drought stress. Vurayai et al. [116], reported reduced leaf area in drought-stressed bambara groundnut plants due to turgor reduction, which is one of the earliest physiological responses to water stress.

Studying morpho-physiological characteristics of a plant can indicate whether it might have the ability to withstand drought stress conditions. There is very limited literature available on how bambara groundnut responds to short periods of drought stress imposed at different development stages [116]. This study aims to identify the morpho-physiological effects of drought stress on bambara groundnut at different developmental stages. This study will aid in identifying crucial drought related traits in bambara groundnut which will provide a basis for breeding. Additionally, this study will provide a platform for identifying genes related to the triat of interest through expression quantitative trait loci (eQTL) analysis. Two genotypes of bambara groundnut (DipC and Tiga Nicuru (TN)) derived from landraces were used for the study. These landraces as discussed in 1.1.3 are more tolerant to drought than many other landraces, while potentially having some variation, as they are morphologically and phenotypically distinct. Peanut (*Arachis hypogaea*) (accession GRIF 302) has been used for comparative analysis with bambara groundnut in relation to drought tolerance. Several conventional and molecular breeding techniques have been adopted to improve drought tolerance in peanut [278, 279]. The effects of drought stress on peanut has also been studied at the molecular and cellular level which resulted in

a substantial amount of genomic and proteomic data investigating the mechanism by which peanut respond to drought stress [279]. By keeping peanut as a comparator species, this study will help in developing and understanding the nature of the drought tolerance level of bambara groundnut compared to other well-studied legume.

5.2 Results

5.2.1 Morpho-physiological response of bambara groundnut genotypes (DipC and TN) and Peanut under drought stress

Several morpho-physiological measurements were taken for DipC, TN and peanut in both irrigated (75% field capacity) and drought treatment plot (50-25% field capacity) based on the bambara groundnut descriptor list (IPGRI, IITA, BAMNET, 2000). The measurements were taken during vegetative, reproductive, pod development and after harvesting. Section 3.2 list all the measurements taken throughout the study. Soil moisture for controls (normal irrigation) were kept at 75% field capacity (FC) throughout the study, while for treatment plot at vegetative and reproductive stages soil water was kept at 50% field capacity for week 1 of drought stress and lowered to 25% FC from the beginning of week 2. Whereas, in treatment plot for pod development stage soil water was kept at 50% FC for four weeks and lowered to 25% from fifth week onwards.

5.2.1.1 Relative water content (RWC)

As shown in Figure 5.1, TN showed significantly lower ($p<0.05$) RWC in drought treatment plot compared to control (irrigated) plot at the vegetative and pod development stage, where week 4 (89 DAS) of drought stress at the pod development stage experienced the highest percentage reduction (8.8%) in treatment compared to control plots. However, no significant difference was observed at the reproductive stage. For DipC, significant difference between control and treatment plots was observed at the vegetative, reproductive and pod development stage. Although, no significant difference was

observed in DipC at week 2 of the vegetative stage. On the other hand, peanut did not show any significant difference between control and treatment plot at all three developmental stages (vegetative, reproductive and pod development). From the ANOVA analysis, significant differences ($p<0.05$) between the two genotypes (DipC and TN) for treatments, as well as between the treatments (drought vs control) ($p<0.01$) were observed at pod development stage. However, no significant difference between the genotypes for treatments was observed at the vegetative and reproductive stage (Table 5.4).

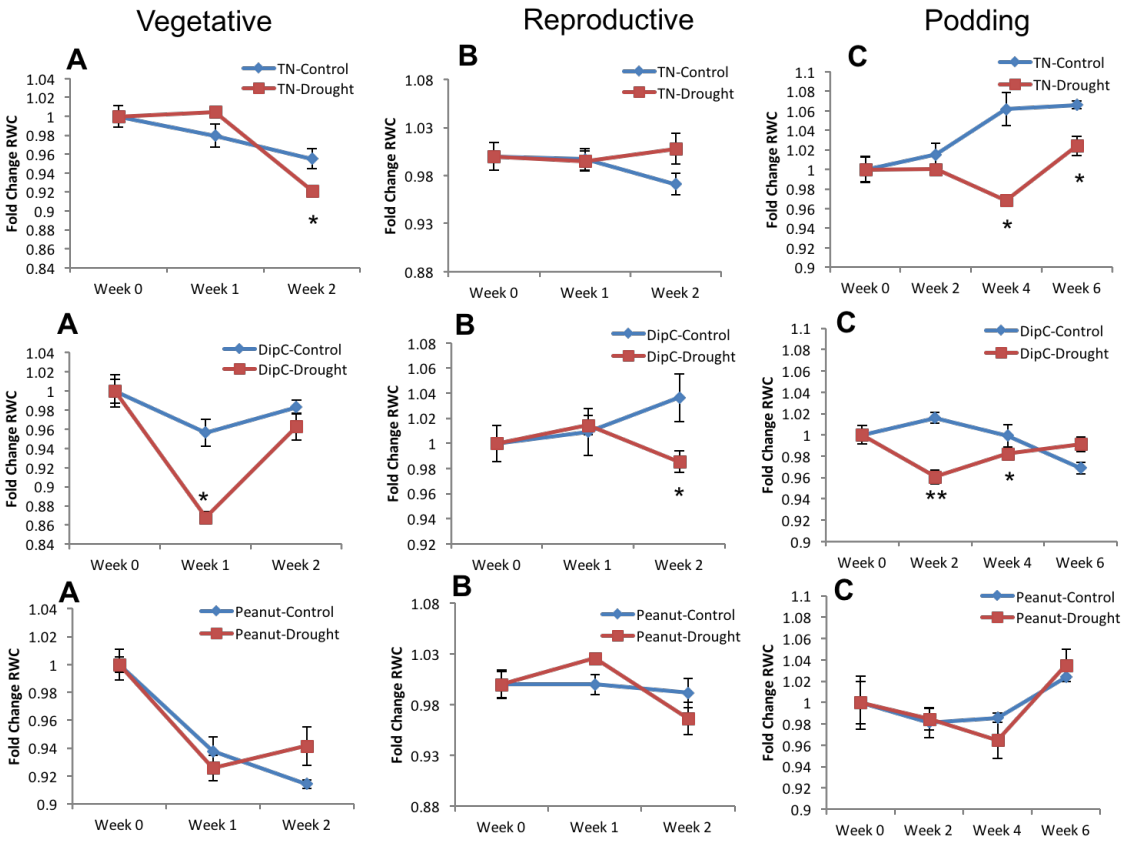

Figure 5.1: **Effect of drought stress on relative water content (%) in TN, DipC and peanut under drought stress.** Column [A], [B] and [C] refers to vegetative, reproductive and pod development (podding) stage respectively. First, second and third row refers to RWC data from TN, DipC and peanut respectively. Values are shown as fold change with respect to week 0 (before drought stress) at each developmental stage when all the plants were at irrigated (control) conditions. Absolute values for week 0 can found on Appendix 15. The measurements over time are non-linear. The blue line indicates the RWC values from control conditions and red line indicates the values from drought conditions. Values shown are the average of three biological replicates. Error bars denote the standard error of the average of three biological replicates. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

5.2.1.2 Photosynthesis

Throughout the drought stress period, the fold change values for photosynthesis in treatment plot were significantly lower ($p<0.05$) compared to control plot in TN,

DipC and peanut at all three developmental stages (Figure 5.2). However, a greater reduction in photosynthesis was observed in treatment plot in peanut at the reproductive stage compared to DipC and TN. On the other hand, there was no significant difference observed between control and treatment plot at the 6th week (103 DAS) of drought stress at the pod development stage in peanut. The highest percentage reductions in treatment plot compared to control plot were observed at the pod development stage in DipC, TN and peanut. Analysis of data using ANOVA, showed a significant difference ($p<0.05$) between the genotypes for treatments at the reproductive stage. Whereas, no significant difference was observed between the genotypes for treatments at the vegetative and pod development stage. In addition, significant difference was observed between the treatments at all three developmental stages (Table 5.4).

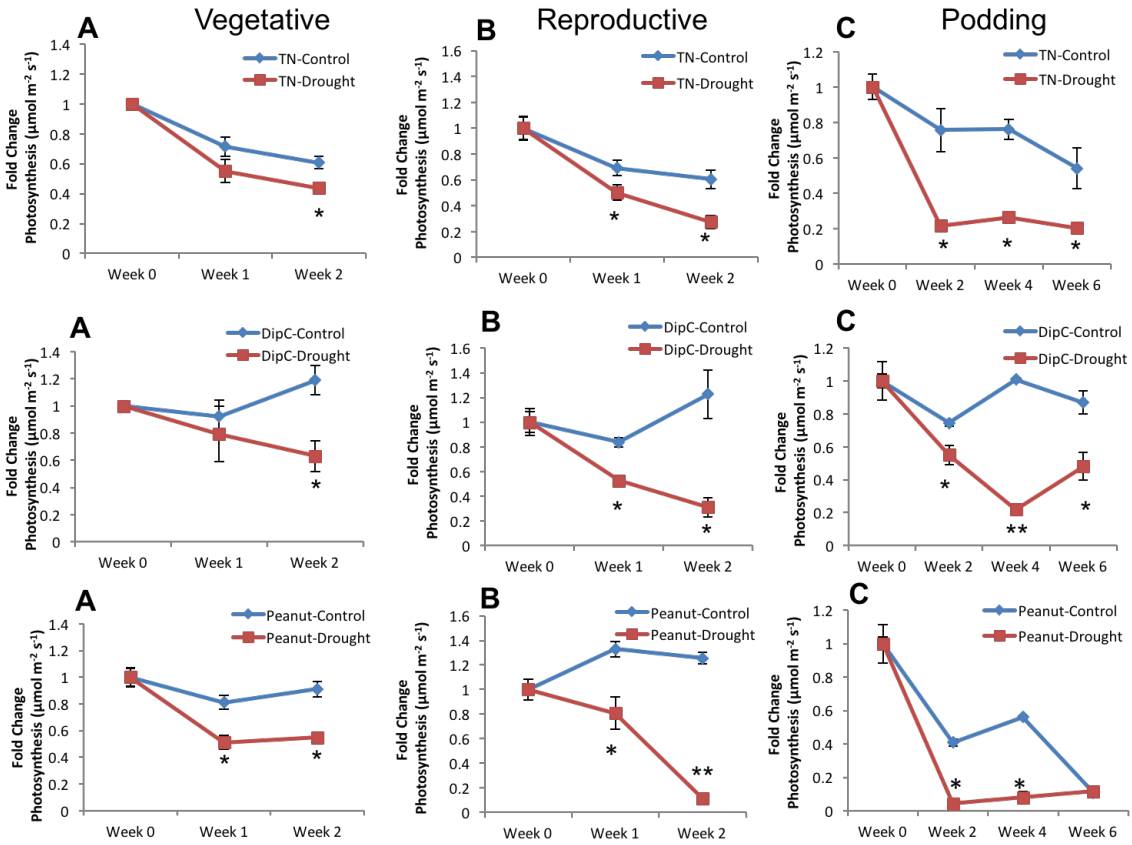

Figure 5.2: **Effect of drought stress on photosynthesis in TN, DipC and peanut under drought stress.** Column [A], [B] and [C] refers to vegetative, reproductive and pod development (podding) stage respectively. First, second and third row refers to photosynthesis data from TN, DipC and Peanut respectively. Values are shown as fold change with respect to week 0 (before drought stress) at each developmental stage when all the plants were at irrigated (control) conditions. Absolute values for week 0 can be found on Appendix 15. The measurements over time are non-linear. The blue line indicates the photosynthesis values from control (irrigated) conditions and red line indicates the values from drought conditions. Values shown are the average of three biological replicates. Error bars denote the standard error of the average of three biological replicates. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

5.2.1.3 Stomatal Conductance (g_s)

Significant reductions ($p<0.05$) in stomatal conductance was observed between treatment and control plots in TN at pod development stage. Whereas, no significant

reduction was seen at the vegetative and reproductive stage (Figure 5.3). On the other hand, DipC showed a significant reduction at the reproductive stage ($p<0.05$) but not in vegetative stage. In the case of the pod development stage, both DipC and TN treatment plot had significantly lower fold change values ($p<0.05$) compared to control plots till week 4 (89 DAS) of drought stress. However, at week 6 (103 DAS), no significant difference was observed between treatment and control plots in both genotypes. In case of peanut significant difference was observed between control and treatment plot at all three stages. Same as DipC and TN, peanut showed no significant difference between control and treatment plot at week 6 (103 DAS). The highest percentage reduction in stomatal conductance in treatment plot compared to control plot was observed in peanut at all three developmental stages. ANOVA analysis showed, significant difference ($p<0.05$) between the genotypes for treatment at the reproductive and pod development stage as well as between the treatments (Table 5.4).

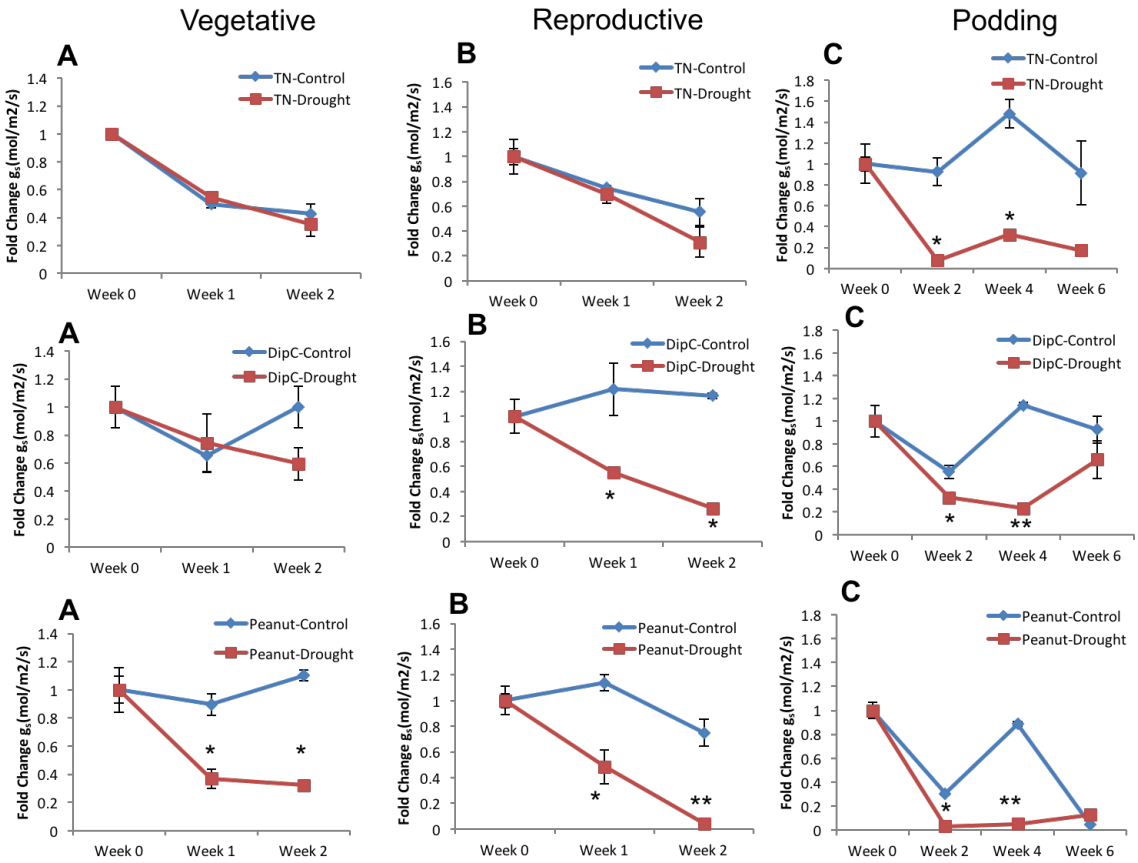

Figure 5.3: **Effect of drought stress on stomatal conductance in TN, DipC and peanut under drought stress.** Column [A], [B] and [C] refers to vegetative, reproductive and pod development (podding) stage respectively. First, second and third row refers to stomatal conductance data from TN, DipC and Peanut respectively. Values are shown as fold change with respect to week 0 (before drought stress) at each developmental stage when all the plants were at irrigated (control) conditions. Absolute values for week 0 can be found on Appendix 15. The measurements over time are non-linear. The blue line indicates the stomatal conductance values from control conditions and red line indicates the values from drought conditions. Values shown are the average of three biological replicates. Error bars denote the standard error of the average of three biological replicates. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

5.2.1.4 Transpiration

Transpiration was significantly reduced ($p<0.05$) in treatment plot compared to control plot at pod development stage in DipC, TN and peanut. Significant reduction was observed in DipC and peanut even at the vegetative and reproductive stage (Figure 5.4). Highest percentage reduction in treatment plot compared to irrigated plants was observed in peanut compared to DipC and TN at all three stages. As seen with the measurement of photosynthesis in peanut, no significant difference were observed between control and treatment plot at the 6th week (103 DAS) of drought stress at the pod development stage. Analysis of data using ANOVA, showed significant difference ($p<0.05$) between the genotypes for treatment at the reproductive and pod development stage, as well as between the treatments. On the other hand, at the vegetative stage, significant difference between the treatments was observed but not between the genotypes for treatments (Table 5.4).

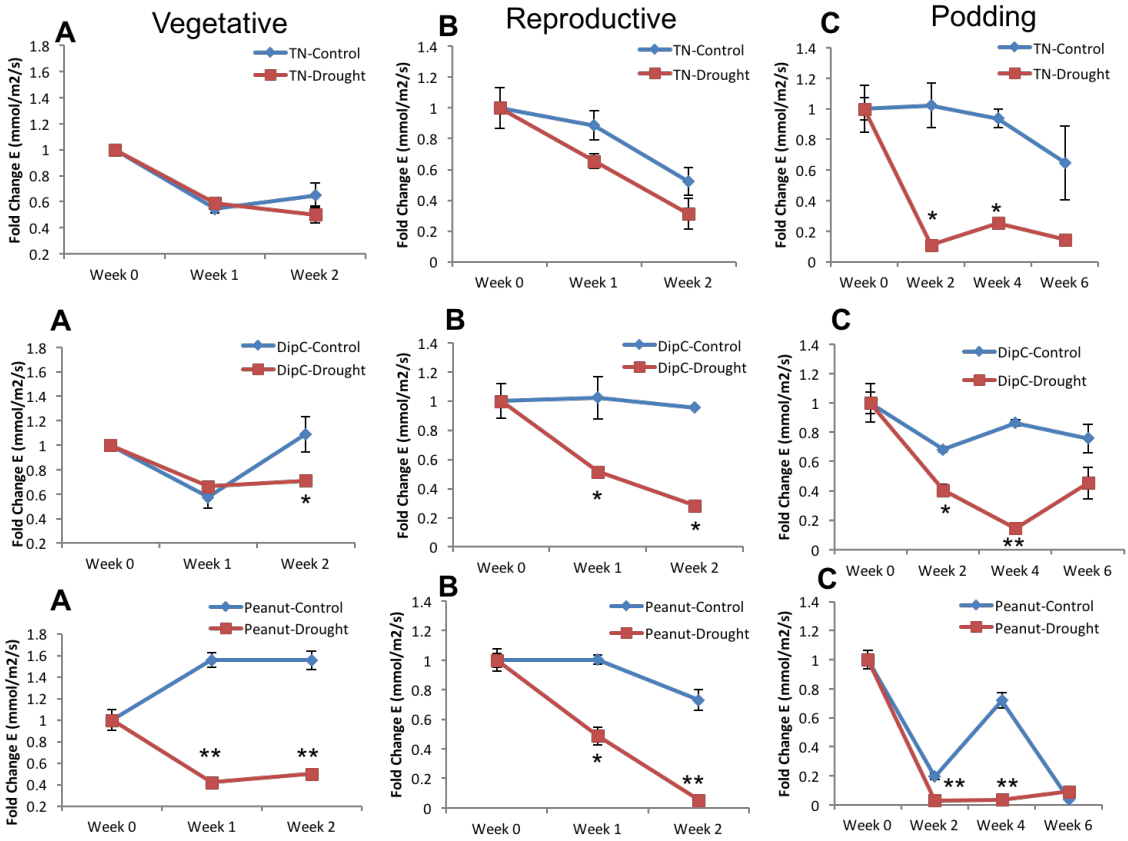

Figure 5.4: **Effect of drought stress on transpiration in TN, DipC and peanut under drought stress.** Column [A], [B] and [C] refers to vegetative, reproductive and pod development (podding) stage respectively. First, second and third row refers to transpiration data from TN, DipC and Peanut respectively. Values are shown as fold change with respect to week 0 (before drought stress) at each developmental stage when all the plants were at irrigated (control) conditions. Absolute values for week 0 can be found on Appendix 15. The measurements over time are non-linear. The blue line indicates the transpiration values from control conditions and red line indicates the values from drought conditions. Values shown are the average of three biological replicates. Error bars denote the standard error of the average of three biological replicates. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

5.2.1.5 SPAD Chlorophyll meter reading (SCMR)

No significant reductions were observed between control and treatment plot at vegetative and reproductive stages in DipC, TN and peanut (Figure 5.5). However, rapid

reduction was seen at the pod development stage in all three (DipC, TN and peanut). Highest percentage reduction in treatment plot compared to control plot was observed at week 6 (103 DAS) of drought stress in TN (77.36%). Significant differences ($p<0.01$) between the genotypes for treatment were observed at the pod development stage from the ANOVA analysis. Significant reduction between treatments ($p<0.05$) was observed at all three developmental stages. Whereas, no significant difference was observed between the genotypes for treatment at the vegetative and reproductive stage (Table 5.4).

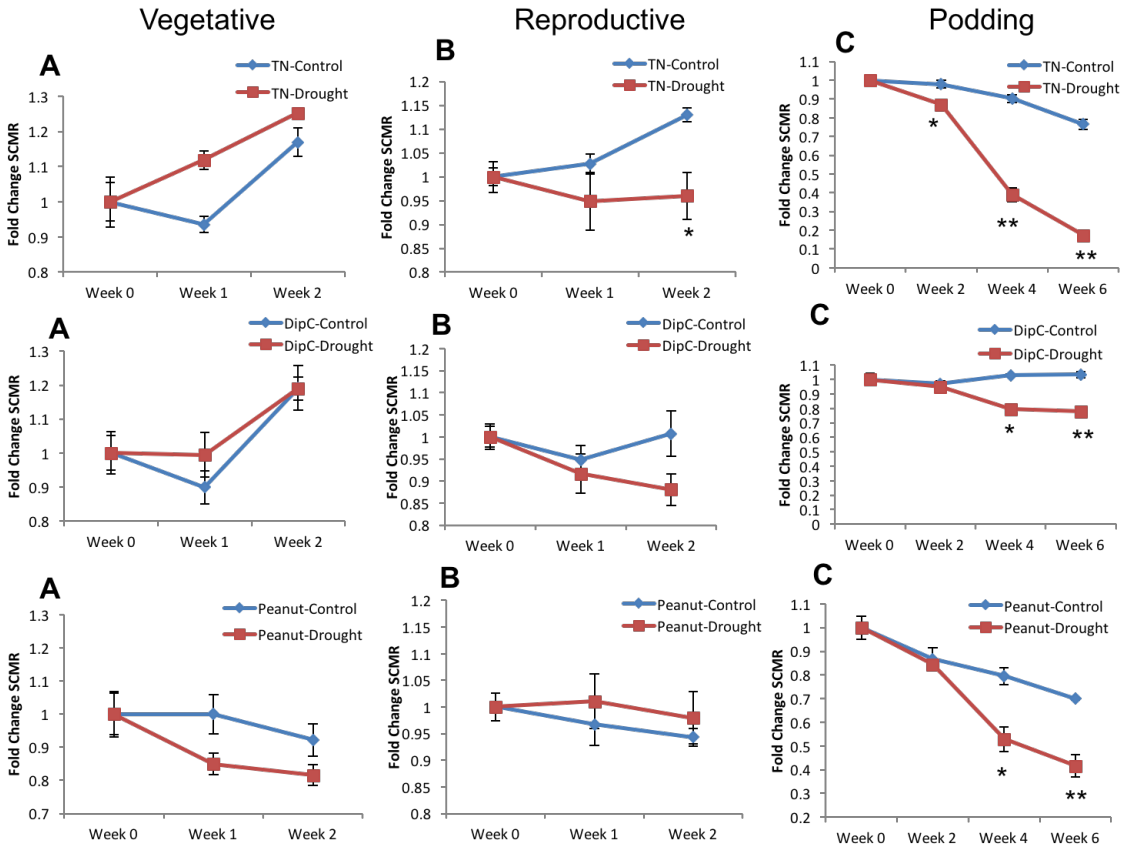

Figure 5.5: **Effect of drought stress on SCMR readings in TN, DipC and peanut under drought stress.** Column [A], [B] and [C] refers to vegetative, reproductive and pod development (podding) stage respectively. First, second and third row refers to SCMR readings from TN, DipC and Peanut respectively. Values are shown as fold change with respect to week 0 (before drought stress) at each developmental stage when all the plants were at irrigated (control) conditions. Absolute values for week 0 can be found on Appendix 15. The measurements over time are non-linear. The blue line indicates the SCMR readings from control conditions and red line indicates the values from drought conditions. Values shown are the average of three biological replicates. Error bars denote the standard error of the average of three biological replicates. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

5.2.1.6 Plant Height

For TN, significant reduction ($p<0.05$) in fold change values for plant height in treatment plot compared to control plot was observed at the vegetative stage (Figure 5.6).

However no significant difference in DipC and peanut was observed at the vegetative stage. At the reproductive and pod development stage, no significant difference was observed in DipC and TN. This is not surprising, as plants at these growth stages (reproductive and pod development) are utilising its energy in producing flowers and pods respectively. Hence, the growth of plant height slows down even in irrigated plants. Analysis of variance showed significant differences ($p<0.05$) between the genotypes for treatments at the vegetative and reproductive stages. Whereas, significant difference ($p<0.05$) between the treatments was only observed at the vegetative and reproductive stage (Table 5.5).

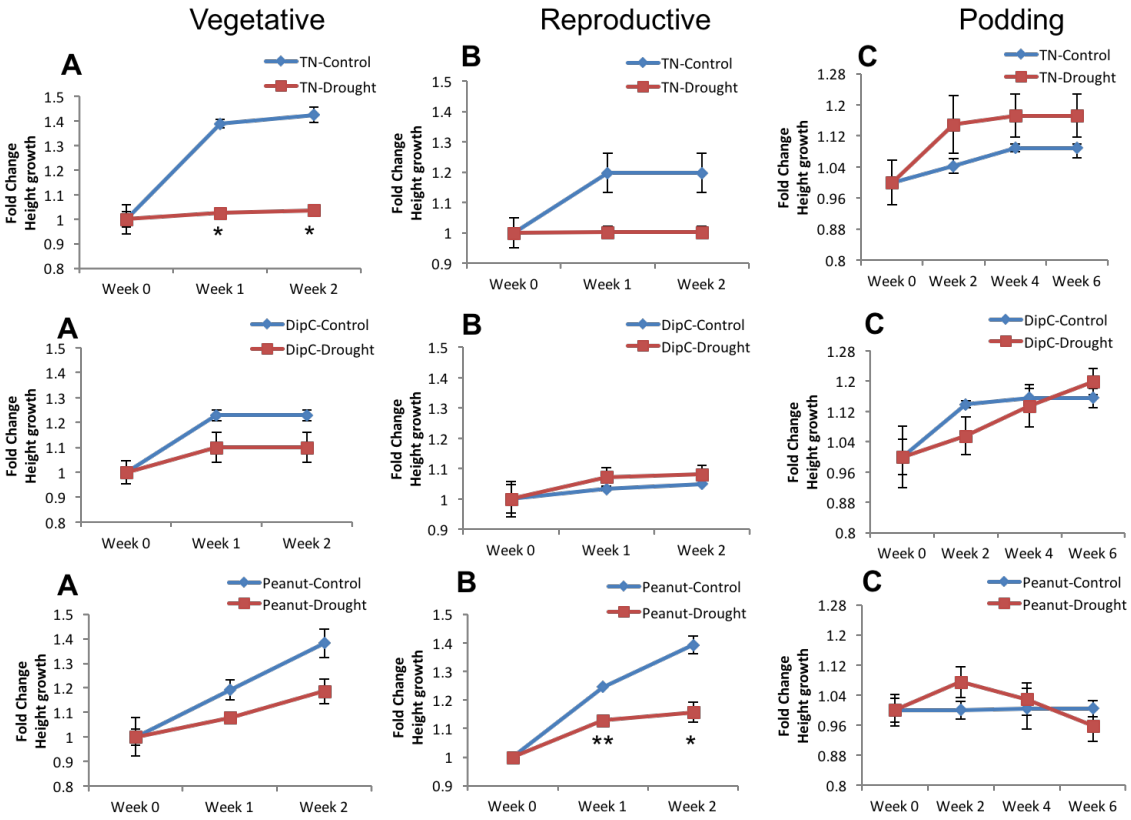

Figure 5.6: **Effect of drought stress on plant height growth in TN, DipC and Peanut.** Column [A], [B] and [C] refers to vegetative, reproductive and pod development (podding) stage respectively. First, second and third row refers to plant height growth from TN, DipC and Peanut respectively. Values are shown as fold change with respect to week 0 (before drought stress) at each developmental stage when all the plants were at irrigated (control) conditions. Absolute values for week 0 can be found in on Appendix 15. The measurements over time are non-linear. The blue line indicates the plant height growth from control conditions and red line indicates the values from drought conditions. Values shown are the average of three biological replicates. Error bars denote the standard error of the average of three biological replicates. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

5.2.1.7 Leaf number

Significant reductions ($p<0.05$) in leaf number for treatment plot compared to control plot was observed in TN, DipC and peanut at the pod development stage (Figure 5.7). Although, a reduced trend in leaf number was seen in treatment plot, from week

2 (75 DAS) onwards no increase in leaf number was seen even in control plot. This is not surprising, as by pod development stage, plants have already completed it's growth in producing leaf and is now utilising its energy in pods formation. For TN, in addition, significant reduction ($p<0.01$) was seen at the vegetative stage but not for DipC and TN. On the other hand, no significant reduction was observed in TN, DipC and peanut at reproductive stage. ANOVA analysis showed significant differences ($p<0.05$) between the genotypes for treatment at the pod development stage but not at vegetative and reproductive stage. On the other hand, significant difference between the treatments was observed at the vegetative and reproductive stage (Table 5.5).

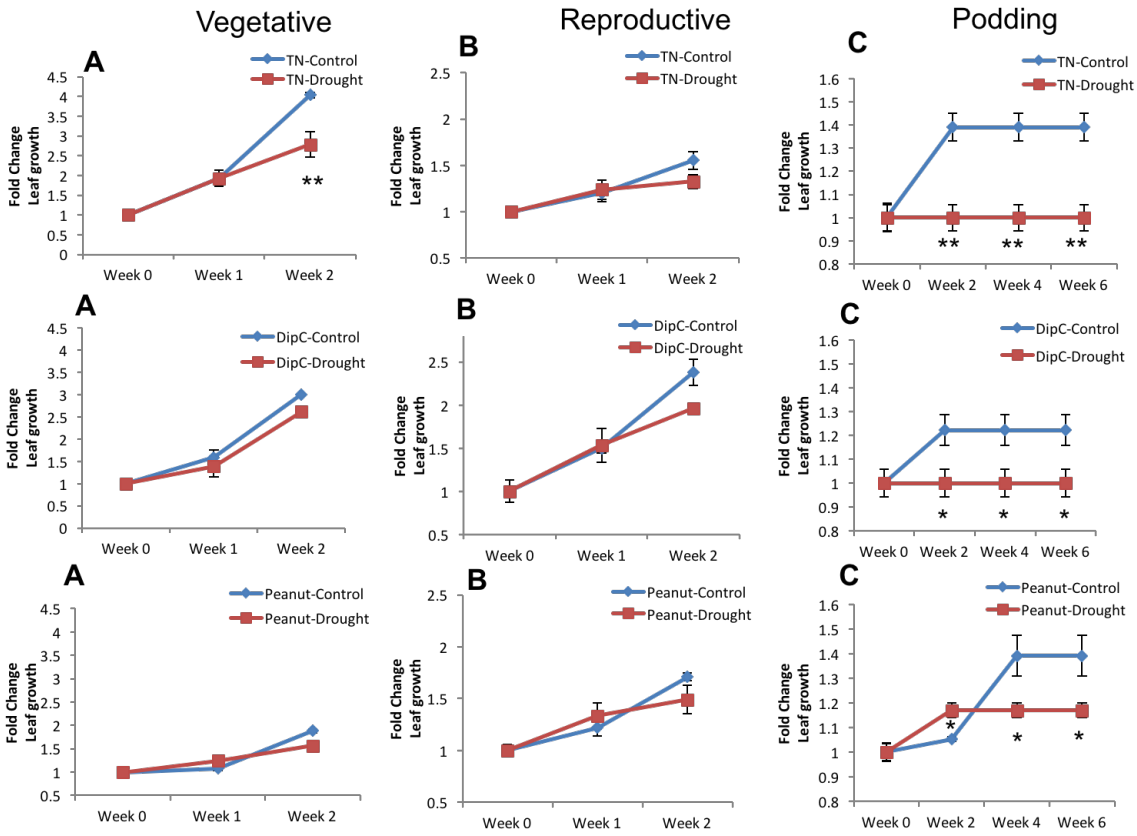

Figure 5.7: **Effect of drought stress on leaf number in TN, DipC and peanut under drought stress.** Column [A], [B] and [C] refers to vegetative, reproductive and pod development (podding) stage respectively. First, second and third row refers to leaf number data from TN, DipC and Peanut respectively. Values are shown as fold change with respect to week 0 (before drought stress) at each developmental stage when all the plants were at irrigated (control) conditions. Absolute values for week 0 can be found on Appendix 15. The measurements over time are non-linear. The blue line indicates the leaf number from control conditions and red line indicates the values from drought conditions. Values shown are the average of three biological replicates. Error bars denote the standard error of the average of three biological replicates. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

5.2.1.8 Leaf Area

In DipC and TN, for vegetative stage, a significant reduction at week 1 (26 DAS) of drought stress was observed in treatment plot compared to control plot (Figure 5.8).

Whereas, no significant reduction in DipC and TN was observed at the reproductive stage. Significant reduction ($p<0.05$) in TN at week 4 (89 DAS) of pod development stage was seen as well. Analysis of data using ANOVA, showed significant differences ($p<0.05$) between the genotypes for treatment at the pod development stage. Significant difference ($p<0.05$) between treatments was observed at the vegetative and pod development stage (Table 5.5).

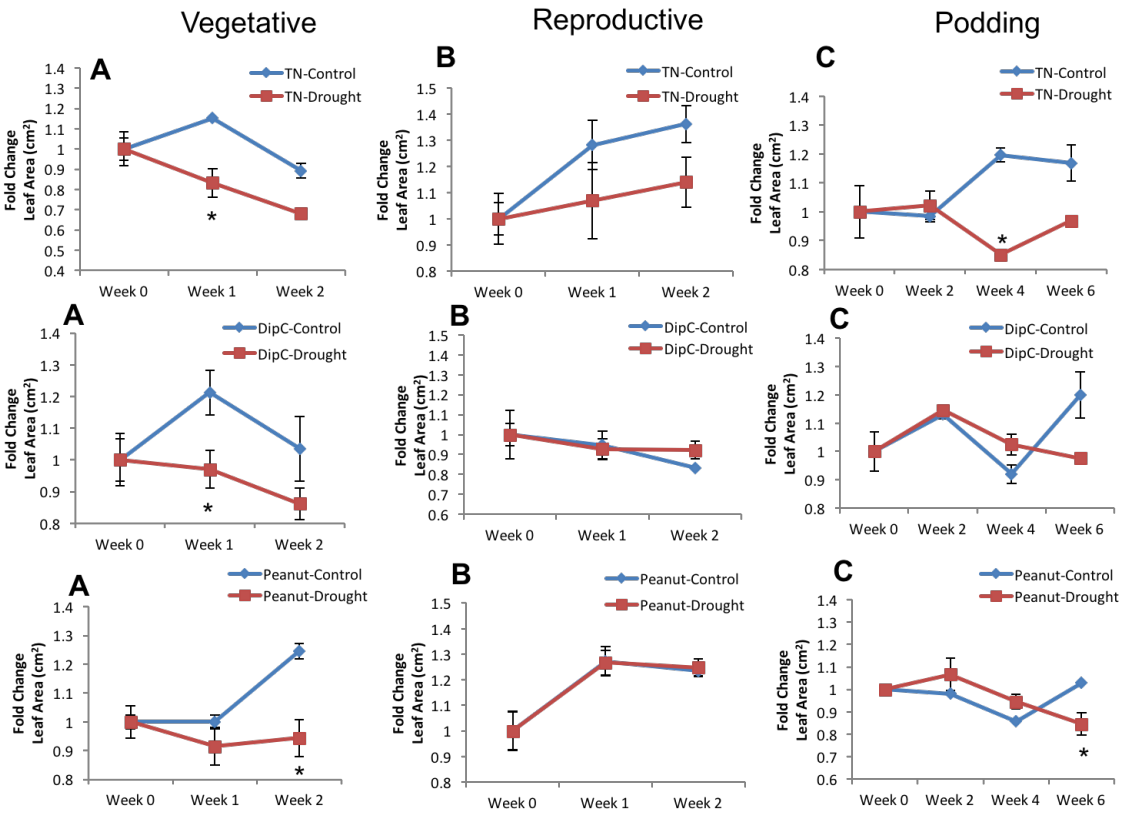

Figure 5.8: **Effect of drought stress on leaf area in TN, DipC and peanut under drought stress.** Column [A], [B] and [C] refers to vegetative, reproductive and pod development (podding) stage respectively. First, second and third row refers to leaf area data from TN, DipC and Peanut respectively. Values are shown as fold change with respect to week 0 (before drought stress) at each developmental stage when all the plants were at irrigated (control) conditions. Absolute values for week 0 can be found on Appendix 15. The measurements over time are non-linear. The blue line indicates the leaf area from control conditions and red line indicates the values from drought conditions. Values shown are the average of three biological replicates. Error bars denote the standard error of the average of three biological replicates. Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

5.2.1.9 Flower number

Drought treatment from just after the onset of flowering to seed maturation caused significant reductions in flower number in treatment plot compared to control plot

Table 5.1: **Morpho-physiological measurements taken for control and treatment plots at vegetative stage**

Plant materials	Measurements weeks	Treatment	RWC	SPAD	Photosynthesis	g_s	Transpiration	Leaf Area	Plant Height	Leaf number
TN	Week 1 (50% FC)	Control	0.97	0.94	0.71	0.50	0.55	1.15	1.39	1.92
		Drought	1.00	1.12	0.55	0.55	0.59	0.83*	1.02*	1.93
	Week 2 (25% FC)	Control	0.96	1.17	0.61	0.43	0.65	0.89	1.42	4.04
		Drought	0.92*	1.25	0.43*	0.35	0.50	0.68	1.03*	2.78**
DipC	Week 1 (50% FC)	Control	0.96	0.90	0.92	0.65	0.57	1.21	1.23	1.58
		Drought	0.86*	0.99	0.79	0.74	0.66	0.97*	1.10	1.38
	Week 2 (25% FC)	Control	0.98	1.19	1.19	1.00	1.09	1.04	1.23	3.00
		Drought	0.96	1.19	0.63*	0.60	0.70*	0.86	1.10	2.62
Peanut	Week 1 (50% FC)	Control	0.94	1.00	0.81	0.90	1.56	1.00	1.19	1.08
		Drought	0.93	0.85	0.50*	0.36*	0.42**	0.91*	1.08	1.24
	Week 2 (25% FC)	Control	0.91	0.92	0.91	1.10	1.56	1.25	1.38	1.88
		Drought	0.94	0.82	0.54*	0.32*	0.49**	0.94*	1.19	1.56

Values are shown as fold change with respect to week 0 (before drought stress).

Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

Table 5.2: **Morpho-physiological measurements taken for control and treatment plots at reproductive stage**

Plant materials	Measurements weeks	Treatment	RWC	SPAD	Photosynthesis	g_s	Transpiration	Leaf Area	Plant Height	Leaf number
TN	Week 1 (50% FC)	Control	1.00	1.03	0.69	0.75	0.89	1.28	1.20	1.20
		Drought	1.00	0.95	0.50*	0.70	0.65	1.07	1.00	1.24
	Week 2 (25% FC)	Control	0.97	1.13	0.60	0.55	0.52	1.36	1.20	1.56
		Drought	1.01	0.96*	0.27*	0.31	0.31	1.14	1.00	1.33
DipC	Week 1 (50% FC)	Control	1.01	0.95	0.84	1.22	1.02	0.95	1.03	1.50
		Drought	1.01	0.92	0.52*	0.55*	0.51*	0.93	1.07	1.54
	Week 2 (25% FC)	Control	1.04	1.01	1.23	1.17	0.95	0.83	1.05	2.38
		Drought	0.98*	0.88	0.31*	0.26*	0.28*	0.92	1.08	1.96
Peanut	Week 1 (50% FC)	Control	1.00	0.97	1.33	1.14	1.00	1.27	1.25	1.22
		Drought	1.03	1.01	0.80*	0.48*	0.48*	1.27	1.12*	1.33
	Week 2 (25% FC)	Control	0.99	0.94	1.26	0.75	0.73	1.23	1.39	1.71
		Drought	0.97	0.98	0.11**	0.04**	0.05**	1.25	1.15**	1.49

Values are shown as fold change with respect to week 0 (before drought stress).

Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

Table 5.3: **Morpho-physiological measurements taken for control and treatment plots at pod development stage**

Plant materials	Measurements weeks	Treatment	RWC	SPAD	Photosynthesis	g_s	Transpiration	Leaf Area	Plant Height	Leaf number
TN	Week 2 (50% FC)	Control	1.02	0.98	0.76	0.92	1.02	0.98	1.04	1.39
		Drought	1.00	0.87*	0.21*	0.07*	0.10*	1.02	1.15	1.00**
	Week 4 (50% FC)	Control	1.06	0.90	0.76	1.48	0.94	1.20	1.09	1.39
		Drought	0.96*	0.38**	0.26*	0.32*	0.25*	0.85*	1.17	1.00**
	Week 6 (25% FC)	Control	1.07	0.76	0.54	0.91	0.65	1.17	1.09	1.39
		Drought	1.02*	0.17**	0.20*	0.17	0.15	0.97	1.17	1.00**
DipC	Week 2 (50% FC)	Control	1.02	0.97	0.74	0.55	0.68	1.13	1.14	1.22
		Drought	0.96**	0.95	0.55*	0.32*	0.40*	1.15	1.06	1.00*
	Week 4 (50% FC)	Control	1.00	1.03	1.01	1.14	0.86	0.92	1.16	1.22
		Drought	0.98*	0.79*	0.21**	0.23**	0.14**	1.03	1.13	1.00*
	Week 6 (25% FC)	Control	0.97	1.03	0.87	0.92	0.75	1.20	1.16	1.22
		Drought	0.99	0.77**	0.48*	0.66	0.45	0.98	1.20	1.00*
Peanut	Week 2 (50% FC)	Control	0.98	0.87	0.41	0.30	0.19	0.98	1.00	1.05
		Drought	0.98	0.84	0.04*	0.03*	0.03**	1.07	1.07	1.16*
	Week 4 (50% FC)	Control	0.99	0.79	0.56	0.89	0.72	0.86	1.00	1.39
		Drought	0.96	0.52*	0.08*	0.05**	0.035**	0.95	1.03	1.16*
	Week 6 (25% FC)	Control	1.02	0.70	0.11	0.05	0.04	1.03	1.00	1.39
		Drought	1.03	0.41**	0.12	0.13	0.09	0.84*	0.96	1.16*

Values are shown as fold change with respect to week 0 (before drought stress).

Single and double asterisks indicate that p-value is less than 0.05 and 0.01 respectively, which was assessed by the paired t-test between groups.

in both DipC ($p<0.01$) and TN ($p<0.05$). Greater reductions were observed in TN (39%) compared to DipC (29%).

Table 5.4: **Two-way ANOVA for physiological data for DipC and TN**

	Photosynthesis		g_s		Transpiration		RWC		Chlorophyll content	
Interactions	**F value**	**P value**	**F value**	**P value**	**F value**	**P value**	**F value**	**P value**	**F value**	**P value**
Drought initiated at vegetative stage										
Treatment	11.49	0.003*	0.95	0.340	1.78	0.195	2.77	0.110	4.32	0.049*
Treatment x Genotypes	1.370	0.254	0.700	0.412	0.370	0.548	1.340	0.260	0.980	0.333
Treatment x weeks	4.450	0.024*	2.630	0.095	4.180	0.029*	0.810	0.459	2.050	0.153
Genotypes x weeks	5.030	0.016*	4.170	0.029*	4.140	0.030*	38.810	0.001**	0.690	0.511*
Treatment x Genotypes x weeks	1.800	0.189	1.010	0.381	0.740	0.487	12.470	0.001**	0.250	0.784
Drought initiated at reproductive stage										
Treatment	29.62	0.001**	16.55	0.001**	20.28	0.001**	0.03	0.859	11.19	0.003*
Treatment x Genotypes	4.830	0.039*	7.740	0.011*	4.180	0.050*	1.730	0.201	0.550	0.468
Treatment x weeks	11.400	0.001**	4.800	0.019*	5.180	0.014*	0.070	0.936	4.580	0.022*
Genotypes x weeks	3.410	0.050*	1.150	0.335	1.260	0.303	0.390	0.679	2.100	0.146
Treatment x Genotypes x weeks	2.780	0.084	1.940	0.167	1.250	0.307	2.260	0.128	0.140	0.872
Drought initiated at pod development stage										
Treatment	125	0.001**	70.33	0.001**	81.71	0.001**	33.26	0.001**	270.33	0.001**
Treatment x Genotypes	0.000	0.995	7.410	0.011*	4.450	0.043*	8.660	0.006*	45.360	0.001**
Treatment x weeks	18.440	0.001**	11.670	0.001**	10.820	0.001**	8.120	0.001**	66.550	0.001**
Genotypes x weeks	4.220	0.013*	2.490	0.080	1.890	0.152	9.560	0.001**	61.890	0.001**
Treatment x Genotypes x weeks	4.630	0.009**	1.220	0.321	2.680	0.064	8.860	0.001**	9.190	0.001**

Statistically significant differences (two-way ANOVA) between genotypes, treatment and drought weeks are denoted as *p<0.05 and **p<0.001

Table 5.5: **Two-way ANOVA for morphological data for DipC and TN**

	Leaf Area		Leaf Number		Plant height	
Interactions	**F value**	**P value**	**F value**	**P value**	**F value**	**P value**
Drought initiated at vegetative stage						
Treatment	18.68	0.001**	13.68	0.001**	63.96	0.001**
Treatment x Genotypes	0.280	0.604	1.800	0.193	15.170	0.001**
Treatment x weeks	5.160	0.015*	9.810	0.001**	16.020	0.001**
Genotypes x weeks	1.670	0.211	4.790	0.019*	0.850	0.442
Treatment x Genotypes x weeks	0.090	0.910	3.970	0.034*	3.820	0.038*
Drought initiated at reproductive stage						
Treatment	1.71	0.205	3.26	0.085	5.50	0.028*
Treatment x Genotypes	3.260	0.085	0.370	0.551	11.540	0.003**
Treatment x weeks	0.520	0.600	4.680	0.020*	1.380	0.273
Genotypes x weeks	5.510	0.011*	16.170	0.001**	0.380	0.690
Treatment x Genotypes x weeks	0.960	0.400	0.360	0.705	2.890	0.077
Drought initiated at pod development stage						
Treatment	9.27	0.005*	87.15	0.001**	1.49	0.231
Treatment x Genotypes	4.060	0.050*	6.640	0.015*	3.740	0.062
Treatment x weeks	4.850	0.007**	9.680	0.001**	0.400	0.752
Genotypes x weeks	2.480	0.080	0.740	0.538	0.250	0.860
Treatment x Genotypes x weeks	5.420	0.004**	0.740	0.538	0.920	0.443

Statistically significant differences (two-way ANOVA) between genotypes, treatment and drought weeks are denoted as *p<0.05 and **p<0.001

5.2.1.10 Seed and pod number

Plants under drought stress imposed at the vegetative and reproductive showed significant reductions ($p<0.01$) in pod and seed number in DipC, TN and peanut compared to control plants. Whereas, the pod development stage showed significant reduction ($p<0.01$) for pod and seed number in DipC but not in TN and peanut (Table 5.6). Highest reductions in seed number in treatment plot compared to control plot at the vegetative stage were observed in DipC (68.36%), followed by TN (61.42%) and peanut (60.01%). At the reproductive stage, highest reduction was seen in DipC (97%), followed by peanut (74.62%) and TN (56.25%). Pod development stage showed the maximum reduction of 34.48% in DipC. No significant reduction was observed in TN and peanut at the pod development stage. Drought stress commenced at the vegetative and reproductive stage lead to greater reduction in seed and pod number compared to pod development stage. This is not surprising as the duration of drought period at the vegetative and reproductive was longer compared to pod development stage. The number of seeds per plant destined for final harvest to a large extent determined the differences in yielding levels of the studied crops under drought conditions.

Table 5.6: **Effect of drought stress on seed and pod number**. Data are the means ±S.E. of three replicates.

Plants	No. Seeds (plant $^{-1}$)		Reduction (%)	No. Pods (plant $^{-1}$)		Reduction (%)
	Irrigated	Drought		Irrigated	Drought	
Drought Initiated at vegetative stage						
DipC	26.33±0.33	8.33±1.20**	68.36	38.33±0.33	14.66±0.88**	61.75
TN	23.33±0.33	9±0**	61.42	23..33±0.33	10±0.57**	57.13
Peanut	21.66±2.18	8.66±0.8**	60.01	15±1.15	5.33±0.66**	64.47
Drought initiated at reproductive stage						
DipC	27.66±1.85	0.66±0.33**	97.59	34±0	10.33±0.33**	69.62
TN	21.33±0.33	9.33±0.33**	56.25	22±1.15	10.33±0.88**	53.05
Peanut	22.33±1.20	5.66±0.66**	74.62	16±0.57	4.66±0.66**	70.88
Drought initiated at pod development stage						
DipC	29±1.15	19±0.57**	34.48	53.33±2.02	24±2.30**	55.00
TN	22±2.08	21.33±0.66	3.03	23±1.73	22.33±1.20	2.91
Peanut	19.33±2.33	16.33±2.33	15.52	14.66±0.66	11.33±1.66	22.71

Means with single and double astericks indicate significane level of *p<0.05 and **p<0.01, respectively between groups

5.2.1.11 Seed yield, biomass and harvest index

The drought stress caused significant reduction ($p<0.01$) in seed yield at all three developmental stages for three studied crops (DipC, TN and peanut) (Table 5.7). Drought stress at the vegetative, reproductive and pod development stage resulted in 72.13%, 98.60%, 30.61% reductions in seed yield for DipC, respectively. Whereas, 78.07%, 77.13%, 49.66% reductions were seen in TN, respectively. On the other hand, peanut showed 71.30%, 78.38%, 29.49% reductions in seed yield, respectively. The effect of drought on seed yield is primarily due to significant reductions in seeds per plant (Table 5.6). From the results, DipC had slightly less reduction in yield compared to TN under drought stress at the vegetative and pod development stage. Whereas, DipC had more reduction in yield compared to TN under drought stress at the repro-

ductive stage. In comparison with peanut, both DipC and TN had greater reduction in seed yield under drought stress at the vegetative and pod development stage. However, at the reproductive stage, peanut had more reduction in seed yield compared to TN but less in comparison to DipC.

A significant reduction ($p<0.05$) in shoot biomass was observed in DipC, TN and peanut for drought stress imposed at the vegetative and reproductive stage (Table 5.7). Shoot dry weight at pod development stage for treatment plot plants showed significant difference ($p<0.05$) for TN and peanut but not for DipC. Whereas, significant reduction for harvest index ($p<0.05$) was seen in DipC at all three developmental stages but not in TN and peanut.

Table 5.7: **Effect of drought stress on seed yield, shoot dry weight and harvest index**. Data are the means ±S.E. of three replicates.

Plants	Seed Yield (g plant $^{-1}$)		Reduction	Shoot dry weight (g)		Reduction	Harvest Index (%)		Reduction
	Irrigated	Drought	(%)	Irrigated	Drought	(%)	Irrigated	Drought	(%)
Drought initiated at vegetative stage									
DipC	6.1±0.1	1.7±0.55**	72.13	4.8±0.1	2.8±0.8*	41.67	55.96±0.10	37.29±0.98**	33.37
TN	7.6±0.1	1.66±0.12**	78.07	4.56±0.06	1.2±0.66*	73.72	62.46±0.03	51.66±6.73	17.29
Peanut	7.43±0.49	2.13±0.27**	71.30	10.8±2	2±0.3*	81.48	46.33±1.01	46.10±4.08	0.49
Drought initiated at reproductive stage									
DipC	9.53±0.69	0.13±0**	98.60	8.6±0.51	3.16±0.38**	63.18	52.50±3.32	3.70±1.89**	92.93
TN	7.43±0.18	1.7±0.45**	77.13	6.3±0.11	3.13±0.61*	50.26	54.11±0.39	35.41±7.39	34.56
Peanut	8.63±0.6	1.86±0.1**	78.38	11.16±1.33	3.55±0.60**	68.21	43.99±3.54	35.16±7.49	20.06
Drought initiated at pod development stage									
DipC	8.16±0.17	5.66±0.47*	30.61	10.45±0.63	9.82±0.83	5.98	45.31±2.12	35.10±3.62*	22.53
TN	9.73±0.66	4.9±0.49**	49.66	7.1±1.18	4.5±0.79*	36.62	58.27±3.86	52.61±4.15	9.70
Peanut	7.23±0.64	5.1±0.1*	29.49	12.9±0.80	6.5±0.40**	49.61	35.85±0.79	44.05±1.16**	-22.89

Means with single and double astericks indicate significane level of *$p<0.05$ and **$p<0.01$, respectively between groups.

5.3 Discussion

5.3.1 Experimental set-up and moisture distribution

A pot experiment was set up in growth chambers using three moisture regimes (75% field capacity (FC) as control and 50% and 25% field capacity as drought treatment plots). Screening under natural drought condition in the target environment is complex because of irregular and erratic weather conditions and drought response but screening under controlled stress environments such as growth rooms is more feasible and manageable. Selection response in the target population of environments under natural stress can be considered a correlated response to selection in the managed stress environment [280]. However, for pot experiments it is worthwhile noting that small variations in growing conditions will amplify differences in plant growth when these conditions are below the optimum. Hence, pot experiments under stress conditions are grounds for amplified error variance and artifacts (http://www.plantstress.com/methods/PotExp.htm).

In order to keep the growing conditions consistent, soil induced drought stress was imposed in three different ways: (1) normal irrigation (control) (75% FC); (2) watering at 50% FC and (3) watering at 25% FC. Soil moisture for controls were kept at 75% FC throughout the study, while for treatment plot at the vegetative and reproductive stages soil water was kept at 50% field capacity for week 1 of drought stress and lowered to 25% FC from the beginning of week 2. Whereas in treatment plot for the pod development stage soil water was kept at 50% FC for four weeks and lowered to 25% from fifth week onwards. This is to ensure plants survive the drought stress period for longer time at pod development stage as this stage is longer than the vegetative and reproductive stages. The pots were weighed everyday to maintain FC. Moisture lost through evapo-transpiration was restored by re-watering to keep

constant FC for control (75%) and treatment plots (50-25%).

5.3.2 Effect of drought stress in two bambara groundnut genotypes and peanut at different developmental stages

Drought stress lead to the reduction in relative water content at all three stages (vegetative, reproductive and pod development) in both bambara groundnut genotypes (DipC and TN). Although both genotypes of bambara groundnut had greater reduction in RWC (3-9%) compared to peanut (1-2%), they still managed to maintain relatively high RWC despite the development of drought stress. Similar observation has observed in other studies on bambara groundnut, where under drought bambara groundnut maintained high RWC [90, 116]. Similar results have also been reported in rice (*Oryza sativa*) [281], peanut (*Arachis hypogea*) [282] and maize (*Zea mays*) [283]. This is a crucial trait as it indicates the drought tolerance of bambara groundnut. Species which exhibit restricted changes in RWC per unit reduction of water potential are often considered to be relatively drought resistant [116]. From this study, it is clear that both peanut and bambara groundnut can maintian high RWC under drought, which indicates their tolerance to drought stress. Flexas et al. [284] and Chaves et al. [285], reported relative water content in the leaves, stomatal conductance and transpiration under drought stress are highly correlated.

It has been reported in other studies that drought stress leads to reduction in stomatal conductance, photosynthesis and transpiration [101, 227, 286–288]. A significant reduction in stomatal conductance in bambara groundnut under drought stress in both genotypes observed in this study is in accord with the observations reported by [90, 116]. It is thought that bambara groundnut maintains its turgor through a combination of osmotic adjustment and effective stomatal regulation of water loss [90]. Stomata are known to respond to chemical signals such as ABA produced by dehy-

drating roots and stomatal regulation is closely linked to soil moisture content [289]. It is known from the previous studies that stomatal conductance is more sensitive to drought stress in comparison to RWC [276]. In this study, stomatal conductance for the treatment plot remains constantly low in comparison with irrigated (control) plants. These observations imply that the regulation of stomatal closure is one of physiological response to drought stress in bambara groundnut [24]. Reduction in stomatal conductance followed a constant decreasing trend in drought stressed plants at all three developmental stages (vegetative reproductive and pod development) (see tables 5.1 5.2 and 5.3). The results are consistent with the observations reported by Vurayai et al. [116]. However, at week 6 (103 DAS) of pod development stage, no significant difference was observed between treatment and control plots in both DipC and TN. Chai et al. [12], reported similar observation where bambara groundnut at 65 DAS to 72 DAS of drought stress showed rapid decline in stomatal conductance, followed by a relatively slow and steady decline between 72 DAS and 84 DAS. Collinson et al. [90], stated that stress-induced stomata closure is believed to be accompanied by osmotic adjustment. Stomatal conductance show little or no change once the decline of stomatal conductance reaches a threshold value due to drought stress, as its speculated that plants are keeping their stomata open for carbon uptake while maintaining their plant water status by osmotic adjustment [90]. From the results, it indicates that both DipC and TN have shown drought avoidance properties where under drought stress, there was significant reduction in stomatal conductance. Similar results has been observed in maize (*Zea mays*) [115], sugarcane (*Saccharum spp.*) [290], and chickpea (*Cicer arietinum*) [291]. It is worthwhile noting that given the low of number data points in each stage might limit the potential of observing trends that might be of relevance. On the other hand, drought stressed plants in peanut have shown greater reductions in stomatal conductance at vegetative (70.58%) and reproductive (94.24%) stages in

comparison to TN (17.75% and 54.55%, respectively). However TN showed greater reduction at pod development stage (91%) than peanut (90%) (Figure 5.3). On the other hand, peanut had greater reductions in stomatal conductance compared to DipC (40.46%, 77.32% and 40.68%, respectively) at all three stages. It worthwhile noting that peanut control plants after week 4 (89 DAS) at pod development stage showed a rapid decline in all gas exchange measurements (stomatal conductance, photosynthesis and transpiration). This could be the effect of the growing conditions and the interaction of pot size with plant size (http://www.plantstress.com/methods/PotExp.htm). Larger plants use more water than smaller plants. It could be possible that pot size (10 litres) and the amount of water given to peanut control plants was not sufficient after a certain increase in growth (http://www.plantstress.com/methods/PotExp.htm). In this study, this hypothesis was supported by peanut control plants which also showed symptoms of wilting and lack of nutrients. However, both genotypes and peanut have shown similar declining trends, where under drought stress significant reductions in stomatal conductance were observed, which indicates their ability to response to water limiting conditions through very similar drought mechanisms. Similar observations has been reported for drought study in peanut, where reduction in stomatal conductance has been observed [292].

Decreased stomatal conductance is known to lower net carbon dioxide assimilation rate, intercellular carbon dioxide and chloroplastic carbon dioxide tension. This carbon dioxide insufficiency will lead to reduction in photosynthetic efficiency and dry matter production that may have negative impact on plant growth and yield [116]. In this study, with stomatal conductance a constant reduction in photosynthesis and transpiration was observed in both genotypes as well as in peanut. This reduction in photosynthesis may be due to reduced substrate availability for plants to photosynthesise caused by stomatal closure [113]. A decrease in the conductance of mesophyll

cells due to drought stress leads to low conductance carbon dioxide and reduction in photosynthesis [276]. Rapid reductions in photosynthesis and stomatal conductance in peanut is consistent with the observations reported by Bhagsari et al. [293], implying regulation of stomata closure for water loss is one of the mechanism employed by plants to regulate water loss. It is clear that drought stress leads to stomatal closure, followed by parallel decrease in photosynthesis. However, stomatal conductance is not just controlled by soil water availability but there are other complex interactions of intrinsic and external factors involved [93]. Thus the need for further research is imperative to fully understand the mechanism involved in stomatal regulation under drought stress in plants. Throughout the drought study, a significant reduction ($p<0.05$) in transpiration in the treatment plot was observed at vegetative, reproductive and pod development stage for DipC and peanut, whereas for TN, reduction was only seen at pod development stage. Drought avoidance is the mechanism achieved by plants through reducing water loss from transpiring leaves [276]. Reduction in transpiration rate has been reported with the increase in duration of drought stress indicating that the plants under moisture stress are postponing tissue dehydration [276, 277]. From the results, it suggest that DipC and peanut has shown properties of postponing tissue dehydration by showing significant reduction in transpiration at all three developmental stages. Whereas TN, only showed properties of postponing tissue dehydration at the pod development stage. Significant reduction ($p<0.05$) in chlorophyll content from SCMR readings in DipC and TN at the pod development stage are in agreements with other studies where water stress lead to reduction in chlorophyll content of rosemary leaves [294].

Results of yield showed that despite bambara groundnut is known to be drought tolerant, water stress still able to effect yield in both DipC and TN. Similar results has been reported in literature for bambara groundnut [14, 113], which all reported reduced

seed yield in bambara groundnut landraces in response to limited water availability. However, what has been found remarkable is the ability of bambara groundnut to still produce yield in an very limiting water conditions (50-25% field capacity) (Table 5.7). An representation of the indicative growth of DipC, TN and peanut at different developmental stages under drought stress and irrigated conditions is shown in Figure 5.9. According to Berchie et al. [14], this indicates the ability of bambara groundnut to be drought tolerant and further justifies the need for more research on the crop. Results of yield from this study, for both DipC and TN, showed that despite DipC having more reductions in seed and pod number compared to TN (Table 5.6), it had less reduction in seed yield for drought stress imposed at vegetative and pod development stage (Table 5.7). This implies that although fertility of DipC reduced more than TN, the pod filling efficiency of DipC in the vegetative and pod development state is less affected by drought stress in comparison to TN. Peanut, on the other hand, showed the maximum relative reduction in biomass production under drought stress at all three developmental stages in comparison to DipC and TN, which indicates its inefficiency to produce biomass under drought stress. Similar observations for peanut under drought stress has been reported [282]. Despite peanut having greater reductions in biomass compared to DipC and TN under drought stress, it has been able to experience less reductions in seed yield in comparison to DipC and TN at the vegetative and pod development stage. However, TN had lesser reduction in seed yield at the reproductive stage compared to peanut. Peanut produced more yield under drought stress at all three stages: vegetative (2.13 g plant^{-1}); reproductive (1.86 g plant^{-1}) and pod development (5.01 g plant^{-1}) compared to TN (1.66 g plant^{-1}, 1.7 g plant^{-1} and 4.9 g plant^{-1}, respectively). On the other hand, peanut produced more yield at vegetative (2.13 g plant^{-1}) and reproductive stage (1.86 g plant^{-1}) than DipC (1.7 g plant^{-1} and 0.13 g plant^{-1}, respectively) but lesser at the pod development stage

(5.01 g plant^{-1}) compared to DipC (5.66 g plant^{-1}).

5.4 Conclusion and potential traits for future programmes

This study showed that despite water limiting conditions, bambara groundnut genotypes can still produce yield. Drought stress lead to reductions in growth indices of plant height, leaf number and leaf area, thus minimising water losses. In addition, bambara groundnut genotypes show classic drought avoidance mechanisms through stomatal closure, thus reducing transpirational losses. Drought stress was shown to reduce seed yield through reduction in seed and pod number. Although both DipC and TN were shown to be productive under drought stress at the vegetative and pod development stage, DipC was highly sensitive to drought stress at the reproductive stage with a 98% overall reduction in seed yield. Furthermore, both bambara groundnut and peanut have shown similar declining trends, where under drought stress significant reductions in stomatal conductance, photosynthesis and transpiration were observed, which indicates their ability to response to water limiting conditions through very similar drought mechanisms. In addition, reductions in seed yield was observed in both genotypes as well as peanut.

As the need to identify traits for genetic improvement for drought tolerance in bambara groundnut increases, this physiological study will assist in dissecting the complexity of the responses of this plant to drought stress. In addition with physiology data, the transcriptomic data has also been generated for this drought study (refer to chapter 6). Identification of differentially expressed genes between the conditions will help to reveal potential components in stress response. The observed reduction in stomatal conductance under drought stress in both genotypes at all three devel-

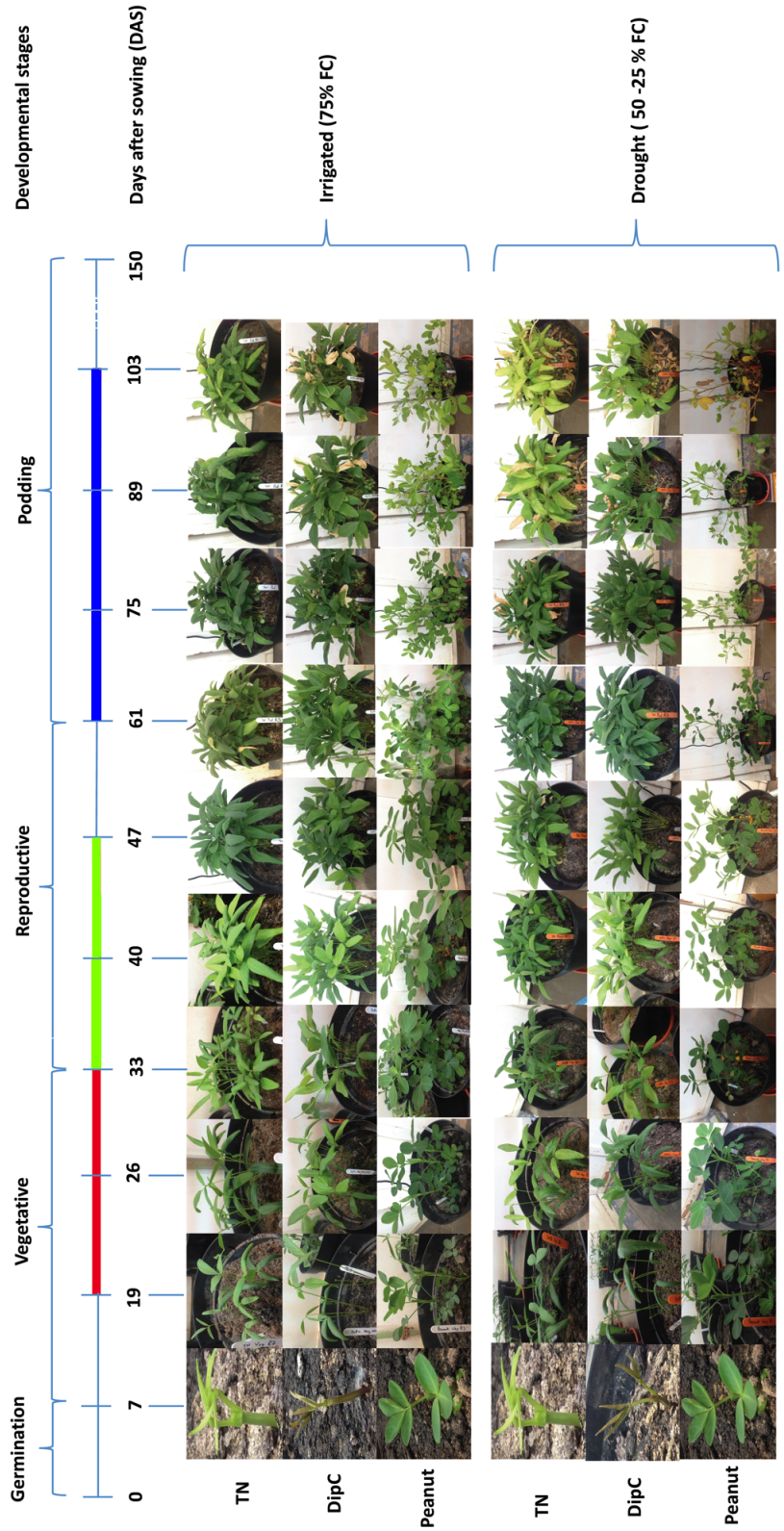

Figure 5.9: **Indicative growth of TN, DipC and Peanut at different development stages under irrigated (control) and drought conditions**. The red, green and blue bar denotes the days at which morpho-physiological measurements were taken for vegetative, reproductive and pod development (podding) stage respectively. Each stage is a separate experiment, where drought start time for each stage differs. Each stage represents different plants. The length of the drought period for vegetative and reproductive was two weeks, with morpho-physiological measurements taken at a interval of 7 days over a period of two weeks. For pod development stage, drought stress was six weeks long, with morpho-physiological measurements taken at a 14 days interval over a span of six weeks.

opmental stages from this study will help in identifying genes related to the trait using transcriptomic approaches to study gene expression. From the physiological point of view, there was significant reduction in photosynthesis under drought stress in both genotypes at all three developmental stages. However, the plants still managed to photosynthesise even with minimal water supply (25% FC). The identification of genes related to photosynthesis trait from the transcriptomic study will explain the observed phenomenon and will assist in understanding the molecular mechanism governing drought response in bambara groundnut.

Chapter 6

Effects of drought stress on the transcriptome of bambara groundnut at different developmental stages

6.1 Introduction

Drought stress as discussed in chapter 2 is a major abiotic stress that inhibits proper plant growth and crop productivity. Several transcriptomic studies have been done on plant species with the aim to understand the molecular mechanisms underlying the ability of crops to produce viable yields under drought conditions [93, 176, 295–299]. To mitigate drought stress plants have developed various strategies, such as generation of larger and deeper root systems [289], regulation of stomatal closure to reduce water loss [285], accumulation of compatible solutes and protective proteins and increased expression levels of various osmoprotectants and antioxidants[93]. The products of the drought-inducible genes can be classified into two groups: (1) the genes that produce products (either directly or indirectly) with functions in drought tolerance. These include molecules such as late embryogenesis abundant (LEA) proteins, osmoprotectants, antioxidants, chaperones, antifreeze proteins, mRNA-binding proteins, key enzymes for osmolyte biosynthesis, water channel proteins, sugar and

proline transporters, detoxification enzymes, and several proteases; and (2) the gene regulatory proteins, which functions in regulating the signal transduction and gene expression of entire responses to drought. These include, various transcription factors (such as MYB, WRKY, bZIP, Zinc-fingers, NAC, AREB, DREB2, HOX) , protein kinases, protein phosphatases, enzymes involved in phospholipid metabolism and other signaling molecules such as calmodulin-binding protein. The transcription factors can govern expression of a certain drought inducible gene either cooperatively or independently [98, 99, 300–302].

With the availability of NGS technologies, it is now possible to study the transcriptome of a species whose genome has not been sequenced yet [42] which will greatly aid in elucidating regulatory networks [43] and understanding the genetic basis of complex traits [44] in non-model species (such as bambara groundnut). To successfully assemble a transcriptome without the aid of reference genome requires robust computational method and substantial sequence data. This is achieved via *de novo* transcriptome assembly [41], which is used to assemble short RNA-Seq reads without a reference genome. It has been widely used for transcriptomic studies in organisms whose genome information is not available [46].

An approach to identifying genes conferring drought avoidance and tolerance is to study species that are already very tolerant to semi-arid conditions. In this regard, bambara groundnut is a potential candidate (refer to 2.1.1 for details). Two bambara groundnut genotypes; DipC and TN, derived from different landraces were used for this experiment. This study aims to: (1) identify genes and expression patterns at different developmental/growth stages and time points in response to drought stress which enable bambara groundnut to grow under semi-arid conditions; (2) to compare the transcriptome of the two genotypes of bambara groundnut to identify what is common and how they differ in their response to drought stress; (3) understanding

the drought mechanisms underlying the ability of crops to produce viable yields under drought conditions and (4) identifying the genes for the trait of interest in relation to drought stress obtained from physiology data. No study has been done on the transcriptome profiling of bambara groundnut under drought stress yet. This study will build upon the data already obtained from the physiology work for this experiment (refer to chapter 5).

6.2 Results

The experiment was divided into three developmental stages - vegetative, reproductive and pod development. The drought was imposed at 19, 33 and 61 DAS for vegetative, reproductive and pod developmental stages, respectively. Each stage is a separate experiment (see 3.2.2 for details on experimental design) . Leaf sampling was done at the start and end period of the drought treatment (drought and control plants) for DipC and TN, for the extraction of RNA. For the gene expression study, leaf samples from vegetative and reproductive stage were send for RNA-sequencing (RNA-seq). This chapter studies the gene expression profiling of the two genotypes under drought stress at the vegetative and reproductive stages. Table 6.1 list the days at which leaf sampling was done for the gene expression study.

Table 6.1: **RNA sampling for gene expression study.**

Developmental stages	**Measurements at DAS***	**Drought weeks**	**Field Capacity (%)**
Vegetative stage (7-33 Days)	19 DAS*	Week 0	75%
	26 DAS*	Week 1	50%
	33 DAS*	Week 2	25%
Reproductive stage (33-61 Days)	33 DAS*	Week 0	75%
	40 DAS*	Week 1	50%
	47 DAS*	Week 2	25%

*DAS = Days after sowing. The days at which RNA sampling was done are highlighted in red.

6.2.1 Read quality checking, statistics and annotation

Illumina HiSeq4000 platform was used to generate 75bp paired-end sequence reads for RNA-seq analysis. Table 6.2 shows the number of reads for each samples before and after quality trimming. The trimmed reads were further used to assemble the transcriptome. For transcriptome assembly all reads were combined into a single

input in order to obtain a consolidated set of contigs (transcripts) which will aid in comparing the transcriptomes across all samples. Transcriptome assembly resulted in 56,772 non-redundant potential gene sequences. The tool "dammit" (version 0.2.2) was used to annotate the assembled sequences, resulting in 43,467 annotated transcripts. Trimmed sequences/reads were mapped back to the assembled transcriptome using HISAT2 for expression estimation. StringTie was then used to assemble the read alignments into potential transcripts/contigs for each sample. This was followed by the StringTie "merge" step to assemble the sequences into non-redundant transcripts which are consistent across the multiple RNA-seq samples. The StringTie merge step produced an annotation (GTF) file which identified 54,225 potential "gene" sequences in total across all samples. Read counts were generated for these 54,225 sequences using htseq-count. A workflow of the RNA-seq analysis for this study can be found in chapter 3 figure 3.4.

Table 6.2: **Breakdown of RNA-seq reads for leaf samples of DipC and TN under control (irrigated) and drought conditions at different time points and developmental stages.** The number of raw reads generated from RNA-seq is shown as raw reads. Trimmed reads are reads that were kept after quality checking and the ones that did not pass the quality checking processes are the discarded reads.

Samples	**Raw Reads**	**Trimmed reads**	**Trimmed reads(%)**	**Discarded Reads**	**Discarded reads (%)**
Vegetative stage					
DipCcontrol1-19	31652142	31650447	99.995	1695	0.005
DipCcontrol2-19	36004105	36001845	99.994	2260	0.006
DipCcontrol3-19	30401159	30399651	99.995	1508	0.005
DipCcontrol1-33	32720112	32718449	99.995	1663	0.005
DipCcontrol2-33	33406349	33404285	99.994	2064	0.006
DipCcontrol3-33	27427307	27425929	99.995	1378	0.005
DipCDrought1-33	29563236	29561964	99.996	1272	0.004
DipCDrought2-33	30177880	30176336	99.995	1544	0.005
DipCDrought3-33	28895797	28894390	99.995	1407	0.005
TNcontrol1-19	29372457	29371479	99.997	978	0.003
TNcontrol2-19	32038373	32037440	99.997	933	0.003
TNcontrol3-19	30896649	30895068	99.995	1581	0.005
TNcontrol1-33	28586049	28584751	99.995	1298	0.005
TNcontrol2-33	32446405	32444773	99.995	1632	0.005
TNcontrol3-33	28887835	28886401	99.995	1434	0.005
TNDrought1-33	32136028	32133951	99.994	2077	0.006
TNDrought2-33	26742166	26740893	99.995	1273	0.005
TNDrought3-33	31415186	31413648	99.995	1538	0.005
Reproductive stage					
DipCcontrol1-33	32256354	32254821	99.995	1533	0.005
DipCcontrol2-33	31976704	31975292	99.996	1412	0.004
DipCcontrol3-33	30602089	30600768	99.996	1321	0.004
DipCcontrol1-47	31374748	31372927	99.994	1821	0.006
DipCcontrol2-47	25586518	25585260	99.995	1258	0.005
DipCcontrol3-47	29816306	29814992	99.996	1314	0.004
DipCDrought1-47	27162028	27160654	99.995	1374	0.005
DipCDrought2-47	28763000	28761075	99.993	1925	0.007
DipCDrought3-47	31940347	31938626	99.995	1721	0.005
TNcontrol1-33	32859355	32857952	99.996	1403	0.004
TNcontrol2-33	36287704	36286071	99.995	1633	0.005
TNcontrol3-33	38239405	38237525	99.995	1880	0.005
TNcontrol1-47	25857225	25855972	99.995	1253	0.005
TNcontrol2-47	28732455	28730845	99.994	1610	0.006
TNcontrol3-47	30781092	30779350	99.994	1742	0.006
TNDrought1-47	30338549	30337174	99.995	1375	0.005
TNDrought2-47	32700209	32698617	99.995	1592	0.005
TNDrought3-47	29668885	29666570	99.992	2315	0.008

6.2.2 Gene expression under control (irrigated) conditions

It is relevant to consider the state of the genotypic transcriptomes at different time points and developmental stages before any drought treatment has taken place. The genes with TMM (Trimmed mean of M-values) normalised average counts per million (CPM) > 50 across all biological replicates were chosen as highly expressed under irrigated conditions at different time points and growth stages. CPM is the descriptive measures for the expression level of a gene under a certain condition. At this threshold, ~1000 genes were left out of 54,225 genes across all samples. At this cut-off each genotype (DipC and TN) had a little over 45 genes models with annotations related to abscisic acid (ABA) signaling and drought responses in all four control conditions (i.e. control conditions at 19 and 33 DAS at vegetative stage and control conditions at 33 and 47 DAS at reproductive stage) (Appendix 16-23). These include homologues of many of the genes involved in the ABA signaling pathway (such as protein PP2C family, bZIP, MYB transcription factors), osmoprotectants (such as aspartate proteases and beta-fructofuranosidase (glycoside-hydrolase family)), several antioxidants (belonging to peroxidase, glutaredoxin, thioredoxin, ferritin and glutathione S-transferases (GST) gene family) and various drought induced genes such as late embryogenesis abundant (LEA), heat shock proteins (HSP), myo-inositol-1-phosphate synthase (MIPS) and dehydrins gene family.

6.2.3 Differentially expressed genes at vegetative stage

At vegetative stage, for both DipC and TN, RNA-seq was done at time points 19 and 33 DAS (refer to 3.6 for RNA sampling details). A total of 1882 and 550 genes were found to be significantly differentially expressed (DE) (FDR < 0.05, log fold change (log FC) ≥ 1) in control (irrigated) condition at 33 DAS compared to control

(irrigated) condition at 19 DAS in DipC and TN, respectively, through differential expression analysis using edgeR (Appendix 24 and 25). Among these genes, expression of 1299 and 274 genes were down-regulated and 583 and 276 genes were up-regulated, respectively. Of these differentially expressed genes (DEG), 70 genes were found to be overlapping between DipC and TN (figure 6.4 and Appendix 26). On the other hand, significantly DEG in drought compared to control condition at 33 DAS for DipC and TN accounts for 586 and 1899 genes, respectively (Appendix 27 and 28). Among these genes, 196 and 1012 were down-regulated and 390 and 887 were up-regulated under drought stress, respectively. Of these genes, only 139 genes were found to be overlapping between DipC and TN (Table 6.3 and figure 6.4). TN showed higher numbers of DEG in drought vs control condition at 33 DAS in comparison to DipC. Whereas DipC, showed higher number of DEG at control-33 DAS vs control-19 DAS condition compared to TN (Figure 6.4). This indicates that TN has activated more genes when subjected to drought stress in comparison to DipC, whereas DipC has activated more genes when comparison between two control conditions were made at different time points. FDR is the expected proportion of erroneous rejections among all rejections [303], thus using lower FDR thresholds reduces the chance of significant values being false positives. Genes were selected for subsequent analysis by satisfying a log FC ≥ 1 and FDR < 0.05. A multidimensional scaling (MDS) plot was used to identify the similarity or dissimilarity between samples. The sample is represented as a point on the plot, and the distance between points corresponds to the similarities between them. This method calculates distances between samples based on log2 fold changes between every pair. Thus, samples of the same experimental condition are closer to one another, while dissimilar samples are further apart [304]. Figures 6.1 and 6.2 show that each biological replicate set are clustered in their respective experimental condition and time-point in DipC and TN, respectively. The MA plot shows the log-fold

change (i.e., the log ratio of normalized expression levels between two experimental conditions (i.e. developmental stage: control-33 DAS vs control-19 DAS) and (Treatment: drought-33 DAS vs control-33 DAS)) and against the log CPM in both DipC and TN (Figure 6.3).

The top 100 up and down regulated genes according to log FC in drought compared to control (irrigated) conditions at time point 33 DAS in DipC and TN are listed in Tables 6.4 and 6.5. Based on the DEG analysis in drought versus control condition at time point 33 DAS, drought-related genes could be divided into following categories:

1**. ABA signal transduction associated gene expression**

Reduced water availability causes the production of ABA, the phyto-hormone which initiates stomatal closure and influences other aspects of plant growth and physiology. It is responsible for regulating a broad range of genes during drought [305]. In this study, changes in gene expression of genes that are involved in triggering ABA signaling were observed in DipC and TN plants under drought stress. In DipC, two homologues of protein phosphatase 2C (PP2C) gene family were up-regulated and three were down-regulated under drought stress compared to control at 33 DAS (Appendix 27). In addition one gene in the ABA_WDS family and a homologue of a known sub-type of cytochrome p450 family involved in ABA catabolism [306, 307] were also induced under drought in DipC. In addition, family of transcription factors (TFs) related to ABA dependent pathways were differentially expressed (either up or down) under drought stress including six MYB genes, three NAC genes and one WRKY gene [99]. On the other hand, for TN, three homologues of PP2C gene family, one ABA_WDS gene and twelve homologues of cytochrome p450 gene family were up-regulated under drought stress (Appendix 28). In addition, family of transcription factors including 22 MYB genes, three bZIP genes, four NAC genes and four WRKY genes were differentially expressed (either up or down) in TN under drought stress at

33 DAS.

2. **Gene expression of osmoprotectants**

Role of osmoprotectants under drought stress has been widely studied in plants [93, 99, 156, 285, 308–311]. It plays a vital role in plant defense against drought stress by contributing towards continued water uptake and maintenance and membrane protection [306]. In this study, one trehalose phosphatase gene, two aspartate proteases genes, one raffinose synthase gene and five beta-fructofuranosidase (glycoside hydrolase family) genes were up-regulated under drought stress compared to control at 33 DAS in DipC (Appendix 27). In addition, three beta-fructofuranosidase genes were down-regulated. On the other hand, for TN, one aspartate proteases gene, two raffinose synthase genes and five beta-fructofuranosidase genes were up-regulated in response to drought stress at 33 DAS (Appendix 28). In addition, six aspartate proteases and 24 beta-fructofuranosidase genes were down-regulated.

3. **Expression of genes involved in antioxidant production**

Drought stress causes the production of reactive oxygen species (ROS), which leads to cell damage and oxidative stress. Upon a reduction in the amount of available water, plants close their stomata (plausibly via ABA signaling), which decreases the CO_2 influx, leading to formation of ROS [93]. Accumulation of antioxidants is a drought/damage tolerance strategy against ROS generated by plants [93, 246, 296, 312, 313]. In this study, antioxidants belonging to gene family of peroxidase, glutathione S-transferases (GST), glutaredoxin, thioredoxin and ferritin were induced under drought stress compared to control condition at 33 DAS. Up-regulation of genes encoding three homologues of peroxidase genes, one GST gene, two glutaredoxin genes and one ferritin gene were found in drought stressed plants at 33 DAS in DipC (Appendix 27). In addition, an homologue of GST and thioredoxin gene was down-regulated. In TN, homologues of four peroxidase genes, eight GST genes, two glutaredoxin genes, one

ferritin gene and two thioredoxin genes were up-regulated in response to drought stress at 33 DAS (Appendix 28). In addition, homologues of eight peroxidase genes and two GST genes were down-regulated.

4. **Expression of genes related to photosynthesis and glycolysis**

Photosynthesis and glycolysis are the basic physiological processes that provide ATP and intermediates for plant metabolism [306]. In this study, the expression level of a gene homologue encoding photosystem II - PsbP (photosystem II subunit P-1) family of proteins and another homologue related to photosynthesis belonging to phosphoribulokinase (PRK) family, which is a calvin cycle component, were highly repressed under drought stress compared to the control at 33 DAS in DipC (Appendix 27). Whereas the other gene related to photosynthesis such as light-harvesting complexes (LHC) was up-regulated. In case of TN, three light-harvesting complexes genes were highly down-regulated and one tri-carboxylic acid (TCA) cycle component namely phosphoenolpyruvate carboxylase (PEPcase) gene was up-regulated under drought stress at 33 DAS (Appendix 28).

With respect to glycolysis, several genes were up-regulated in DipC under drought stress compared to control at 33 DAS including, one pyruvate kinase (PK), two gene homologues of GDP-fucose-o-fucosyltransferase (O-fuct) family, one glucose-6-phosphate dehydrogenase (G6PD) gene and one glyceraldehyde 3-phosphate dehydrogenase (Gp_dh) gene (Appendix 27). Whereas one homologue of O-fuct gene family was down-regulated. On the other hand, one homologue of O-fuct gene family, two fructose-bisphosphate aldolase (glycolytic) genes and one Gp_dh gene were up-regulated under drought stress in TN (Appendix 28). In addition, one homologue of O-fuct gene family was down-regulated.

5. **Cell wall associated responses**

In this study, expression levels of number of cell wall associated genes were either

induced or repressed under drought stress compared to control at 33 DAS. An homologue of UDP-glucosyl transferase (UDPGT) gene family, two cellulose synthase genes and two fasciclin genes were found to be highly expressed under drought stress at 33 DAS in DipC (Appendix 27). Whereas one cell wall related gene xyloglucan endotransglycosylase (XET) was down-regulated. On the other hand, seven homologues of UDP-glucosyl transferase gene family were up-regulated under drought stress in TN. In addition, thirteen fasciclin genes, seven cellulose synthase gene and five XET genes were down-regulated (Appendix 28).

6. **Other genes induced by drought stress**

Other genes related to drought such as late embryogenesis abundant (LEA), heat-shock proteins (HSP), Lipoxygenase and Myo-inositol-1-phosphate synthase (MIPS) were up-regulated under drought stress compared to the control at 33 DAS in both DipC and TN (Appendix 27 and 28). In addition, several transporters (such as ABC, MFS1, MFS2, sodium ion, iron, sulfate and sugar transporters) were induced under drought stress. MFS are multidrug transporters that regulates stomatal movements and polar auxin transport by modulating potassium and proton fluxes in plant [314]. Numerous protein kinases (such as serine/threonine) and calmodulin related genes, were up-regulated under drought in both DipC and TN, which perform diverse and important functions in plant signal transduction (such as stomatal opening/closing, ABA response, pH regulation, nitrogen storage, plant development and defense) [93, 249, 315]. Furthermore, apart from ABA, other hormones such as cytokinin binding and auxin were induced under drought stress in both genotypes.

Out of the 139 genes that were found to be common between DipC and TN in response to drought at 33 DAS, a number of drought-related genes such as family of genes associated with oxidative stress (glutaredoxin, GST, ferritin and peroxidase), four osmotic adjustment related genes (raffinose synthase, aspartate proteases and

beta-fructofuranosidase), one photosynthesis related genes (LHC) and other drought-related genes such as LEA, lipoxygenase, MIPS and HSP were differentially expressed in both DipC and TN (Table 6.3)

Furthermore, the enriched GO-terms were assigned to three principal categories: molecular function (6 GO terms), biological process (37 GO-terms) and cellular component (5 GO-terms) in DipC under drought stress at 33 DAS (Appendix 29). The top significant GO-terms assigned (having at least 3 read counts) for biological process in DipC were related to: plant-type cell wall organization or biogenesis (GO:0071554), response to oxidative stress (GO:0006979), response to hydrogen peroxide (GO:0042542), response to light intensity (GO:0009642), response to reactive oxygen species (GO:0000302), cellular carbohydrate catabolic process (GO:0044275), cell wall organisation- (GO:0071555) and response to heat (GO:0009408) (see fig. 6.5 for top GO-terms assigned) (refer to Appendix 29 for the whole list of GO-terms assigned). Enriched GO-terms assigned for TN plants under drought stress at 33 DAS to three principal categories were: molecular function (39 GO terms), biological process (76 GO-terms) and cellular component (11 GO-terms) (Appendix 30). The top GO-terms assigned for biological process in TN were related to xylan biosynthetic process (GO:0045492), hemicellulose metabolic process (GO:0010410), cell wall organization or biogenesis (GO:0071554), cell wall polysaccharide metabolic process (GO:0010383), response to water deprivation (GO:0009414), response to oxidative stress (GO:0006979), secondary metabolic process (GO:0019748), response to abiotic stimulus (GO:0009628), response to oxidation-reduction process (GO:0055114) and external encapsulating structure organisation (GO:0045229) (see fig. 6.5 for top GO-terms assigned) (refer to Appendix 30 for the whole list of GO-terms assigned).

In addition, the enriched GO-terms assigned under biological process category when comparisons between control conditions (Control-33 DAS vs Control-19 DAS)

were made at different time points in DipC were mainly related to polysaccharide metabolic process (GO:0005976), hydrogen peroxide catabolic process (GO:0042744), response to external stimulus (GO:0009605), cell wall organisation (GO:0071555), secondary metabolic process (GO:0019748) and carbohydrate metabolic process (GO:0005975) (Appendix 31). Whereas, for TN, the GO-terms assigned were mainly related to: response to cytokinin (GO:0009735), chloroplast organisation (GO:0009658), cytokinin activated signaling pathway (GO:0009736) and plastid localisation (GO:0051644) (Appendix 32).

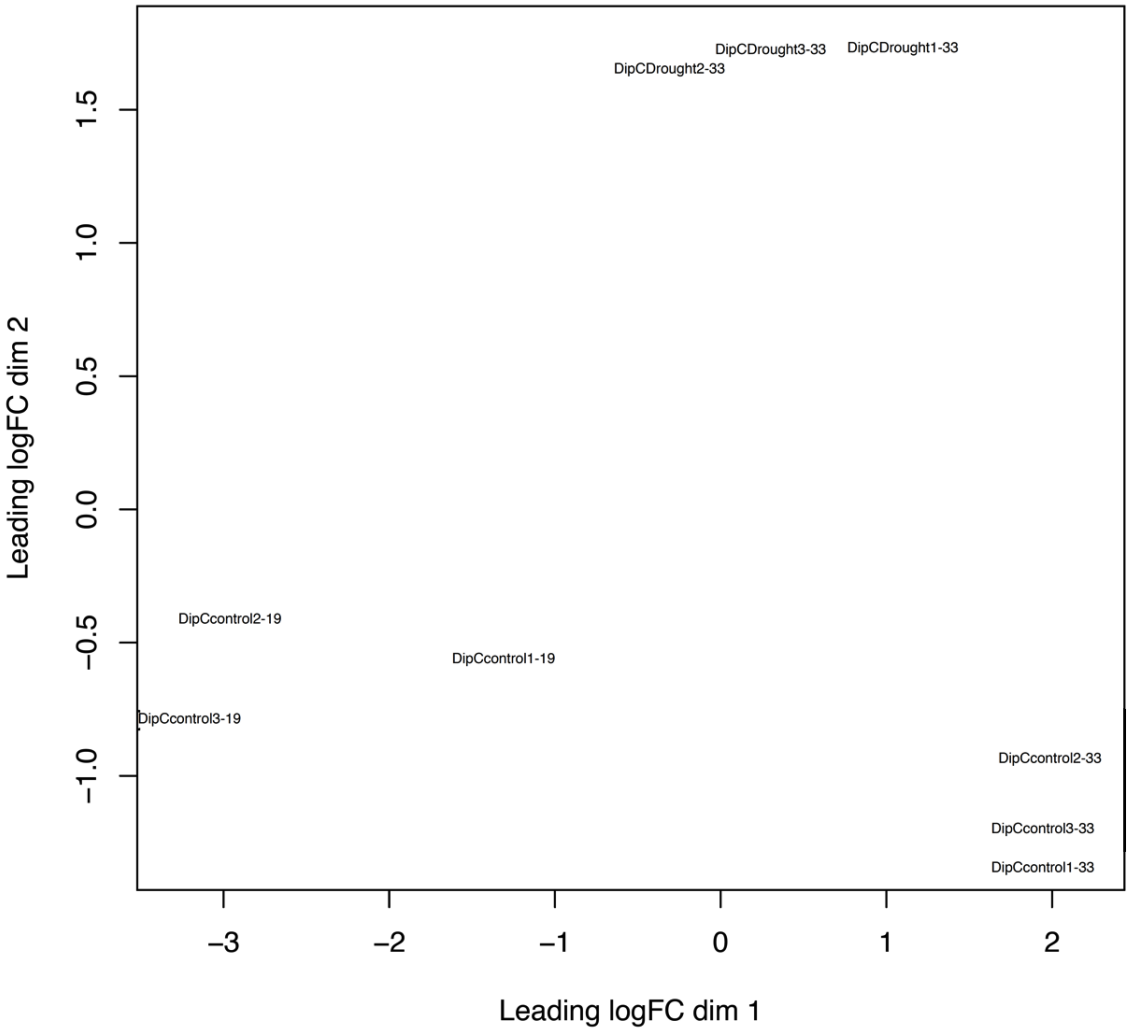

Figure 6.1: **Multidimensional scaling plot on differential gene expression of DipC samples at different conditions and time points at vegetative stage.** Replicates for control (irrigated) conditions at time points 19 and 33 DAS and drought conditions at time point 33 DAS are clustered in their respective experimental condition and time-point. This shows that samples between experimental conditions and time points were distinctively different.

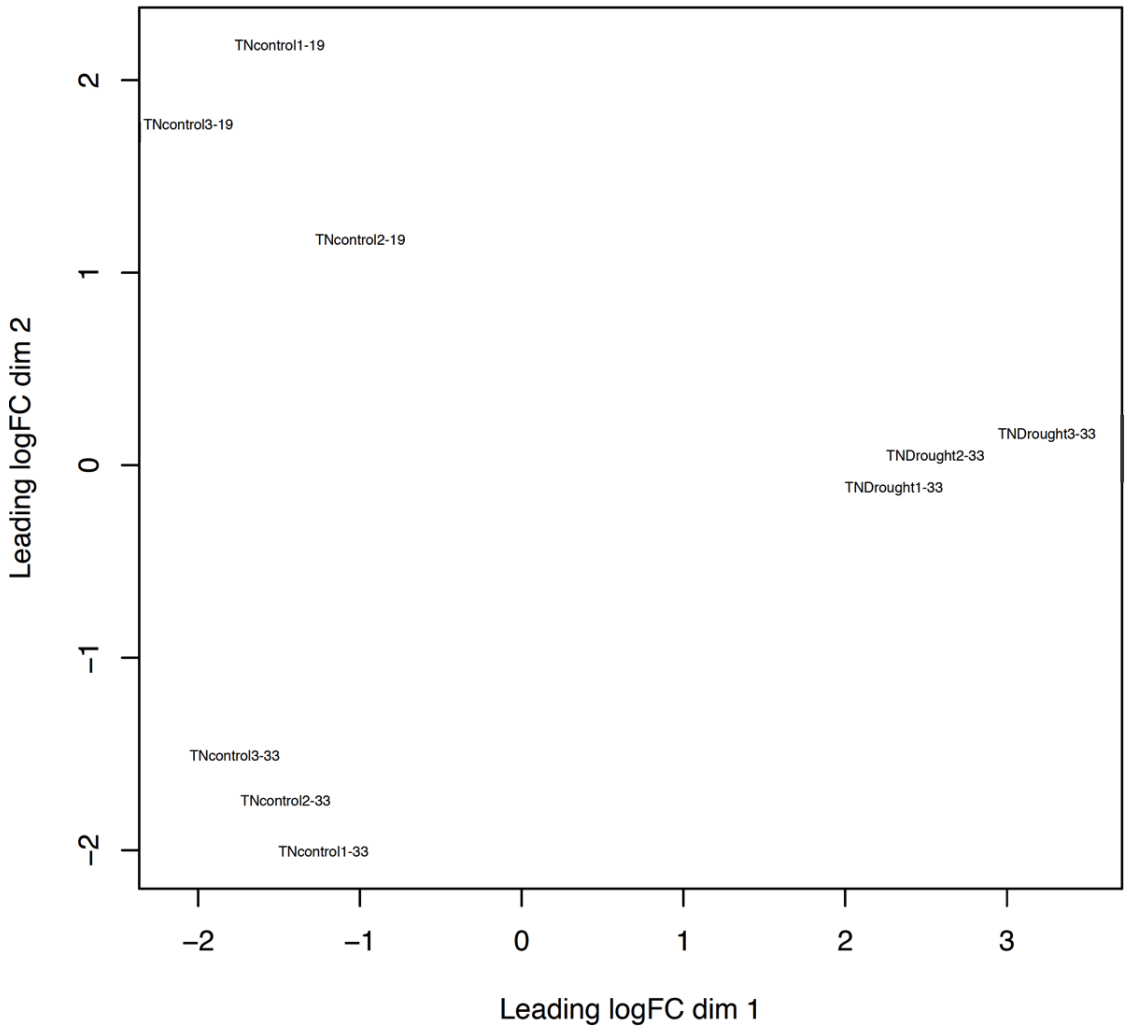

Figure 6.2: **Multidimensional scaling plot on differential gene expression of TN samples at different experimental conditions and time points at vegetative stage.** Replicates for control (irrigated) conditions at time point 19 and 33 DAS and drought conditions at time point 33 DAS are clustered in their respective experimental condition and time-point. This shows that samples between experimental conditions and time points were distinctively different.

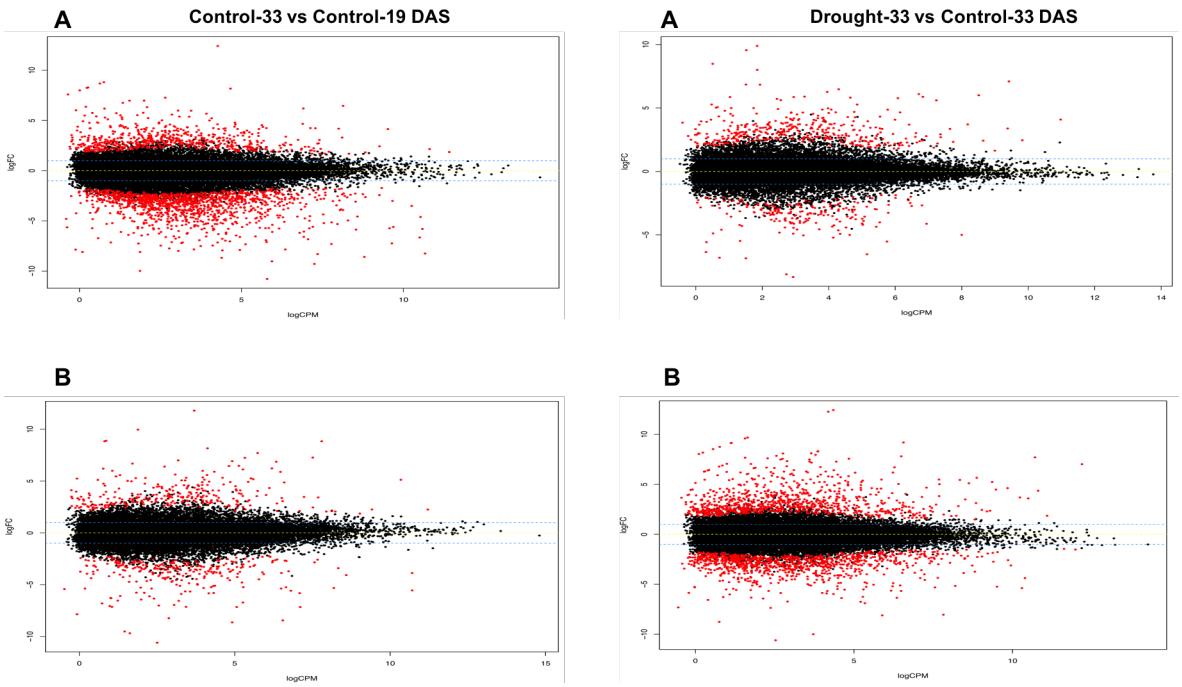

Figure 6.3: **MA plot comparing log FC and average log CPM on differential gene expression of Control-33 DAS vs Control-19 DAS and Drought-33 DAS vs Control-33 DAS at vegetative stage.** MA plot shows the log-fold change (i.e., the log ratio of normalized expression levels between two experimental conditions and against the log counts per million (CPM). [A] and [B] refers to plot from DipC and TN respectively. Experimental conditions of differential expression are listed on top of the figure. Those genes selected as differentially expressed (with FDR < 0.05, log fold change (logFC) ≥ 1) are highlighted as red dots.

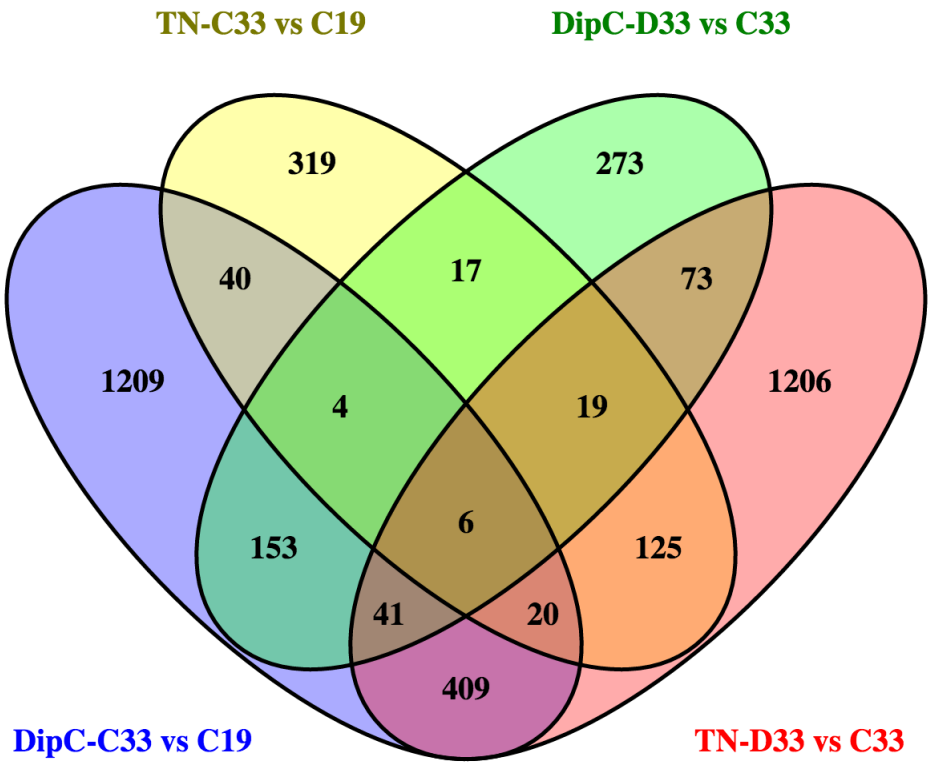

Figure 6.4: **Comparison of number of differentially expressed genes between DipC and TN at different experimental conditions.** A Venn diagram showing the number of differentially expressed genes in DipC (DipC-C33 vs C19 and DipC-D33 vs C33) and TN (TN-C33 vs C19 and TN-D33 vs C33) at different experimental conditions. C and D refer to control and drought respectively. 33 and 19 are the days after sowing at which the sampling was done. TN showed higher number of DE genes between TN-D33 vs C33 condition in comparison to DipC. Whereas DipC showed higher number of DE genes at DipC-C33 vs C19 condition.

Table 6.3: **Overlapping genes between DipC and TN in drought versus control (irrigated) condition at time point 33 DAS**

Gene-ID	Annotation	Regulation under drought in DipC	Regulation under drought in TN	Gene description in relation to drought stress
UoN.bamGnut.52130	HSP20	UP	UP	Induced by drought stress
UoN.bamGnut.45701	Unknown	UP	UP	
UoN.bamGnut.6545	Kunitz_legume	UP	UP	
UoN.bamGnut.52601	Peptidase_C1	UP	UP	
UoN.bamGnut.51451	Extensin_2	UP	UP	
UoN.bamGnut.2022	COX2_TM	UP	UP	
UoN.bamGnut.40032	PdomMRNAr1.1-08458.1	UP	UP	
UoN.bamGnut.38756	NA	UP	UP	
UoN.bamGnut.51739	BAG domain	UP	UP	Modulators of chaperone activity
UoN.bamGnut.43944	Raffinose_synthase	UP	UP	Osmoprotectant
UoN.bamGnut.48316	PQ-loop	UP	UP	
UoN.bamGnut.39903	Tryp_alpha_amyl	UP	UP	
UoN.bamGnut.43172	Lipoxygenase	UP	UP	Induced by drought stress
UoN.bamGnut.44527	Glyco_transf_5	UP	UP	
UoN.bamGnut.38939	Rhodanese	UP	UP	
UoN.bamGnut.47465	Betafructofuranosidase_14 (glycoside hydrolase_14)	UP	UP	Hydrolysis of sucrose- osmoprotectant
UoN.bamGnut.26149	Kunitz_legume	UP	UP	
UoN.bamGnut.45698	Acid Phosphatase B	UP	UP	Known to act under stress by maintaining a certain level of inorganic phosphate in plant cells
UoN.bamGnut.49985	Myo-inositol-1-phosphate synthase (MIPS)	UP	UP	Induced by ABA, drought
UoN.bamGnut.22131	Kunitz_legume	UP	UP	
UoN.bamGnut.18596	DOCK-C2	Down	UP	
UoN.bamGnut.46716	Kunitz_legume	UP	UP	
UoN.bamGnut.35993	RRM_7	Down	UP	
UoN.bamGnut.44278	DUF789	UP	UP	
UoN.bamGnut.46640	Glutaredoxin	UP	UP	Antioxidant/ROS scavenging
UoN.bamGnut.1420	ANTH	Down	UP	
UoN.bamGnut.8317	SAC3_GANP	UP	UP	
UoN.bamGnut.52158	GLE1	Down	UP	
UoN.bamGnut.27002	Carbonic anhydrase (Pro_CA)	UP	UP	Related to photosynthesis and plays role in ROS scavenging
UoN.bamGnut.39387	PLAT	UP	UP	
UoN.bamGnut.31006	Serine hydroxymethyltransferase (SHMT)	UP	UP	Related to oxidative stress. Functions in photorespiratory pathway and controlling the cell damage

UoN.bamGnut.8156	Lectin_legB	UP	UP	
UoN.bamGnut.51670	MreB_Mbl	UP	UP	
UoN.bamGnut.31696	HSP20	UP	UP	Induced by drought stress
UoN.bamGnut.41189	Kunitz_legume	UP	UP	
UoN.bamGnut.52453	Glutaredoxin	UP	UP	Antioxidant/ROS scavenging
UoN.bamGnut.40510	LEA_3	UP	UP	Induced by drought stress
UoN.bamGnut.54120	Glutathione S-transferase-2, N-terminal domain (GST)	Down	UP	Antioxidant/ ROS scavenging
UoN.bamGnut.4720	Lipoxygenase	UP	UP	Induced by drought stress
UoN.bamGnut.40504	WRI1(AP2-ERF)	UP	UP	Transcription
UoN.bamGnut.36919	Ferritin	UP	UP	Antioxidant/ ROS scavenging
UoN.bamGnut.32627	OPT	UP	UP	
UoN.bamGnut.42104	Unknown	Down	UP	
UoN.bamGnut.31521	Peptidase_C26	UP	UP	
UoN.bamGnut.2471	NA	UP	UP	
UoN.bamGnut.7659	Protein kinase	UP	UP	Signal transduction
UoN.bamGnut.23049	MIF4G	Down	UP	
UoN.bamGnut.27016	HSP20	UP	UP	Induced by drought stress
UoN.bamGnut.30393	ABC1	Down	UP	Cellular transporter
UoN.bamGnut.3344	p450	UP	UP	
UoN.bamGnut.39648	Extensin_1	UP	UP	
UoN.bamGnut.11038	mRNA_cap_C	UP	UP	
UoN.bamGnut.50762	PDDEXK_6	UP	UP	
UoN.bamGnut.27625	DnaJ	UP	UP	
UoN.bamGnut.54018	Aminotran_4	UP	UP	
UoN.bamGnut.20849	NA	UP	UP	
UoN.bamGnut.26151	S1-P1_nuclease	UP	UP	
UoN.bamGnut.41454	Hydrolase	UP	UP	
UoN.bamGnut.52505	PAR1	UP	UP	
UoN.bamGnut.47082	ECH_1	UP	UP	
UoN.bamGnut.40612	Lipase_3	UP	UP	
UoN.bamGnut.47740	Abhydrolase_1	UP	UP	
UoN.bamGnut.11968	Acid_phosphat_B	UP	UP	Known to act under stress by maintaining a certain level of inorganic phosphate in plant cells
UoN.bamGnut.47186	Unknown	UP	UP	
UoN.bamGnut.53110	Pollen_allerg_1	UP	UP	
UoN.bamGnut.23408	NUDIX	UP	UP	
UoN.bamGnut.43256	A6R773_AJECN	UP	UP	

UoN.bamGnut.47051	Acyltransferase	UP	UP	
UoN.bamGnut.35343	LRR_4	UP	UP	
UoN.bamGnut.40082	FAD_binding_3	UP	UP	
UoN.bamGnut.38941	NA	Down	UP	
UoN.bamGnut.20107	Pollen_allerg_1	UP	UP	
UoN.bamGnut.31054	EamA	UP	UP	
UoN.bamGnut.53840	ADH_zinc_N	UP	UP	
UoN.bamGnut.20540	Hydrolase_4	UP	UP	
UoN.bamGnut.20857	Pirin	UP	UP	
UoN.bamGnut.23286	Sulfate_transp	Down	Down	
UoN.bamGnut.17222	Light-harvesting complexes of green plants (Chloroa_b-bind)	UP	Down	Related to photosynthesis and plays a positive rolein guard cell signaling in response to ABA
UoN.bamGnut.998	Cupin_1	UP	Down	
UoN.bamGnut.21107	Aspartate protease (Asp)	UP	Down	Osmoprotectant
UoN.bamGnut.5351	Glyco_transf_8	UP	Down	
UoN.bamGnut.17370	ENSDNOP00000023418	UP	Down	
UoN.bamGnut.48758	Extensin_2	UP	Down	
UoN.bamGnut.31977	PC-Esterase	UP	Down	
UoN.bamGnut.33958	UBN2_3	UP	Down	
UoN.bamGnut.21376	Polyketide_cyc2	Down	Down	
UoN.bamGnut.21759	Cellulose_synthase	UP	Down	cell wall metabolism
UoN.bamGnut.25602	Tubulin	UP	Down	
UoN.bamGnut.3941	Lebercilin	UP	Down	
UoN.bamGnut.20825	GASA	UP	Down	
UoN.bamGnut.50179	ACYPI009798-PA	Down	Down	
UoN.bamGnut.23626	TPT	UP	Down	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.11622	HSP20	Down	Down	Induced by drought stress
UoN.bamGnut.26992	Drought induced 19 protein (Di19), zinc-binding	Down	Down	Transcription
UoN.bamGnut.15071	Polyketide_cyc2	Down	Down	
UoN.bamGnut.10340	DUF642	UP	Down	
UoN.bamGnut.21577	Cu-oxidase_2	UP	Down	copper ion binding. Redox related enzyme
UoN.bamGnut.5639	Peroxidase	UP	Down	Antioxidant/ ROS scavenging
UoN.bamGnut.3198	betafructofuranosidase_19 (glycoside hydrolase_19)	UP	Down	hydrolysis of sucrose- osmoprotectant
UoN.bamGnut.33001	Peptidase_C13	Down	Down	
UoN.bamGnut.10038	MYB_86	UP	Down	Transcription

UoN.bamGnut.22383	DPBB_1	UP	Down	
UoN.bamGnut.5245	DUF547	UP	Down	
UoN.bamGnut.42548	DEAD	UP	Down	
UoN.bamGnut.26120	EKC30732	UP	Down	
UoN.bamGnut.8560	NA	Down	Down	
UoN.bamGnut.17657	Methyltransf_3	UP	Down	
UoN.bamGnut.16499	Adenine_glyco	UP	Down	
UoN.bamGnut.16317	Aminotran_3	UP	Down	
UoN.bamGnut.45815	DOMON	UP	Down	
UoN.bamGnut.4986	Fasciclin	UP	Down	Related to cell wall
UoN.bamGnut.35132	Methyltransferase_11	UP	Down	
UoN.bamGnut.34370	PPR_3	Down	Down	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.26937	Weak chloroplast movement under blue light (WEMBL)	Down	Down	Photosynthesis related gene
UoN.bamGnut.21800	LRR_4	UP	Down	
UoN.bamGnut.11322	NA	UP	Down	
UoN.bamGnut.31476	SEO_N	UP	Down	
UoN.bamGnut.49481	WD40	UP	Down	
UoN.bamGnut.9228	zf-RING_2	UP	Down	Transcription
UoN.bamGnut.5468	Fasciclin	UP	Down	Related to cell wall
UoN.bamGnut.11930	Peptidase_M10	UP	Down	
UoN.bamGnut.44510	WD40	UP	Down	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.8660	Ap2-ERF domain	Down	Down	Transcription
UoN.bamGnut.11591	FAD_binding_4	Down	Down	
UoN.bamGnut.12631	MORN	Down	Down	
UoN.bamGnut.22613	B3	Down	Down	
UoN.bamGnut.47303	LEA_2	Down	Down	Induced by drought stress
UoN.bamGnut.1776	AAA_29	UP	Down	
UoN.bamGnut.33047	ABC_transport	Down	Down	Cellular transporter
UoN.bamGnut.41899	Proton_antipo_M	Down	Down	
UoN.bamGnut.9081	GMC_oxred_C	UP	Down	
UoN.bamGnut.1472	Amidase	Down	Down	
UoN.bamGnut.6748	Ferric_reduct	Down	Down	
UoN.bamGnut.48904	Glycos_transf_1	UP	Down	
UoN.bamGnut.17348	RAD-like 6 (MYB TF)	Down	Down	Transcription
UoN.bamGnut.13977	AGLA009383-RA	UP	Down	
UoN.bamGnut.18415	TPR_12	UP	Down	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.17350	Unknown	Down	Down	
UoN.bamGnut.21321	ParA	UP	Down	

Table 6.4: **Top UP-regulated genes in drought compared to control condition at time point 33 DAS in DipC and TN.**

Gene_ID	logFC	FDR	Annotation	Gene description in relation to drought stress
UP-regulated in DipC				
UoN.bamGnut.44662	9.89	0.00000	DEAD BOX Helicase_C	RNA metabolism, response to ABA, drought, signal transduction
UoN.bamGnut.52258	9.55	0.00000	HEAT_2 (ARM-repeat)	Protein sorting via ubiquitination
UoN.bamGnut.34377	8.48	0.00000	DUF21	
UoN.bamGnut.44278	8.00	0.00000	DUF789	
UoN.bamGnut.21369	7.09	0.00001	Kunitz_legume	
UoN.bamGnut.15773	6.84	0.00000	Alternative oxidase (AOX)	Maintains photosynthesis during drought by promoting energy balance in the chloroplast.
UoN.bamGnut.9081	6.83	0.00000	Glucose-methanol-choline oxidoreductase, C	Glycolysis/carbohydrate metabolism
UoN.bamGnut.52291	6.47	0.00000	Hemopexin	
UoN.bamGnut.15899	6.26	0.00000	B3	
UoN.bamGnut.26149	6.09	0.00000	Kunitz_legume	
UoN.bamGnut.49985	6.00	0.00000	Myo-inositol-1-phosphate synthase (MIPS)	Induced by ABA, drought
UoN.bamGnut.19390	5.90	0.00001	Iron transport multicopper oxidase fio1	Iron transport
UoN.bamGnut.44892	5.88	0.00057	Translationally controlled tumor protein (TCTP)	Stomatal closure via ABA signaling
UoN.bamGnut.892	5.86	0.00000	Calcium-dependent protein kinase (CDPK) family protein	Signal transduction. Plant drought stress signaling
UoN.bamGnut.38284	5.86	0.00000	Na+/H+ exchangers (NHX)	Sodium ion transport and potassium ion homeostasis
UoN.bamGnut.24223	5.82	0.00015	Adenine Phosphoribosyl-Transferase (APT1)	
UoN.bamGnut.2022	5.77	0.00086	COX2_TM	
UoN.bamGnut.12534	5.72	0.00000	Fibronectin type III	
UoN.bamGnut.16854	5.61	0.00000	Protein kinase	Signal transduction
UoN.bamGnut.21321	5.60	0.00606	ParA (ATPase)	
UoN.bamGnut.22864	5.27	0.00000	Amidase	
UoN.bamGnut.44527	5.27	0.00189	Glyco_transf_5	Maintaining cell homeostasis and regulating plant growth and development. Regulation of lignin content.
UoN.bamGnut.25928	5.07	0.00000	Prenyltransferase	
UoN.bamGnut.43162	5.02	0.00000	EMB1444 (bHLH)	Transcription
UoN.bamGnut.35343	4.99	0.00000	LRR4	Protein phosphorylation, regulates drought stress by activating the expression of ABA

UoN.bamGnut.48636	4.99	0.00000	PTS_2-RNA	
UoN.bamGnut.29765	4.97	0.00358	PP2C	Accelerates growth of inflorescence stems through the activation of cell proliferation and expansion.
UoN.bamGnut.23436	4.92	0.00001	Putative dentin sialophosphoprotein-related	
UoN.bamGnut.19461	4.88	0.00003	SacI homology domain	Cell wall maintanence
UoN.bamGnut.12158	4.87	0.00000	Brr2p DEAD/DExH box ATP-dependent RNA helicase	
UoN.bamGnut.27002	4.86	0.00591	Carbonic anhydrase (Pro_CA)	Related to photosynthesis and plays role in ROS scavenging
UoN.bamGnut.36776	4.78	0.00000	Arp2/3 complex,	
UoN.bamGnut.18415	4.69	0.00211	TPR_12	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.18429	4.66	0.00911	HEAT_2 (ARM-repeat)	Protein sorting via ubiquitination
UoN.bamGnut.17222	4.65	0.00249	Light-harvesting complexes of green plants (Chloroa_b-bind)	Related to photosynthesis and plays a positive rolein guard cell signaling in response to ABA
UoN.bamGnut.9444	4.65	0.00238	Unknown	
UoN.bamGnut.41454	4.64	0.00009	Hydrolase	
UoN.bamGnut.47930	4.62	0.00001	Nasvi2EG004087t1	
UoN.bamGnut.549	4.60	0.00037	LRR_4	Protein phosphorylation, regulates drought stress by activating the expression of ABA
UoN.bamGnut.54073	4.55	0.00018	Nodulin-like	
UoN.bamGnut.16251	4.54	0.00209	Sister chromatid cohesion C-terminus	
UoN.bamGnut.19086	4.53	0.00000	K_transport	Osmotically regulates water movement in and outside of cells . It plays a major role in guard cell regulation under drought stress
UoN.bamGnut.5098	4.49	0.00014	Pyruvate kinase_C, alpha/beta domain	Enzyme involved in glycolysis.
UoN.bamGnut.24293	4.39	0.00002	Lecithin retinol acyltransferase	
UoN.bamGnut.3802	4.39	0.00004	Triose-phosphate	Transport activity
UoN.bamGnut.15855	4.33	0.00001	Hydrolase_4	
UoN.bamGnut.25086	4.32	0.00003	MarR	
UoN.bamGnut.18603	4.29	0.00947	Ribosomal_L32p	
UoN.bamGnut.52070	4.27	0.00000	Root phototropism protein 3 (NPH3)	Response to light stimulus, gravitropism
UoN.bamGnut.19291	4.26	0.00004	Multipcopper oxidase	Copper ion binding. Redox related enzyme
UoN.bamGnut.49481	4.23	0.00649	WD40	Repeats families in plants allow them to rapidly acclimatize to adverse conditions

UoN.bamGnut.54221	4.21	0.00201	B3	
UoN.bamGnut.4725	4.19	0.01223	LRR_4	
UoN.bamGnut.11370	4.18	0.00606	HAD	
UoN.bamGnut.51451	4.15	0.00154	Extensin_2	
UoN.bamGnut.31006	4.13	0.02744	Serine hydroxymethyltransferase	Related to oxidative stress. Functions in photorespiratory pathway and controlling the cell damage
UoN.bamGnut.49011	4.09	0.00300	Queuine tRNA-ribosyltransferase	
UoN.bamGnut.45698	4.09	0.00000	Acid Phosphatase B	Acid phosphatase is known to act under stress by maintaining a certain level of inorganic phosphate in plant cells
UoN.bamGnut.6545	4.08	0.00012	Kunitz_legume	
UoN.bamGnut.15578	4.08	0.00013	DUF247	
UoN.bamGnut.21108	4.07	0.00015	Zinc-binding dehydrogenase	
UoN.bamGnut.7185	4.07	0.02506	Trehalose-phosphatase	Osmoprotectant. Starch and sucrose metabolism.
UoN.bamGnut.36312	4.06	0.00367	Replication factor RFC1 C terminal domain	
UoN.bamGnut.46713	4.04	0.00006	GDP-fucose protein O-fucosyltransferase (O-Fuct)	Glycolysis/carbohydrate metabolism
UoN.bamGnut.3717	4.02	0.00026	Probable lipid transfer 2	
UoN.bamGnut.44600	3.99	0.00000	NB-ARC	
UoN.bamGnut.48066	3.98	0.00015	BRCT	
UoN.bamGnut.23408	3.96	0.00606	NUDIX	
UoN.bamGnut.26564	3.96	0.00201	Glucose-6-phosphate dehydrogenase (G6PD_C)	Plays role in glycolysis to produce energy in the form of ATP and NADH
UoN.bamGnut.45701	3.95	0.00003	NA	
UoN.bamGnut.25328	3.94	0.00140	NAC	Transcription
UoN.bamGnut.23078	3.91	0.00002	Histidine phosphatase 1	
UoN.bamGnut.43172	3.89	0.00020	Lipoxygenase 1	Induced by drought stress
UoN.bamGnut.53922	3.87	0.00104	DnaJ	
UoN.bamGnut.25503	3.86	0.00004	EDAN000067-RA	
UoN.bamGnut.49066	3.84	0.00018	Protein kinase	Signal transduction
UoN.bamGnut.23090	3.84	0.03035	Amino_oxidase	
UoN.bamGnut.20798	3.84	0.02517	Peroxidase	Antioxidant/ROS scavenging. Conservation and protection of cellular constituents
UoN.bamGnut.35465	3.83	0.00014	LTP_2	
UoN.bamGnut.48316	3.81	0.00128	PQ-loop	
UoN.bamGnut.19428	3.77	0.00098	RIC1	

UoN.bamGnut.23583	3.76	0.04093	NA	
UoN.bamGnut.27164	3.76	0.00284	Plant_trans	Cellular transporter
UoN.bamGnut.47895	3.75	0.00030	NA	
UoN.bamGnut.4614	3.74	0.00057	GDP-fucose protein O-fucosyltransferase (O-Fuct)	Glycolysis/carbohydrate metabolism
UoN.bamGnut.23947	3.73	0.00038	MYB_6	Transcription
UoN.bamGnut.3344	3.72	0.00026	p450	
UoN.bamGnut.47465	3.71	0.00047	Beta-fructofuranosidase 14 (glycoside hydrolase 14)	Hydrolysis of sucrose- Osmoprotectant
UoN.bamGnut.8981	3.70	0.00144	DnaJ	
UoN.bamGnut.44510	3.70	0.01943	WD40	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.46772	3.70	0.00079	Peroxidase	
UoN.bamGnut.37637	3.70	0.02757	Na+/H+ exchangers (NHX)	Sodium ion transport and potassium ion homeostasis
UoN.bamGnut.22131	3.69	0.00013	Kunitz_legume	
UoN.bamGnut.14864	3.69	0.02910	Tetraspannin	
UoN.bamGnut.24283	3.69	0.00000	basic helix-loop-helix (bHLH)	Transcription
UoN.bamGnut.37408	3.69	0.00014	DUF179	
UoN.bamGnut.22716	3.69	0.04368	Agenet	
UoN.bamGnut.4649	3.67	0.00719	NAC	Transcription
UoN.bamGnut.24715	3.66	0.00494	Unknown	
UP-regulated in TN				
UoN.bamGnut.682	12.44	0.00000	Glutathione S-transferase, C-terminal domain (GST)	Antioxidant/ ROS scavenging
UoN.bamGnut.39360	12.28	0.00000	Unknown	
UoN.bamGnut.53377	9.69	0.00000	zf-C3HC4_3	Transcription
UoN.bamGnut.52163	9.60	0.00000	Thaumatin	
UoN.bamGnut.52130	9.20	0.00000	HSP20	Induced by drought stress
UoN.bamGnut.38233	9.15	0.00000	Unknown	
UoN.bamGnut.15605	9.13	0.00000	Protein kinase domian	Signal transduction
UoN.bamGnut.45701	8.76	0.00000	Unknown	
UoN.bamGnut.53091	8.53	0.00000	ABC_transport	Cellular transporter
UoN.bamGnut.51344	8.32	0.00000	HelroP129652	
UoN.bamGnut.28590	8.30	0.00000	Extensin_2	
UoN.bamGnut.11823	8.18	0.00000	UBA	
UoN.bamGnut.48031	8.11	0.00000	Dehydrin/LEA	Induced by drought stress
UoN.bamGnut.46852	8.09	0.00000	Bet_v_1	

UoN.bamGnut.52834	8.04	0.00000	MYB_6	Transcription
UoN.bamGnut.50795	7.96	0.00000	D-mannose binding lectin	Immune system
UoN.bamGnut.6648	7.86	0.00000	Unknown	
UoN.bamGnut.29480	7.81	0.00000	Succinyl-CoA ligase like flavodoxin domain	
UoN.bamGnut.6545	7.71	0.00000	Kunitz_legume	
UoN.bamGnut.20230	7.67	0.00000	Protein kinase domian	Signal transduction
UoN.bamGnut.42913	7.61	0.00000	Protein kinase domian	Signal transduction
UoN.bamGnut.52601	7.57	0.00000	Peptidase_C1	
UoN.bamGnut.49070	7.55	0.00000	ENSPVAP00000008033	
UoN.bamGnut.44087	7.49	0.00000	The pentatricopeptide repeat_1 (PPR)	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.41260	7.49	0.00000	Transferase	
UoN.bamGnut.37116	7.33	0.00000	Zinc finger AN1	Transcription
UoN.bamGnut.40803	7.28	0.00000	GED	
UoN.bamGnut.42037	7.27	0.00000	LEA_1	Induced by drought stress
UoN.bamGnut.7677	7.18	0.00000	ENSMODP00000036472	
UoN.bamGnut.50258	7.06	0.00000	mir-399	
UoN.bamGnut.51719	7.05	0.00000	Probable lipid transfer 2	
UoN.bamGnut.7343	7.02	0.00000	Acid phosphatase B	Acid phosphatase is known to act under stress by maintaining a certain level of inorganic phosphate in plant cells
UoN.bamGnut.51497	6.89	0.00000	Unknown	
UoN.bamGnut.51856	6.81	0.00000	mir-399	
UoN.bamGnut.51461	6.78	0.00000	UBN2_2	
UoN.bamGnut.43813	6.74	0.00000	HSP20	Induced by drought stress
UoN.bamGnut.52142	6.71	0.00000	Unknown	
UoN.bamGnut.49776	6.68	0.00000	Calcium-dependent protein kinase (CDPK) family protein	Signal transduction. Plant drought stress signaling
UoN.bamGnut.39866	6.67	0.00000	Unknown	
UoN.bamGnut.51451	6.64	0.00000	Extensin_2	
UoN.bamGnut.35142	6.62	0.00000	Nup188	
UoN.bamGnut.15016	6.45	0.00000	Tubulin-tyrosine ligase family	
UoN.bamGnut.45667	6.43	0.00000	ABC2_membrane	Cellular transporter
UoN.bamGnut.16477	6.42	0.00000	MIF4G	
UoN.bamGnut.51111	6.40	0.00000	Dirigent	
UoN.bamGnut.54261	6.30	0.00000	Unknown	
UoN.bamGnut.42294	6.29	0.00000	Unknown	
UoN.bamGnut.54259	6.28	0.00000	Beta-fructofuranosidase 31 (glycoside hydrolase 31)	Hydrolysis of sucrose - Osmoprotectant

UoN.bamGnut.51926	6.28	0.00000	DUF3774	
UoN.bamGnut.53190	6.27	0.00000	TCOGS2:TC014111-PA	
UoN.bamGnut.41882	6.25	0.00000	RING-H2 zinc finger domain	
UoN.bamGnut.53678	6.24	0.00000	RVT_2	
UoN.bamGnut.50295	6.20	0.00000	SPX	
UoN.bamGnut.40500	6.20	0.00000	AAA_16	
UoN.bamGnut.2022	6.19	0.00010	Cytochrome c oxidase subunit 2	Role in cell apoptosis
UoN.bamGnut.38337	6.17	0.00000	LEA_4	Induced by drought stress
UoN.bamGnut.20704	6.09	0.00000	PCI	
UoN.bamGnut.38394	6.07	0.00000	Sulfate_transport	Cellular transporter
UoN.bamGnut.53482	6.06	0.00000	G3ATJ1_SPAPN	
UoN.bamGnut.2423	6.02	0.00000	Nucleoporin autopeptidase_2	
UoN.bamGnut.42356	6.01	0.00000	LRR_4	
UoN.bamGnut.40032	5.98	0.00000	PdomMRNAr1.1-08458.1	
UoN.bamGnut.53106	5.98	0.00000	Acyl-CoA dehydrogenase_1, C-terminal domain	
UoN.bamGnut.51210	5.92	0.00000	Beta-fructofuranosidase 18 (glycoside hydrolase 18)	Hydrolysis of sucrose - Osmoprotectant
UoN.bamGnut.38756	5.91	0.00000	Unknown	
UoN.bamGnut.52823	5.87	0.00000	CBG23024	
UoN.bamGnut.8123	5.85	0.00000	S2J9K3_MUCC1	
UoN.bamGnut.13863	5.82	0.00000	PI3_PI4_kinase	Signal transduction
UoN.bamGnut.16103	5.80	0.00016	GBA2_N	
UoN.bamGnut.46488	5.80	0.00006	NADH dehydrogenase (ubiquinone)	Establishment of the electrochemical potential used to produce ATP transporters for sodium and pH homeostasis.
UoN.bamGnut.52162	5.77	0.00000	S1-P1_nuclease	
UoN.bamGnut.46252	5.74	0.00000	HSP20	Induced by drought stress
UoN.bamGnut.39120	5.72	0.00000	EST1_DNA_bind	
UoN.bamGnut.51739	5.66	0.00000	BAG domain	Modulators of Chaperone activity
UoN.bamGnut.11638	5.66	0.00000	Hydrolase_4	
UoN.bamGnut.52447	5.60	0.00000	Unknown	
UoN.bamGnut.52859	5.60	0.00001	HAD	
UoN.bamGnut.46816	5.56	0.00000	LRR_8	
UoN.bamGnut.41199	5.54	0.00000	Cytochrome P450	Role in phytohormone biosynthesis and catabolism, and secondary metabolism
UoN.bamGnut.1452	5.52	0.00000	Terpene synthase C terminal domain	Its role is to synthesize terpenes
UoN.bamGnut.27805	5.50	0.00001	GYF	
UoN.bamGnut.43944	5.49	0.00000	Raffinose_synthase	Osmoprotectant

UoN.bamGnut.39428	5.48	0.00000	Calmodulin_binding	Signal transduction
UoN.bamGnut.21817	5.48	0.00000	Ceramidase	
UoN.bamGnut.45742	5.47	0.00000	Sec3_C	
UoN.bamGnut.13020	5.45	0.00034	AAA_29	
UoN.bamGnut.48316	5.43	0.00000	PQ-loop	
UoN.bamGnut.23323	5.43	0.00000	Cpn60_TCP1	
UoN.bamGnut.43836	5.38	0.00000	FBpp0160707	
UoN.bamGnut.46119	5.36	0.00000	J9VLD2_CRYNH	
UoN.bamGnut.38252	5.31	0.00000	Methyltransf_16	
UoN.bamGnut.42115	5.28	0.00000	Multi_copper_oxidase_3	copper ion binding. Redox related enzyme
UoN.bamGnut.52261	5.26	0.00000	Unknown	
UoN.bamGnut.51074	5.24	0.00000	LEA_3	Induced by drought stress
UoN.bamGnut.39903	5.24	0.00000	Plant lipid transfer protein, (LTP family)	Responsible for the shuttling of phospholipids and other fatty acid groups between cell membranes
UoN.bamGnut.43172	5.23	0.00000	Lipoxygenase	Induced by drought stress
UoN.bamGnut.45350	5.23	0.00011	Lyase_1	
UoN.bamGnut.38683	5.23	0.00000	Methyltransf_2	

Table 6.5: **Top down-regulated genes in drought compared to control condition at time point 33 DAS in DipC and TN**

Gene_ID	logFC	FDR	Annotation	Gene description in relation to drought stress
Down-regulated in DipC				
UoN.bamGnut.50173	-8.32	0.00000	Unknown	
UoN.bamGnut.8282	-8.10	0.00000	EamA	Transporter
UoN.bamGnut.17677	-6.84	0.00000	M3BXM6_SPHMS	
UoN.bamGnut.33450	-6.80	0.00000	Protein kinase_Tyrosine	Signal transduction. Related to cell division
UoN.bamGnut.6437	-6.53	0.00000	PsbP	Photosynthesis related gene
UoN.bamGnut.19320	-6.35	0.00000	REF	
UoN.bamGnut.32908	-5.57	0.00000	Na+/H+ exchangers (NHX)	Sodium ion transport and potassium ion homeostasis
UoN.bamGnut.44453	-5.52	0.00000	bHLH-MYC_N	Transcription
UoN.bamGnut.28497	-5.47	0.00000	Unknown	
UoN.bamGnut.18186	-5.21	0.00000	tRNA_lig_CPD	
UoN.bamGnut.42028	-5.04	0.00011	Unknown	
UoN.bamGnut.53530	-5.00	0.00014	Phosphoenolpyruvate carboxykinase	Photosynthesis related gene
UoN.bamGnut.26937	-5.00	0.00048	Weak chloroplast movement under blue light (WEMBL)	Photosynthesis related gene
UoN.bamGnut.40700	-4.92	0.00926	Ank_2	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.14721	-4.90	0.01414	Protein kinase_Tyrosine	Signal transduction. Related to cell division
UoN.bamGnut.5763	-4.88	0.00000	Ion_transport	Cellular transporter
UoN.bamGnut.50179	-4.82	0.00008	ACYPI009798-PA	
UoN.bamGnut.54120	-4.64	0.00026	Glutathione S-transferase, N-terminal domain (GST)	Antioxidant/ ROS scavenging
UoN.bamGnut.17348	-4.59	0.00000	RAD-like 6 (MYB TF)	Transcription
UoN.bamGnut.47383	-4.58	0.00154	DUF760	
UoN.bamGnut.46909	-4.56	0.00358	Suf	
UoN.bamGnut.29455	-4.45	0.00027	G9NYD7_HYPAI	
UoN.bamGnut.47712	-4.42	0.00649	DUF3591	
UoN.bamGnut.6805	-4.42	0.00003	Sin3a_C	
UoN.bamGnut.2247	-4.36	0.00028	Protein kinase	Signal transduction
UoN.bamGnut.12996	-4.34	0.00000	Cytokin-binding	Cell division, hormone signaling
UoN.bamGnut.4723	-4.30	0.00008	Tau95	
UoN.bamGnut.23848	-4.25	0.00492	Aspartate/ornithine carbamoyltransferase	
UoN.bamGnut.27114	-4.22	0.00010	Methyltransf_11	
UoN.bamGnut.1472	-4.21	0.00170	Amidase	

UoN.bamGnut.51799	-4.12	0.00280	Proton_antipo_M	
UoN.bamGnut.42342	-4.12	0.00274	DUF1666	
UoN.bamGnut.38518	-4.10	0.00412	RNA_pol_L_2	
UoN.bamGnut.33334	-4.10	0.01790	Unknown	
UoN.bamGnut.36694	-4.06	0.00003	SPU_008979-tr	
UoN.bamGnut.23049	-4.05	0.00186	MIF4G	
UoN.bamGnut.27231	-4.05	0.00036	Unknown	
UoN.bamGnut.18596	-4.03	0.01769	DOCK-C2	
UoN.bamGnut.48225	-4.01	0.00000	Hydrolase_3	
UoN.bamGnut.49353	-3.98	0.02381	Peptidase_S9_N	
UoN.bamGnut.38457	-3.95	0.00616	LRR_8	
UoN.bamGnut.16381	-3.94	0.00149	TPR_11	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.51068	-3.91	0.01176	EFX63927	
UoN.bamGnut.14565	-3.91	0.00015	RVT_1	
UoN.bamGnut.36349	-3.91	0.00050	Methyltransferase_29	
UoN.bamGnut.18181	-3.90	0.00014	DUF936	
UoN.bamGnut.42104	-3.88	0.01189	Unknown	
UoN.bamGnut.32557	-3.85	0.00829	SAP	
UoN.bamGnut.51042	-3.80	0.03490	p450	
UoN.bamGnut.1716	-3.80	0.00040	PP2C	Negative regulator of ABA signaling
UoN.bamGnut.18659	-3.78	0.00154	Hydrolase	
UoN.bamGnut.2113	-3.76	0.00463	Sin3a_C	
UoN.bamGnut.12631	-3.74	0.01462	MORN	
UoN.bamGnut.33014	-3.74	0.02181	LBP_BPI_CETP	
UoN.bamGnut.23721	-3.73	0.01972	DUF3506	
UoN.bamGnut.52560	-3.72	0.00227	Voltage_CLC	
UoN.bamGnut.33047	-3.70	0.00907	ABC_transport	Cellular transporter
UoN.bamGnut.48794	-3.70	0.00042	Xyloglucan endo-transglycosylase_C (XET)	With cellulose, it forms a network that strengthens the cell wall.
UoN.bamGnut.42484	-3.69	0.00130	XPG_N	
UoN.bamGnut.30940	-3.69	0.00975	ENSRNOP00000040000	
UoN.bamGnut.3638	-3.60	0.00001	Methyltransferase_11	
UoN.bamGnut.7974	-3.57	0.00092	Unknown	
UoN.bamGnut.17350	-3.56	0.00001	Unknown	
UoN.bamGnut.10898	-3.56	0.00001	Tubulin	
UoN.bamGnut.20899	-3.55	0.00534	DnaJ	
UoN.bamGnut.10535	-3.52	0.00465	Retroviral aspartyl protease_2 (RVP_2)	

UoN.bamGnut.28528	-3.49	0.04192	EamA	Transporter
UoN.bamGnut.32211	-3.46	0.04794	RRM_1	
UoN.bamGnut.41899	-3.46	0.02991	Proton_antipo_M	
UoN.bamGnut.11591	-3.44	0.02363	FAD_binding_4	
UoN.bamGnut.10673	-3.43	0.02427	Thioredoxin_6	Antioxidant/ ROS scavenging
UoN.bamGnut.31586	-3.42	0.00593	EF-hand_5	
UoN.bamGnut.17876	-3.42	0.00513	Oxidored_FMN	
UoN.bamGnut.35205	-3.42	0.00753	Unknown	
UoN.bamGnut.25716	-3.40	0.00641	DUF547	
UoN.bamGnut.38410	-3.40	0.03021	Calmodulin_binding	Signal transduction
UoN.bamGnut.37267	-3.40	0.03664	Protein kinase_Tyrosine	Signal transduction. Related to cell division.
UoN.bamGnut.46804	-3.37	0.02440	zf-CCCH	Transcription
UoN.bamGnut.4995	-3.36	0.00097	PPR_2	
UoN.bamGnut.33972	-3.34	0.00886	ENSONIP00000017366	
UoN.bamGnut.50657	-3.34	0.02744	Beta-fructofuranosidase 100 (glycoside hydrolase 100)	
UoN.bamGnut.8660	-3.33	0.00077	AP2-ERF domain	Transcription
UoN.bamGnut.34370	-3.31	0.00959	PPR_3	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.14537	-3.30	0.04153	Unknown	
UoN.bamGnut.17182	-3.30	0.02317	F-box-like	
UoN.bamGnut.52158	-3.29	0.01774	GLE1	
UoN.bamGnut.46666	-3.29	0.02849	HEAT_2	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.21841	-3.28	0.01086	RVT_1	
UoN.bamGnut.27710	-3.28	0.01701	Voltage_CLC	
UoN.bamGnut.15589	-3.28	0.03396	Beta-fructofuranosidase 17 (glycoside hydrolase 17)	Hydrolysis of sucrose- osmoprotectant
UoN.bamGnut.53896	-3.26	0.01545	Peptidase_M28	
UoN.bamGnut.49601	-3.25	0.03785	Abhydrolase_2	
UoN.bamGnut.26712	-3.24	0.03709	RBM39linker	
UoN.bamGnut.32688	-3.23	0.04707	Glyco_transf_8	
UoN.bamGnut.28815	-3.21	0.00662	Phosphoribulokinase (PRK)	Enzyme involved in calvin cycle- related to photosynthesis
UoN.bamGnut.31253	-3.21	0.02865	Tryp_alpha_amyl	
UoN.bamGnut.26992	-3.20	0.00794	Drought induced 19 protein (Di19), zinc-binding	Transcription
UoN.bamGnut.35210	-3.20	0.00236	Lipase_3	
UoN.bamGnut.46473	-3.20	0.01523	PDT	
UoN.bamGnut.17086	-3.19	0.04836	Cation_ATPase_N	Cellular transporter
UoN.bamGnut.39874	-3.18	0.03865	Nexin_C	

Down-regulated in TN				
UoN.bamGnut.21994	-10.60	0.00000	NT-C2	
UoN.bamGnut.8669	-10.00	0.00000	PMR5N	
UoN.bamGnut.28701	-8.76	0.00000	5_nucleotidide	
UoN.bamGnut.25669	-8.11	0.00000	Aldose_epim	
UoN.bamGnut.21321	-8.05	0.00000	ParA	
UoN.bamGnut.35437	-7.36	0.00000	DLH	
UoN.bamGnut.22562	-7.32	0.00000	Phosphoenolpyruvate carboxyki-nase_PEPCK_ATP	Photosynthesis related gene
UoN.bamGnut.23494	-6.89	0.00000	Fasciclin	Related to cell wall
UoN.bamGnut.34339	-6.75	0.00000	Med23	
UoN.bamGnut.21901	-6.72	0.00000	NPH3_ nonphototropic hypocotyl 1	Signal transduction. Related to phototropism - a light-activated serine-threonine protein kinase.
UoN.bamGnut.50515	-6.56	0.00000	Unknown	
UoN.bamGnut.15431	-6.43	0.00000	Aluminium activated malate transporter (ALMT)	Cellular transporter
UoN.bamGnut.30413	-6.38	0.00098	CRAL_TRIO	
UoN.bamGnut.44444	-6.35	0.00015	DnaJ_C	
UoN.bamGnut.16147	-6.33	0.00000	Hpt	
UoN.bamGnut.30416	-6.28	0.00000	Chorismate_binding	
UoN.bamGnut.4610	-6.25	0.00001	Haloacid dehalogenase-like hydrolase_2	Osmoprotectant
UoN.bamGnut.41632	-6.02	0.00000	Peroxidase	ROS scavenging/antioxidant
UoN.bamGnut.34826	-5.97	0.00015	Mito_carrier	
UoN.bamGnut.8986	-5.91	0.00000	PB10360-PA	
UoN.bamGnut.15067	-5.87	0.00000	Unknown	
UoN.bamGnut.52668	-5.86	0.00002	ATP-synt_ab_N	Energy metabolism
UoN.bamGnut.17351	-5.86	0.00000	Unknown	
UoN.bamGnut.9935	-5.67	0.00046	D8QDI4_SCHCM	
UoN.bamGnut.26296	-5.67	0.00015	Haloacid dehalogenase, N terminal	
UoN.bamGnut.29983	-5.67	0.00004	Sulfate_transport	Cellular transporter
UoN.bamGnut.17350	-5.64	0.00000	Unknown	
UoN.bamGnut.31174	-5.57	0.00000	VARLMGL	
UoN.bamGnut.32104	-5.57	0.00000	NA	
UoN.bamGnut.18415	-5.54	0.00001	Tetratricopeptide repeat_12 (TPR)	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.16443	-5.52	0.00000	Fasciclin	Related to cell wall
UoN.bamGnut.9082	-5.51	0.00000	Glucose-methanol-choline oxidoreductase family	
UoN.bamGnut.25416	-5.47	0.00000	SHN3 (AP2-ERF domain)	Transcription

UoN.bamGnut.1002	-5.45	0.00000	POX	
UoN.bamGnut.13977	-5.38	0.00000	AGLA009383-RA	
UoN.bamGnut.22485	-5.31	0.00001	Glucuronosyltransferase (UDPGT)	Involved in the formation of β- glucans which is involved in formation of cell wall
UoN.bamGnut.13966	-5.30	0.00000	MIP (aquaporin)	Water transport
UoN.bamGnut.9295	-5.26	0.00000	Calcium-dependent protein kinase (CDPK) family protein	Signal transduction. Plant drought stress signaling
UoN.bamGnut.31955	-5.25	0.00000	ABC_transport	Cellular transporter
UoN.bamGnut.24640	-5.24	0.00000	Tryp_alpha_amyl	
UoN.bamGnut.17348	-5.23	0.00000	RAD-like 6 (MYB TF)	Transcription
UoN.bamGnut.2072	-5.21	0.00000	Unknown	
UoN.bamGnut.8752	-5.20	0.00000	Myb_DNA-binding	Transcription
UoN.bamGnut.50781	-5.20	0.00000	PAM2	
UoN.bamGnut.33191	-5.16	0.00000	AECH14963-PA	
UoN.bamGnut.19806	-5.12	0.00001	Unknown	
UoN.bamGnut.2075	-5.12	0.00000	Unknown	
UoN.bamGnut.48904	-5.11	0.00010	Glycos_transf_1	
UoN.bamGnut.25812	-5.07	0.00001	Zip	
UoN.bamGnut.44681	-5.04	0.00029	COX2_TM	
UoN.bamGnut.22141	-5.02	0.00000	S3BYK0_OPHP1	
UoN.bamGnut.24800	-4.98	0.00000	bHLH-MYC_N	Transcription
UoN.bamGnut.6748	-4.89	0.00000	Ferric_reductase	
UoN.bamGnut.47216	-4.88	0.00004	Methyltransferase_3	
UoN.bamGnut.32962	-4.88	0.00120	EDAN002454-RA	
UoN.bamGnut.25161	-4.88	0.00000	Bm8519b	
UoN.bamGnut.3509	-4.88	0.00002	Multipcopper oxidase	copper ion binding. Redox related enzyme
UoN.bamGnut.25093	-4.86	0.00024	Adenine_glyco	
UoN.bamGnut.25315	-4.86	0.00000	LFUL016095-RA	
UoN.bamGnut.31347	-4.84	0.00000	FAD_binding_7	
UoN.bamGnut.29652	-4.83	0.00011	WD40	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.1472	-4.82	0.00222	Amidase	
UoN.bamGnut.26827	-4.80	0.00000	Pkinase_Tyrosine	Signal transduction. Related to cell division
UoN.bamGnut.10193	-4.80	0.00004	TPX2	
UoN.bamGnut.13811	-4.79	0.00001	Betafructofuranosidase_18 (glycoside hydrolase_18)	hydrolysis of sucrose- osmoprotectant
UoN.bamGnut.8813	-4.76	0.00001	DUF239	
UoN.bamGnut.12329	-4.74	0.00214	RSN1_7TM	
UoN.bamGnut.9452	-4.71	0.00001	NA	
UoN.bamGnut.8862	-4.69	0.00001	TB2_DP1_HVA22	
UoN.bamGnut.25659	-4.68	0.00002	Cytochrom_B561	
UoN.bamGnut.20995	-4.66	0.00000	RAD-like 6 (MYB TF)	Transcription

UoN.bamGnut.30181	-4.64	0.00002	The pentatricopeptide repeat_1 (PPR)	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.47941	-4.63	0.00003	Unknown	
UoN.bamGnut.47220	-4.63	0.00000	Methyltransferase_16	
UoN.bamGnut.40519	-4.60	0.00001	Unknown	
UoN.bamGnut.6118	-4.59	0.00870	Rad60-SLD	
UoN.bamGnut.20831	-4.57	0.00003	Unknown	
UoN.bamGnut.11895	-4.56	0.00000	HAD	
UoN.bamGnut.13256	-4.56	0.00041	Glyco_transf_8	
UoN.bamGnut.7971	-4.55	0.00001	WD40	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.9165	-4.52	0.00076	Exostosin	
UoN.bamGnut.9081	-4.51	0.00094	GMC_oxred_C	
UoN.bamGnut.14582	-4.51	0.00005	Unknown	
UoN.bamGnut.49007	-4.51	0.00011	P34-Arc	
UoN.bamGnut.20142	-4.50	0.00001	Unknown	
UoN.bamGnut.20761	-4.49	0.00001	Peroxidase	Antioxidant/ ROS scavenging
UoN.bamGnut.25573	-4.48	0.00358	Unknown	
UoN.bamGnut.27822	-4.42	0.00002	MR5N	
UoN.bamGnut.10148	-4.40	0.00859	DEAD BOX	
UoN.bamGnut.44411	-4.39	0.00009	NA	
UoN.bamGnut.41899	-4.39	0.00018	Proton_antipo_M	
UoN.bamGnut.34250	-4.38	0.00006	Ribophorin_II	
UoN.bamGnut.31338	-4.38	0.00000	Lipase_GDSL	
UoN.bamGnut.19526	-4.37	0.00000	Acyl-thio_N	
UoN.bamGnut.33047	-4.36	0.00132	ABC_transport	Cellular transporter
UoN.bamGnut.18512	-4.36	0.00000	NA	
UoN.bamGnut.18561	-4.35	0.00002	PC-Esterase	
UoN.bamGnut.18520	-4.35	0.00007	SEO_N	
UoN.bamGnut.13405	-4.33	0.00173	RRM_5	
UoN.bamGnut.5967	-4.31	0.00000	YABBY	
UoN.bamGnut.40265	-4.26	0.00046	ENSMLUP00000000489	

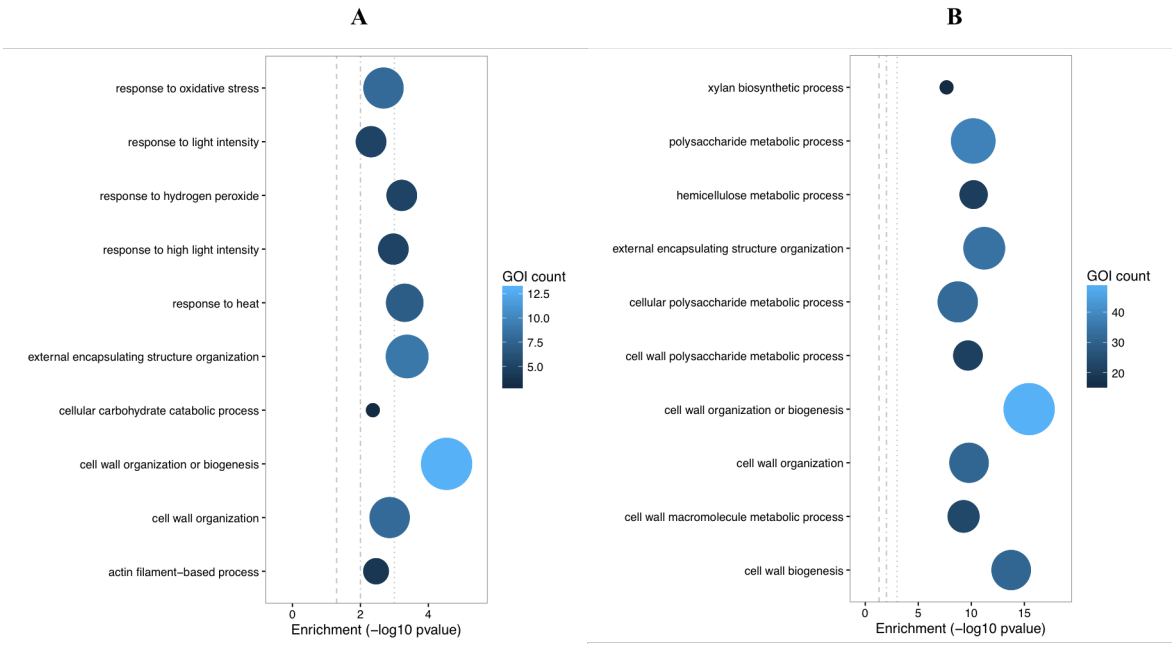

Figure 6.5: **Gene ontology classification of differentially expressed genes in drought stress under the 'biological process' category.** The top significant GO terms (having atleast 3 read counts) associated with differentially expressed transcripts in drought stress samples at 33 DAS. [A] and [B] refer to GO-terms assigned for DipC and TN, respectively. The p-value cutoff was set at 0.01. The GO-terms assigned are from Arabidopsis gene ontology reciprocal blast hits. GOI = Number of genes of interest with the respective gene ontology (GO) term.

6.2.3.1 Hierarchical cluster analysis

To visualise global gene expression across all the samples, a heat map was generated using heatmap.2. Genes with FDR < 0.05 and a log FC ≥ 1 were used for clustering analysis. In DipC, control samples at 33 DAS were more clustered together with drought samples at 33 DAS in comparison to control samples at 19 DAS, as shown by dendograms (fig. 6.6). This shows that the differences in the DipC sample gene expression between different time points is greater than the treatment (Drought vs Control) within the same time point. On the other hand for TN, control samples at different time-points (19 and 33 DAS) were clustered together in comparison to drought samples at 33 DAS, which indicates that the differences in gene expression

among treatments is greater than different time points (fig. 6.6). The heat map shows genes differentially expressed between different experimental conditions (Control-33 vs Control-19 and Drought-33 vs Control-33), where blue represents genes with higher gene expression, and red represents the ones with lower expression.

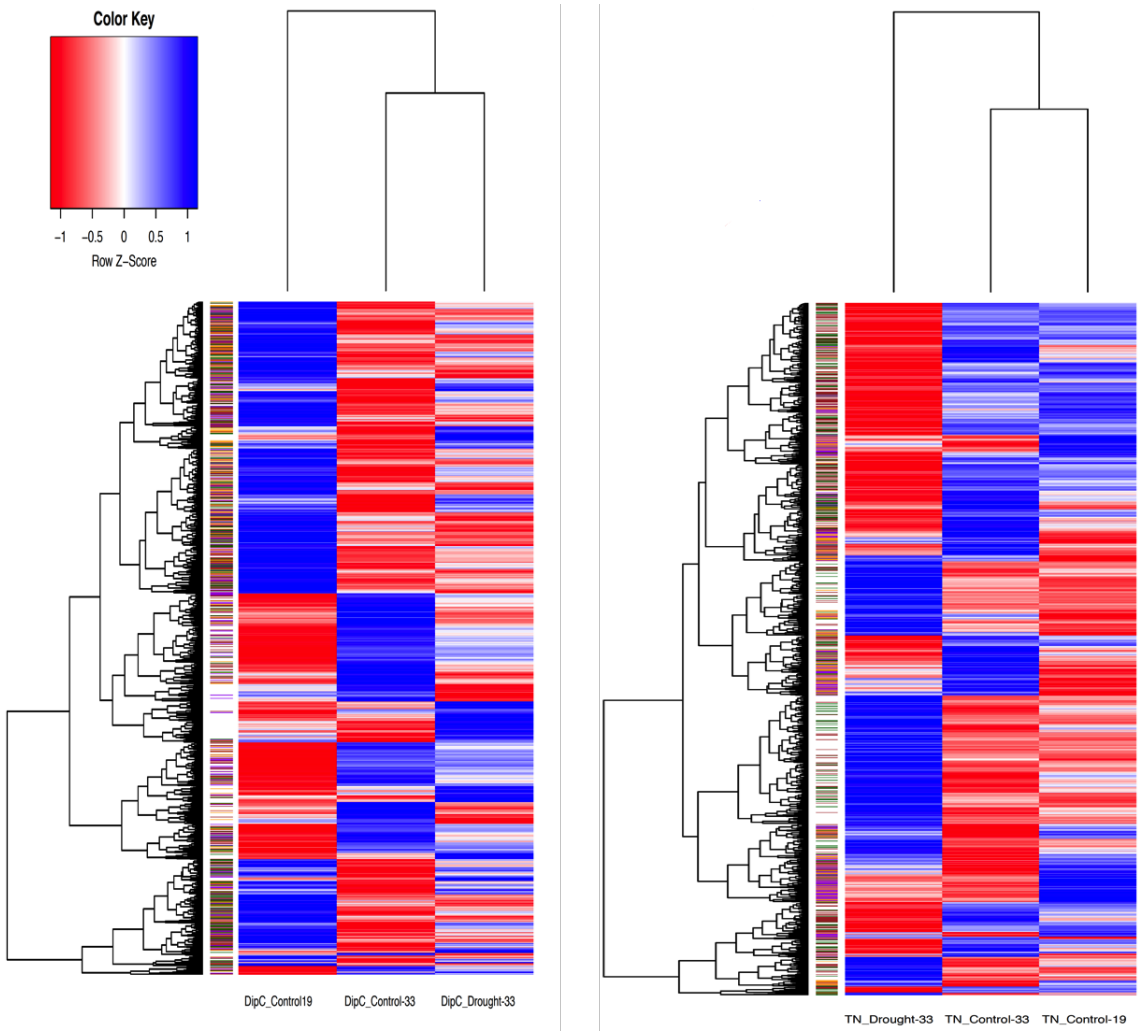

Figure 6.6: **Hierarchical clustering of global gene expression between different experimental conditions.** Heat map shows genes differentially expressed between control conditions at 19 and 33 DAS and drought condition at 33 DAS, where blue represents genes with higher gene expression and red represents the ones with lower expression. Genes were filtered with FDR < 0.05 and a log FC ≥ 1, where all samples had read count > 1 in all biological replicates. Clustering of samples within each experimental condition denoted by the dendrogram.

6.2.4 Differentially expressed genes at reproductive stage

At reproductive stage, for both DipC and TN, RNA-seq was done at time points 33 and 47 DAS (refer to 3.6 for RNA sampling details). A total of 1266 and 893 genes were found to be significantly differentially expressed in control (irrigated) condition at 47 DAS compared to control (irrigated) condition at 33 DAS in DipC and TN, respectively, through differential expression analysis using edgeR (Appendix 33 and 34). Among these genes, expression of 652 and 532 genes were down-regulated and 614 and 361 were up-regulated, respectively. Of these differentially expressed genes (DEG), 245 genes were found to be overlapping between DipC and TN (fig. 6.10 and Appendix 35). On the other hand, significantly DEG in drought compared to control condition at 47 DAS for DipC and TN accounts for 378 and 654 genes, respectively (Appendix 36 and 37). Among these genes, 193 and 252 were down-regulated and 185 and 402 were up-regulated under drought stress, respectively. Of these genes, only 65 genes were found to be overlapping between DipC and TN (Table 6.6 and fig. 6.10).

MDS figure shows that replicates are clustered in their respective experimental condition and time-point in DipC and TN, respectively (fig. 6.7 and 6.8). MA plot shows the log-fold change (i.e., the log ratio of normalized expression levels between two experimental conditions (i.e. control-47 DAS vs control-33 DAS) and (drought-47 DAS vs control-47 DAS)) and against the log CPM in both DipC and TN (fig. 6.9).

Top 100 up and down regulated genes according to log FC in drought compared to control (irrigated) condition at time point 47 DAS in DipC and TN are listed in tables 6.7 and 6.8. Based on the DEG analysis in drought versus control condition at time point 47 DAS, drought-related genes are divided into following categories:

1. ABA signal transduction associated gene expression

At reproductive stage in DipC, two gene homologues of protein phosphatase 2C (PP2C) gene family were differentially expressed in response to drought at 47 DAS.

Among the two DEG, one gene was up-regulated and other gene was down-regulated. In addition, family of transcription factors associated with ABA dependent pathway such as MYB and NAC genes were differentially expressed (either up or down) under drought stress (Appendix 36). On the other hand, three homologues of PP2C gene family were differentially expressed in TN in response to drought stress, of which two were up-regulated and one was down-regulated. In addition, six gene homologues of a known sub-type of cytochrome p450 family involved in ABA catabolism, transcription factors including seven MYB genes, five NAC genes, one bZIP gene and two WRKY genes were differentially expressed (either up or down) in TN (Appendix 37).

2. **Gene expression of osmoprotectants**

The genes related to osmotic adjustment that were up-regulated under drought stress compared to control at 47 DAS in both DipC and TN were mainly related to raffinose synthase and beta-fructofuranosidase (Appendix 36 and 37). However, unlike drought stress at the vegetative stage, other osmoprotectants such as trehalose phosphatase and aspartate proteases genes were not differentially expressed (either up or down) at the reproductive stage in both DipC and TN.

3. **Expression of genes involved in antioxidant production**

Gene homologues in the GST and glutaredoxin family were highly repressed under drought stress compared to control at 47 DAS in DipC (Appendix 36). On the other hand, TN showed high expression of antioxidants associated with genes in the GST, peroxidase, superoxide dismutase (Sod_Fe) and ferritin families in response to drought at 47 DAS (Appendix 37).

4. **Expression of genes related to photosynthesis and glycolysis**

Down-regulation of genes related to photosynthesis (genes in the PsbP and PRK family) were found in DipC in response to drought stress compared to control at 47 DAS (Appendix 36). On the other hand, TN showed up-regulation of an homologue

of PsbP gene family under drought stress. However, the LHC gene belonging to the light reaction pathway was down-regulated in TN in response to drought (Appendix 37).

With respect to glycolysis, only two homologues of O-fuct gene family were differentially expressed in DipC in response to drought at 47 DAS (Appendix 36). The expression level of the two homologues of O-fuct gene family were down-regulated under drought in DipC. On the other hand, the only homologue of O-fuct gene family found to be differentially expressed in response to drought in TN was up-regulated (Appendix 37).

5. Cell wall associated responses

With five gene homologues (related to cell wall) differentially expressed in DipC in response to drought at 47 DAS: three belonging to cellulose synthase and two to UDPGT gene family. Among the three DEG associated with cellulose synthase, two were up-regulated and one was down-regulated. Whereas of the two homologues of UDPGT gene family differentially expressed, one was up-regulated (Appendix 36). On the other hand, two cellulose synthase and two homologues of UDPGT gene family were up-regulated in TN in response to drought. In addition, one cellulose synthase and three homologues UDPGT gene family were down-regulated (Appendix 37).

6. Other genes induced by drought stress

Other genes related to drought such as LEA and HSP were up-regulated under drought stress compared to control at 47 DAS in DipC (Appendix 36). Several transporters (ABC, iron, and sugar transporters) were induced under drought stress. A gene (jas) related to jasmonate signaling pathway, involved in plant disease and defense was also induced under drought stress. Protein tyrosine phosphatase; an enzyme known to play an important role in stomatal closure through ABA signaling [316] was highly induced under drought stress in DipC. In addition, numerous protein ki-

nases (serine/threonine) were up-regulated under drought which perform diverse and important functions in plant signal transduction. On the other hand, several LEA, dehydrins, HSP, transporters (ABC, sulfate and sugar transporters), protein kinases and calmodulin related genes were up-regulated in response to drought in TN (Appendix 37). Furthermore, apart from ABA, other hormone such as auxin was induced under drought stress in both genotypes.

Out of the 65 genes that were found to be common between DipC and TN in response to drought at 47 DAS, some drought-related genes such as glutaredoxin associated with oxidative stress, raffinose synthase related to osmotic adjustment (UoN.bamGnut.43944), one photosynthesis related genes (PsbP), allene oxide cyclase (UoN.bamGnut.1864) associated with jasmonate biosynthesis, two cell wall related genes (UDPGT and cellulose synthase), two transcription factors (AUX-IAA and HOX) and other drought and ABA induced genes such as calmodulin(UoN.bamGnut-.39868), protein tyrosine phosphatase (UoN.bamGnut.27176), LEA3 (UoN.bamGnut-.51074), HSP20 (UoN.bamGnut.51074), PIP5K (UoN.bamGnut.24742) and MIPS were differentially expressed in both genotypes (see Table 6.6).

Furthermore, the enriched GO-terms were assigned to three principal categories: molecular function (12 GO terms), biological process (39 GO-terms) and cellular component (5 GO-terms) in DipC under drought stress at 47 DAS (Appendix 38). The top significant GO-terms assigned (having at least 3 read counts) for biological process in DipC were related to: response to water deprivation (GO:0009414), response to abiotic stimulus (GO:0009628), response to chemical (GO:0042221) and response to inorganic substance (GO:0010035) (see fig. 6.11 for top GO-terms assigned) (refer to Appendix 38 for the whole list of GO-terms assigned). On other hand, enriched GO-terms assigned for TN plants under drought stress at 47 DAS to three principal categories were: molecular function (6 GO terms), biological process (31 GO-terms)

and cellular component (5 GO-terms) (Appendix 39). The top GO-terms assigned for biological process in TN were related to: response to wounding (GO:0009611), response to water deprivation (GO:0009414), response to oxygen containing compound (GO:1901700), response to jasmonic acid (GO:0009753), response to inorganic substance (GO:0010035), response to chemical (GO:0042221), jasmonic acid signaling mediated pathway (GO:0009867) and cellular response to jasmonic acid stimulus (GO:0071395) (see Fig. 6.11 for top GO-terms assigned) (refer to Appendix 39 for the whole list of GO-terms assigned).

In addition, the enriched GO-terms assigned under the biological process category when compared between control conditions (Control-47 DAS vs Control-33 DAS) were made at different time points in DipC were mainly related to polysaccharide metabolic process (GO:0005976), polysaccharide catabolic process (GO:0000272), carbohydrate metabolic process (GO:0005975), regulation of reactive oxygen species metabolic process (GO:2000377) and response to external stimulus (GO:0009605) (refer to Appendix 40 for the whole list of GO-terms assigned). Whereas, for TN, the GO-terms assigned were mainly related to: response to ethylene (GO:0009723), response to abiotic stimulus (GO:0009628), oxidation-reduction (GO:0055114), ethylene-activated signaling pathway (GO:0009873) and anion transport (GO:0006820) (refer to Appendix 41 for the whole list of GO-terms assigned).

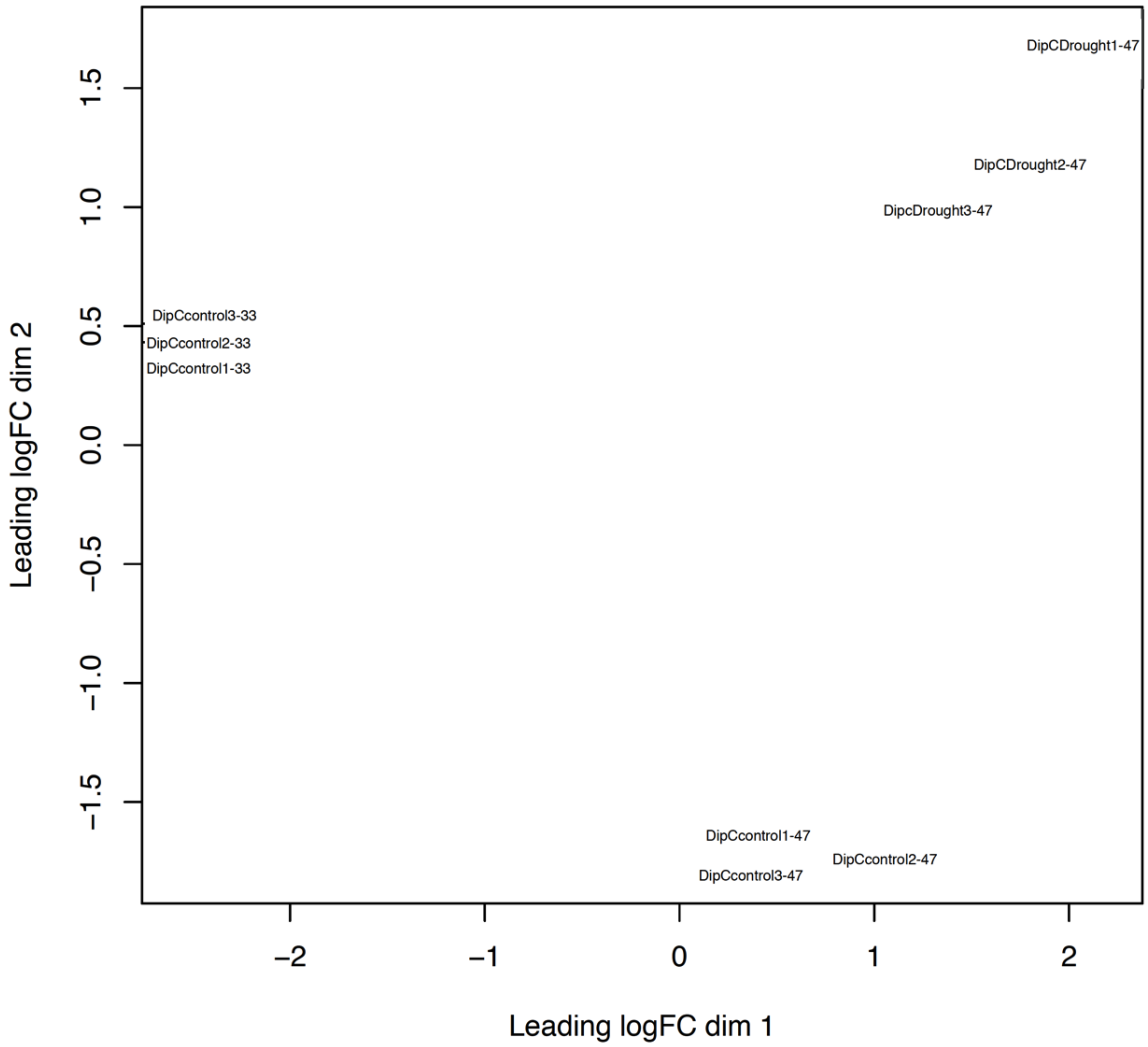

Figure 6.7: **Multidimensional scaling plot on differential gene expression of DipC samples at different conditions and time points at reproductive stage.** Replicates for control (irrigated) conditions at time points 33 and 47 DAS and drought conditions at time point 47 DAS are clustered at parallel level in their respective experimental condition and time-point. This shows that samples between experimental conditions and time points were distinctively different.

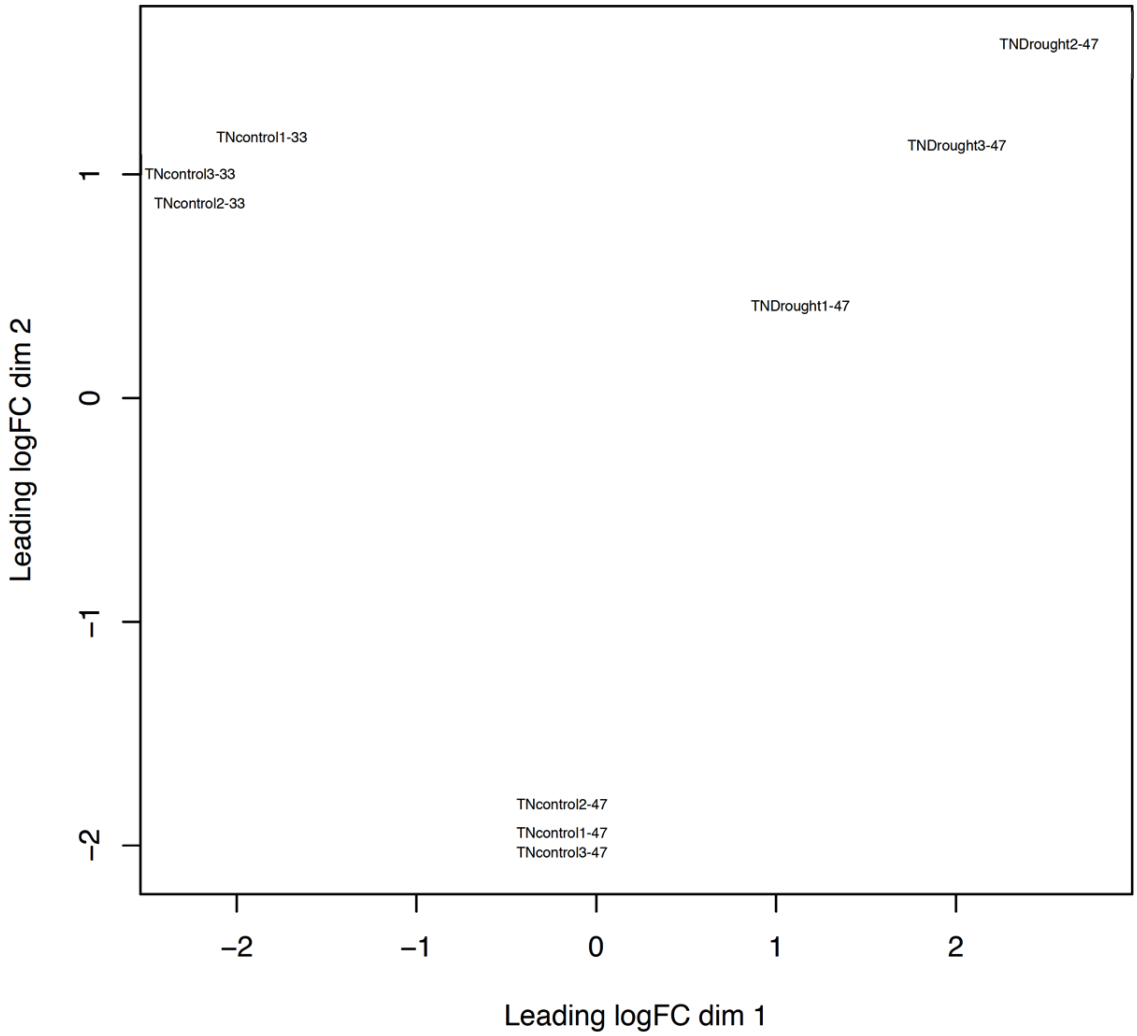

Figure 6.8: **Multidimensional scaling plot on differential gene expression of TN samples at different conditions and time points at reproductive stage.** Replicates for control (irrigated) conditions at time points 33 and 47 DAS and drought conditions at time point 47 DAS are clustered at parallel level in their respective experimental condition and time-point. This shows that samples between experimental conditions and time points were distinctively different.

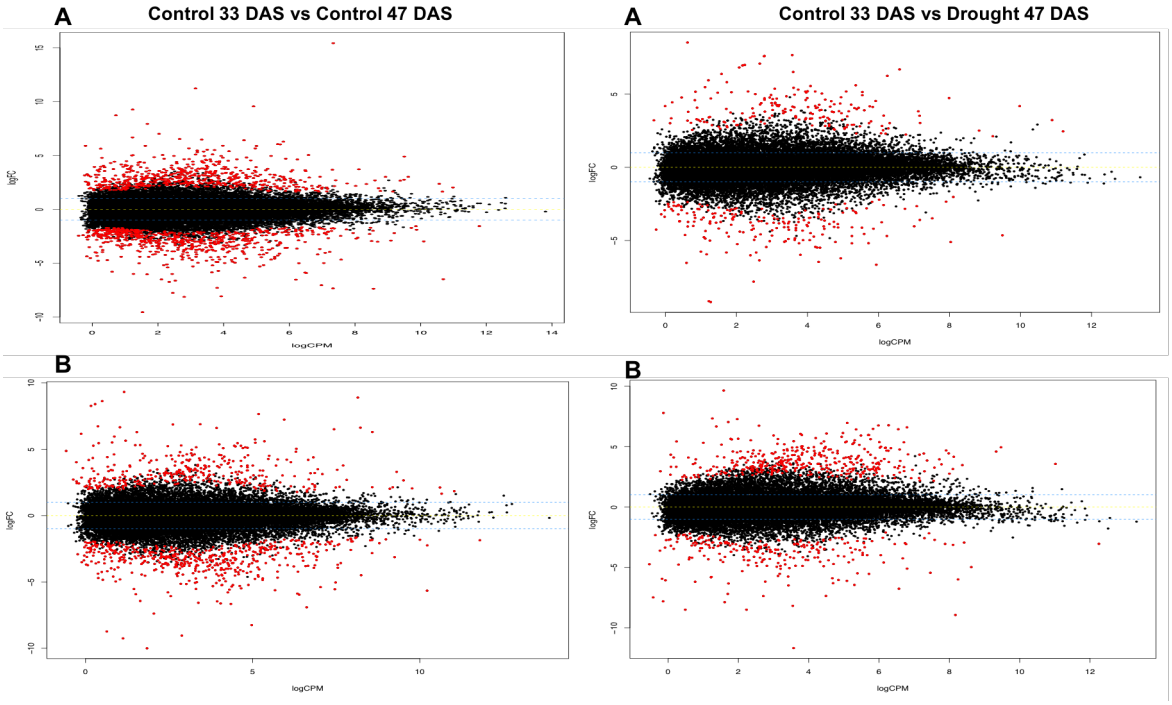

Figure 6.9: **MA plot comparing log FC and average log CPM on differential gene expression of Control-47 DAS vs Control-33 DAS and Drought-47 DAS vs Control-33 DAS at reproductive stage.** MA plot shows the log-fold change (i.e., the log ratio of normalized expression levels between two experimental conditions and against the log counts per million (CPM). [A] and [B] refers to plot from DipC and TN respectively. Experimental conditions of differential expression are listed on top of the figure. Those genes selected as differentially expressed (with FDR < 0.05, log fold change (logFC) ≥ 1) are highlighted as red dots.

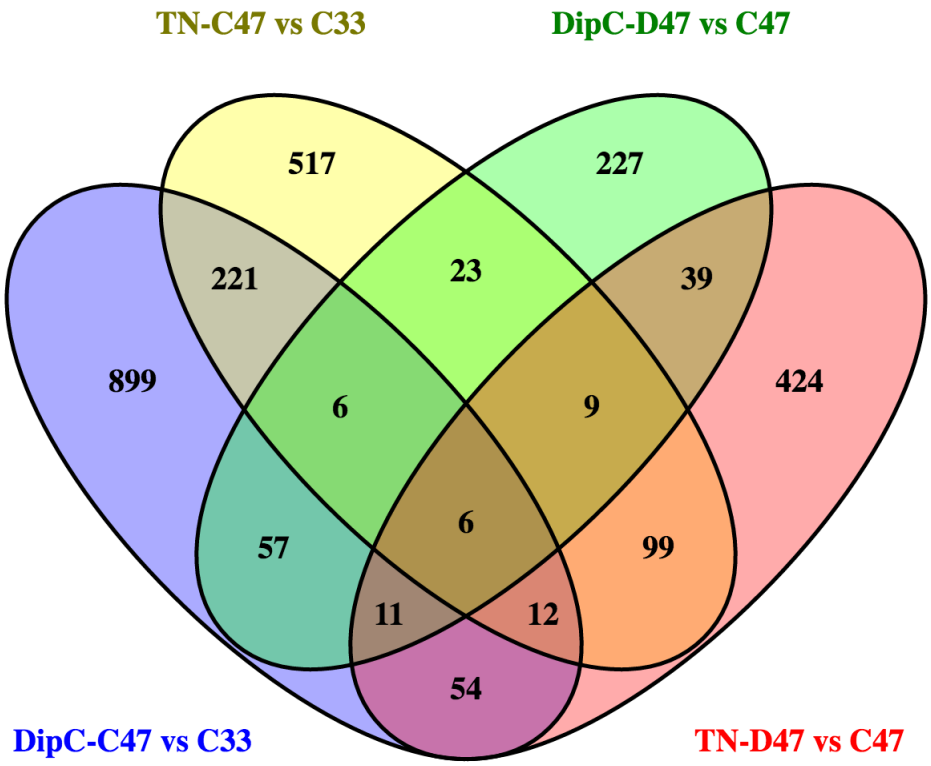

Figure 6.10: **Comparison of number of differentially expressed genes between DipC and TN at different experimental conditions.** A Venn diagram showing the number of differentially expressed genes in DipC (DipC-C47 vs C33 and DipC-D47 vs C47) and TN (TN-C47 vs C33 and TN-D47 vs C47) at different experimental conditions. C and D refer to control and drought respectively. 33 and 47 are the days after sowing at which the sampling was done. TN showed higher number of DE genes between TN-D47 vs C47 condition in comparison to DipC. Whereas DipC showed higher number of DE genes at DipC-C47 vs C33 condition.

Table 6.6: **Overlapping genes between DipC and TN in drought versus control (irrigated) condition at time point 47 DAS**

Gene-ID	**Annotation**	**Regulation under drought in DipC**	**Regulation under drought in TN**	**Gene description in relation to drought stress**
UoN.bamGnut.1864	Allene oxide cyclase	UP	UP	Catalysed step in jasmonate (JA) biosynthesis is important in the wound response
UoN.bamGnut.9499	PITH	Down	UP	
UoN.bamGnut.52874	NA	UP	UP	
UoN.bamGnut.51112	Unknown	UP	UP	
UoN.bamGnut.40032	PdomMRNAr1.1-08458.1	UP	UP	
UoN.bamGnut.43944	Raffinose_synthase	UP	UP	Osmoprotectant
UoN.bamGnut.52142	Unknown	UP	UP	
UoN.bamGnut.42037	LEA_1	UP	UP	Induced by drought stress
UoN.bamGnut.43646	Tryptophan-rich sensory protein (TspO)	UP	UP	TSPO was found to attenuate plant cell porphyria by delta-aminolevulinic acid levels and the accumulation of tetrapyrroles
UoN.bamGnut.39868	Calmodulin_binding	Down	UP	Signal transduction
UoN.bamGnut.39903	Tryp_alpha_amyl	UP	UP	
UoN.bamGnut.18596	DOCK-C2	Down	UP	
UoN.bamGnut.27176	Protein tyrosine phosphatase	UP	UP	ABA signaling. Plays role in stomatal closure.
UoN.bamGnut.14984	BURP	Down	UP	
UoN.bamGnut.46119	J9VLD2_CRYNH	UP	UP	
UoN.bamGnut.13682	PsbP (photosystem II subunit P-1)	Down	UP	Photosynthesis related genes
UoN.bamGnut.51074	LEA_3	UP	UP	Induced by drought stress
UoN.bamGnut.43793	Aldo_ket_red	UP	UP	
UoN.bamGnut.36333	Cellulose_synthase	UP	UP	Cell wall maintenance
UoN.bamGnut.43478	NAC 3	UP	UP	Transcription
UoN.bamGnut.26740	Trm112p	UP	UP	
UoN.bamGnut.35651	Protein kinase tyrosine	UP	UP	Signal transduction
UoN.bamGnut.42115	Cu-oxidase_3	UP	UP	
UoN.bamGnut.43139	Aminotran_3	UP	UP	
UoN.bamGnut.26068	PK-G12rRNA	UP	UP	
UoN.bamGnut.36887	Protein kinase tyrosine	UP	UP	
UoN.bamGnut.41053	DUF668	UP	UP	
UoN.bamGnut.4456	TPR_12	Down	UP	Repeats families in plants allow them to rapidly acclimatize to adverse conditions

UoN.bamGnut.42093	Lipase_3	UP	UP	
UoN.bamGnut.44857	Cellulose_synthase	UP	UP	Cell wall maintenance
UoN.bamGnut.19838	FAD_binding_7	UP	UP	
UoN.bamGnut.51739	BAG domain	UP	UP	Modulators of Chaperone activity
UoN.bamGnut.43813	HSP20	UP	UP	Induced by drought stress
UoN.bamGnut.40358	RPE65	UP	UP	
UoN.bamGnut.39167	AUX_IAA	UP	UP	Transcription
UoN.bamGnut.40849	Homeobox (HOX)	UP	UP	Transcription
UoN.bamGnut.26944	Cu_bind_like	Down	Down	
UoN.bamGnut.50834	HAND	Down	Down	
UoN.bamGnut.4147	DUF3474	Down	Down	
UoN.bamGnut.2307	Hydrolase_3	UP	Down	
UoN.bamGnut.42936	AAA	Down	Down	
UoN.bamGnut.24640	Tryp_alpha_amyl	Down	Down	
UoN.bamGnut.44145	His_Phos_2	Down	Down	
UoN.bamGnut.24742	Phosphatidylinositol-4-phosphate 5-Kinase (PIP5K)	Down	Down	Induced by drought stress and ABA
UoN.bamGnut.25563	Glyco_transf_21	Down	Down	
UoN.bamGnut.33191	AECH14963-PA	Down	Down	
UoN.bamGnut.53892	Cellulose_synthase	Down	Down	Cell wall maintenance
UoN.bamGnut.18405	Protein kinase	UP	Down	signal transduction
UoN.bamGnut.49138	Meth_synthase_2	Down	Down	
UoN.bamGnut.30410	CRAL_TRIO	Down	Down	
UoN.bamGnut.38564	Calreticulin	UP	Down	
UoN.bamGnut.45133	Peptidase_C2	UP	Down	
UoN.bamGnut.2247	Protein kinase	Down	Down	Signal transduction
UoN.bamGnut.26444	Glutaredoxin	Down	Down	antioxidant
UoN.bamGnut.25758	DUF4005	UP	Down	
UoN.bamGnut.17350	Unknown	Down	Down	
UoN.bamGnut.49985	Myo-inositol-1-phosphate synthase (MIPS)	Down	Down	Induced by drought stress and ABA
UoN.bamGnut.12765	ATP-synthase_ab	Down	Down	Energy metabolism
UoN.bamGnut.48171	Fe-S_biosynthesis	Down	Down	
UoN.bamGnut.10508	Protein kinase tyrosine	UP	Down	
UoN.bamGnut.21162	SMC_N	Down	Down	
UoN.bamGnut.11591	FAD_binding_4	Down	Down	
UoN.bamGnut.13344	Glucuronosyltransferase (UDPGT)	Down	Down	Involved in the formation of β- glucans which is involved in formation of cell wall
UoN.bamGnut.51799	Proton_antipo_M	UP	Down	
UoN.bamGnut.47308	NA	Down	Down	

Table 6.7: **Top UP-regulated genes in drought compared to control condition at time point 47 DAS in DipC and TN**

Gene_ID	logFC	FDR	Gene name	Gene description in relation to drought stress
UP-regulated in DipC				
UoN.bamGnut.53657	8.53	0.00000	ABC_membrane	Transporters
UoN.bamGnut.1651	7.67	0.00000	WD40	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.50740	7.61	0.00000	NUMOD3	
UoN.bamGnut.14707	7.59	0.00000	Q2HAQ6_CHAGB	
UoN.bamGnut.30028	7.09	0.00000	DUF3657	
UoN.bamGnut.49956	7.00	0.00000	Pkinase_Tyrosine	Related to cell division
UoN.bamGnut.18755	6.98	0.00000	Multi antimicrobial extrusion protein	
UoN.bamGnut.17272	6.96	0.00000	Auxin response transcription factor (ARF)	Organogenesis and embroyogenesis. mediated by recruitment of AUX-IAA
UoN.bamGnut.46682	6.82	0.00000	N2,N2-dimethylguanosine tRNA methyltransferase	
UoN.bamGnut.42727	6.69	0.00003	Photolyase, FAD binding 7	Photoreactivation enzyme for DNA repair
UoN.bamGnut.8705	6.52	0.00000	zf-C2H2_jaz	Transcription
UoN.bamGnut.11588	6.38	0.00007	NA	
UoN.bamGnut.26767	6.25	0.00035	NADH dehydroge-nase(ubiquinone) (Proton antiporter_C,)	
UoN.bamGnut.18091	5.95	0.00000	R4XEE5_TAPDE	
UoN.bamGnut.52681	5.82	0.00000	Multi antimicrobial extrusion protein (MatE)	
UoN.bamGnut.52874	5.61	0.00010	NA	
UoN.bamGnut.46976	5.56	0.00001	Ethylene-responsive binding factor-associated repression (EAR)	Negatively regulating genes involved in developmental, hormonal, and stress signaling pathways
UoN.bamGnut.11823	5.48	0.00000	UBA	Cell growth
UoN.bamGnut.33477	5.43	0.00082	Kinesin	
UoN.bamGnut.41280	5.37	0.00205	LRR_4	
UoN.bamGnut.49671	5.31	0.00034	Protein kinase_Tyrosine	Related to cell division
UoN.bamGnut.25064	5.20	0.00000	DnaJ	
UoN.bamGnut.36887	5.18	0.00034	Protein kinase_Tyrosine	
UoN.bamGnut.429	5.18	0.00076	AAA_18	Related to cell division

UoN.bamGnut.45068	5.18	0.00005	Q5KDC9_CRYNJ	
UoN.bamGnut.52142	5.18	0.00034	Unknown	
UoN.bamGnut.47254	5.17	0.00000	Glyco_transf_8	
UoN.bamGnut.40179	5.15	0.00002	Unknown	
UoN.bamGnut.47477	5.13	0.00076	Protein kinase_Tyrosine	Related to cell division
UoN.bamGnut.8599	5.11	0.00199	CemA	
UoN.bamGnut.42115	5.10	0.00013	Multi_copper_oxidase_3	Copper ion binding. Redox related enzyme
UoN.bamGnut.34881	5.04	0.00589	ATP-synt_DE_N	Energy metabolism
UoN.bamGnut.9551	5.00	0.00078	CTD_bind	
UoN.bamGnut.32188	4.96	0.00050	BTB	
UoN.bamGnut.31854	4.94	0.00237	CpeT	
UoN.bamGnut.1864	4.93	0.02576	Allene oxide cyclase	Catalysed step in jasmonate (JA) biosynthesis is important in the wound response
UoN.bamGnut.6458	4.86	0.01673	LRR_4	
UoN.bamGnut.38564	4.79	0.00158	Calreticulin	
UoN.bamGnut.18417	4.77	0.00007	BAH	
UoN.bamGnut.21413	4.77	0.00047	Response_reg	
UoN.bamGnut.43646	4.73	0.00168	Tryptophan-rich sensory protein_MBR (TspO)	TSPO was found to attenuate plant cell porphyria by delta-aminolevulinic acid levels and the accumulation of tetrapyrroles
UoN.bamGnut.99	4.73	0.00010	SSU_rRNA_eukarya	
UoN.bamGnut.28996	4.72	0.00043	Cation_ATPase_N	Energy metabolism
UoN.bamGnut.49414	4.71	0.00011	EF-hand_4	
UoN.bamGnut.16103	4.70	0.01675	GBA2_N	
UoN.bamGnut.100	4.66	0.00392	SSU_rRNA_eukarya	
UoN.bamGnut.52880	4.63	0.00630	Unknown	
UoN.bamGnut.40032	4.62	0.00092	PdomMRNAr1.1-08458.1	
UoN.bamGnut.51074	4.60	0.00205	LEA_3	Induced by drought stress
UoN.bamGnut.52304	4.59	0.00018	DUF212	
UoN.bamGnut.37687	4.53	0.00006	PP2C	Accelerates growth of inflorescence stems through the activation of cell proliferation and expansion.
UoN.bamGnut.50152	4.52	0.01292	ThiF	
UoN.bamGnut.40438	4.44	0.00017	EDR1	
UoN.bamGnut.26740	4.42	0.01403	Trm112p	
UoN.bamGnut.53922	4.40	0.00013	DnaJ	
UoN.bamGnut.15840	4.39	0.00163	SMN	
UoN.bamGnut.42373	4.37	0.00255	2OG-FeII_Oxy	
UoN.bamGnut.19838	4.36	0.00126	FAD_binding_7	
UoN.bamGnut.44857	4.31	0.00010	Cellulose_syntase	Cell wall maintenance
UoN.bamGnut.47834	4.29	0.00042	TIR	
UoN.bamGnut.40849	4.28	0.00098	Homeobox (HOX)	Transcription
UoN.bamGnut.6977	4.28	0.02286	Str_synth	

UoN.bamGnut.43813	4.27	0.00070	HSP20	Induced by drought stress
UoN.bamGnut.36333	4.26	0.01649	Cellulose_synt	Cell wall maintenance
UoN.bamGnut.5941	4.24	0.02844	Protein kinase	Signal transduction
UoN.bamGnut.14950	4.20	0.00834	Ion_trans_2	Transporters
UoN.bamGnut.35400	4.19	0.00707	NA	
UoN.bamGnut.51585	4.19	0.00057	mir-399	
UoN.bamGnut.16961	4.18	0.00262	DUF4283	
UoN.bamGnut.44721	4.18	0.00040	ENSEEUP00000011770	
UoN.bamGnut.42788	4.17	0.00010	CCT	
UoN.bamGnut.25573	4.15	0.02134	Unknown	
UoN.bamGnut.230	4.15	0.00567	Transposase_24	
UoN.bamGnut.6571	4.13	0.03869	Triose-phosphate Transporter famil	
UoN.bamGnut.46540	4.12	0.00055	ENSP00000310861	
UoN.bamGnut.1966	4.10	0.00597	Beta-fructofuranoisdae 10 (glycoside hydrolase family)	Hydrolysis of sucrose- Osmoprotectant
UoN.bamGnut.27176	4.09	0.02269	Protein tyrosine phosphatase	ABA signaling. Plays role in stomatal closure.
UoN.bamGnut.24082	4.08	0.00026	Unknown	
UoN.bamGnut.10508	4.07	0.00004	Protein kinase_Tyrosine	Related to cell division
UoN.bamGnut.28084	4.05	0.00170	PB1	
UoN.bamGnut.11038	4.05	0.03959	mRNA_cap_C	
UoN.bamGnut.54121	4.03	0.03590	POTRA	
UoN.bamGnut.20710	3.99	0.00119	NB-ARC	
UoN.bamGnut.52806	3.99	0.00327	DNA_ligase_A_M	
UoN.bamGnut.29245	3.98	0.04313	TIR_2	
UoN.bamGnut.24403	3.92	0.00018	PDDEXK_6	
UoN.bamGnut.42645	3.92	0.02538	BES1_N	regulate brassinosteroid-induced genes. Modulate major components of anti oxidative system
UoN.bamGnut.2307	3.90	0.04679	Hydrolase_3	
UoN.bamGnut.39903	3.90	0.00153	Tryp_alpha_amyl	
UoN.bamGnut.34272	3.90	0.04386	Unknown	
UoN.bamGnut.45094	3.86	0.03449	Protein kinase_Tyrosine	Related to cell division
UoN.bamGnut.18584	3.84	0.02128	zf-RING_11	Transcription
UoN.bamGnut.21379	3.83	0.00001	Unknown	
UoN.bamGnut.13764	3.81	0.04561	TPL-binding domain in jasmonate signaling (jas)	TPL proteins function as negative regulators of jasmonate responses.
UoN.bamGnut.54188	3.80	0.02778	FBpp0173880	
UoN.bamGnut.17859	3.79	0.00069	ABC1	Transporters

UoN.bamGnut.7407	3.79	0.03358	zf-RING_2	Transcription
UoN.bamGnut.33047	3.77	0.02781	ABC_transport	Transporters
UP-regulated in TN				
UoN.bamGnut.38403	9.64	0.00000	Peptidase_S10	
UoN.bamGnut.51033	7.78	0.00000	Protein kinase	Signal transduction
UoN.bamGnut.36833	7.33	0.00000	Ion_trans_2	Transporters
UoN.bamGnut.13538	7.27	0.00000	B2WMX7_PYRTR	
UoN.bamGnut.4277	7.03	0.00000	LRR_4	
UoN.bamGnut.39437	6.96	0.00000	bHLH	Transcription
UoN.bamGnut.1864	6.77	0.00070	Allene_oxide_cyclase	Catalysed step in jasmonate (JA) biosynthesis is important in the wound response
UoN.bamGnut.39648	6.75	0.00000	Extensin_1	
UoN.bamGnut.22901	6.73	0.00000	DUF829	
UoN.bamGnut.48316	6.67	0.00000	PQ-loop	
UoN.bamGnut.26296	6.59	0.00032	Haloacid dehalogenase, N terminal	
UoN.bamGnut.9499	6.46	0.00034	PITH	
UoN.bamGnut.52874	6.44	0.00000	NA	
UoN.bamGnut.45678	6.42	0.00083	Ribosomal_S7e	
UoN.bamGnut.39775	6.12	0.00012	Trp_Tyr_perm	
UoN.bamGnut.27002	6.08	0.00158	Carbonic anhydrase (Pro_CA)	Related to photosynthesis and plays role in ROS scavenging
UoN.bamGnut.53588	6.05	0.00001	PHD	Transcription
UoN.bamGnut.53003	5.96	0.00584	C4Y7D3_CLAL4	
UoN.bamGnut.38616	5.96	0.00108	AhpC-TSA	
UoN.bamGnut.44232	5.96	0.00000	Glyco_transf_5	
UoN.bamGnut.37110	5.92	0.00007	LRR_8	
UoN.bamGnut.43638	5.92	0.00000	Glutathione S-transferase_3, C-terminal domain	Antioxidant
UoN.bamGnut.41438	5.81	0.00002	Bet_v_1	
UoN.bamGnut.9667	5.71	0.00079	RNB	
UoN.bamGnut.51524	5.65	0.00005	GDSL-like Lipase/Acylhydrolase	
UoN.bamGnut.51112	5.64	0.00000	Unknown	
UoN.bamGnut.41797	5.63	0.00000	Fatty acid hydroxylase superfamily	
UoN.bamGnut.4534	5.61	0.00079	AP2	Transcription
UoN.bamGnut.8170	5.59	0.00034	DUF668	
UoN.bamGnut.40032	5.52	0.00000	PdomMRNAr1.1-08458.1	

UoN.bamGnut.14511	5.52	0.00015	Thiamin pyrophosphokinase, catalytic domain	
UoN.bamGnut.39360	5.45	0.00000	NA	
UoN.bamGnut.33147	5.43	0.00264	TCP	
UoN.bamGnut.32863	5.42	0.00001	Abhydrolase_1	
UoN.bamGnut.14952	5.42	0.00367	DUF761	
UoN.bamGnut.2022	5.37	0.00188	COX2_TM	
UoN.bamGnut.35649	5.36	0.00140	DDE_Tnp_4	
UoN.bamGnut.52668	5.35	0.00064	ATP-synt_ab_N	Energy metabolism
UoN.bamGnut.43193	5.33	0.00009	Terpene_synth_C	
UoN.bamGnut.53624	5.31	0.00026	BTB	
UoN.bamGnut.38380	5.26	0.00087	Unknown	
UoN.bamGnut.48031	5.23	0.00002	Dehydrin	Induced by drought stress
UoN.bamGnut.42506	5.21	0.00196	DUF4228	
UoN.bamGnut.19390	5.19	0.00002	FIO1_SCHPO	
UoN.bamGnut.53740	5.16	0.00002	Sakowv30017535m	
UoN.bamGnut.38337	5.14	0.00000	LEA_4	Induced by drought stress
UoN.bamGnut.41260	5.02	0.00018	Transferase	
UoN.bamGnut.2112	5.01	0.00113	Sin3a_C	
UoN.bamGnut.40156	4.99	0.00153	PC-Esterase	
UoN.bamGnut.42294	4.97	0.00008	NA	
UoN.bamGnut.35988	4.96	0.01028	Lipoxygenase	Induced by drought stress
UoN.bamGnut.43944	4.94	0.00001	Raffinose_synthase	Osmoprotectant
UoN.bamGnut.38408	4.94	0.01434	GDP-fucose protein O-fucosyltransferase (O-fuct)	Glycolysis/carbohydrate metabolism
UoN.bamGnut.40340	4.92	0.00004	bHLH	Transcription
UoN.bamGnut.40385	4.91	0.00249	Aa_transport	Transporters
UoN.bamGnut.52142	4.91	0.00002	Unknown	
UoN.bamGnut.40054	4.89	0.00079	bHLH	Transcription
UoN.bamGnut.42584	4.89	0.00003	NA	
UoN.bamGnut.42037	4.88	0.00012	LEA_1	Induced by drought stress
UoN.bamGnut.40279	4.87	0.00373	Stress-antifung	
UoN.bamGnut.39428	4.87	0.00070	Calmodulin_binding	Signal transduction
UoN.bamGnut.13427	4.87	0.00018	Peptidase_C2	
UoN.bamGnut.43646	4.80	0.00001	Tryptophan-rich sensory protein (TspO)	TSPO was found to attenuate plant cell porphyria by delta-aminolevulinic acid levels and the accumulation of tetrapyrroles
UoN.bamGnut.16534	4.78	0.00010	IQ	
UoN.bamGnut.39247	4.77	0.00088	3-beta hydroxysteroid dehydrogenase/isomerase family	

UoN.bamGnut.38433	4.76	0.00229	tify	
UoN.bamGnut.38839	4.76	0.00233	Unknown	
UoN.bamGnut.52130	4.75	0.00079	HSP20	Induced by drought stress
UoN.bamGnut.39758	4.70	0.00442	DIOX_N	
UoN.bamGnut.41199	4.69	0.00021	cytochrom p450	
UoN.bamGnut.38584	4.68	0.00022	ABC_transport	Transporters
UoN.bamGnut.42640	4.68	0.00009	p450	
UoN.bamGnut.47045	4.67	0.02719	zf-rbx1	Transcription
UoN.bamGnut.39868	4.67	0.00454	Calmodulin_binding	Signal transduction
UoN.bamGnut.40510	4.67	0.00070	LEA_3	Induced by drought stress
UoN.bamGnut.52859	4.66	0.00351	HAD	
UoN.bamGnut.38424	4.65	0.00424	C1H9N8_PARBA	
UoN.bamGnut.22033	4.64	0.03582	Phi_1	
UoN.bamGnut.8687	4.64	0.02653	STT3	
UoN.bamGnut.10538	4.64	0.00449	DUF3411	
UoN.bamGnut.28465	4.61	0.00003	Xan_ur_permease	
UoN.bamGnut.49070	4.60	0.00088	ENSPVAP00000008033	
UoN.bamGnut.51577	4.60	0.00292	Protein kinase	Signal transduction
UoN.bamGnut.39903	4.59	0.00000	Tryp_alpha_amyl	
UoN.bamGnut.40573	4.59	0.01011	Unknown	
UoN.bamGnut.41730	4.57	0.00322	Abhydrolase_5	
UoN.bamGnut.41515	4.55	0.00672	Mito_carrier	
UoN.bamGnut.38885	4.53	0.00393	NA	
UoN.bamGnut.38677	4.52	0.00088	Lipase_GDSL	
UoN.bamGnut.5012	4.51	0.00249	DUF946	
UoN.bamGnut.36306	4.47	0.00038	Smp_047650.1:pep	
UoN.bamGnut.45667	4.45	0.00228	ABC2_membrane	Transporters
UoN.bamGnut.47897	4.44	0.00804	RVT_2	
UoN.bamGnut.45337	4.44	0.00000	Ferritin	Antioxidant
UoN.bamGnut.39037	4.44	0.01046	NA	
UoN.bamGnut.24302	4.43	0.03591	Rer1	
UoN.bamGnut.22875	4.43	0.00009	Terpene_synth	
UoN.bamGnut.38976	4.41	0.00207	DUF4228	
UoN.bamGnut.17803	4.40	0.00140	Rxt3	

Table 6.8: **Top down-regulated genes in drought compared to control condition at time point 47 DAS in DipC and TN**

Gene_ID	**logFC**	**FDR**	**Annotation**	**Gene description in relation to drought stress**
Down-regulated in DipC				
UoN.bamGnut.3920	-2.02	0.04602	TauE	
UoN.bamGnut.2481	-2.21	0.04294	Carbonic anhydrase (Pro_CA)	Related to photosynthesis and plays role in ROS scavenging
UoN.bamGnut.19494	-2.35	0.01881	Phosphotransferase enzyme family	The PTS transports sugars/carbohydrate (such as glucose, mannose, and mannitol) into the cell
UoN.bamGnut.53972	-2.36	0.03399	Pro_isomerase	
UoN.bamGnut.3921	-2.37	0.04779	TauE	
UoN.bamGnut.49691	-2.38	0.04756	DUF5082	
UoN.bamGnut.31458	-2.48	0.02763	Hydrolase	
UoN.bamGnut.21615	-2.49	0.03779	NA	
UoN.bamGnut.8852	-2.53	0.03663	Q5AP56_CANAL	
UoN.bamGnut.23039	-2.53	0.04864	Glutathione S-transferase-2, N-terminal domain (GST)	Antioxidant/ ROS scavenging
UoN.bamGnut.3879	-2.55	0.04876	Unknown	
UoN.bamGnut.4351	-2.56	0.02872	NA	
UoN.bamGnut.21721	-2.58	0.01649	Unknown	
UoN.bamGnut.8071	-2.60	0.04238	Tetratricopeptide repeat_14 (TPR)	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.21052	-2.62	0.03512	NA	
UoN.bamGnut.26944	-2.62	0.04150	Cu_binding_like	
UoN.bamGnut.20079	-2.65	0.02872	NA	
UoN.bamGnut.21755	-2.65	0.02317	SapB_2	
UoN.bamGnut.52522	-2.65	0.02778	Unknown	
UoN.bamGnut.22575	-2.66	0.03358	Unknown	
UoN.bamGnut.48454	-2.70	0.02778	PC-Esterase	
UoN.bamGnut.10189	-2.72	0.01487	ubiquitin	
UoN.bamGnut.3826	-2.75	0.03371	NA	
UoN.bamGnut.45338	-2.79	0.03240	Phosphatidylethanolamine-binding protein	
UoN.bamGnut.30410	-2.79	0.03171	CRAL_TRIO	
UoN.bamGnut.36701	-2.80	0.03719	tRNA-synt_1	
UoN.bamGnut.24732	-2.83	0.03358	NmrA	
UoN.bamGnut.26992	-2.83	0.02120	Drought induced 19 protein (Di19), zinc-binding	Transcription
UoN.bamGnut.26591	-2.84	0.01120	NA	

UoN.bamGnut.5476	-2.84	0.02649	MYB_46	Transcription
UoN.bamGnut.21236	-2.85	0.04015	LRR_8	
UoN.bamGnut.10023	-2.85	0.01438	NA	
UoN.bamGnut.15007	-2.86	0.03358	ANAPC4_WD40	
UoN.bamGnut.35577	-2.87	0.04008	DUF3474	
UoN.bamGnut.8005	-2.88	0.04136	Pkinase_Tyrosine	Related to cell division
UoN.bamGnut.4318	-2.89	0.04294	Methyltransferase_3	
UoN.bamGnut.23132	-2.91	0.03171	YABBY	
UoN.bamGnut.32014	-2.93	0.00594	Unknown	
UoN.bamGnut.24627	-2.93	0.01621	NA	
UoN.bamGnut.9776	-2.94	0.03904	EFX90075	
UoN.bamGnut.44917	-2.95	0.04445	Adaptin_N	
UoN.bamGnut.33620	-2.96	0.00353	F8P6Y8_SERL9	
UoN.bamGnut.22352	-2.97	0.02140	Pectinesterase	
UoN.bamGnut.25416	-2.98	0.01881	SHN3 (AP2-EREBP)	Transcription
UoN.bamGnut.27169	-3.04	0.02510	E1-E2_ATPase	
UoN.bamGnut.17350	-3.05	0.00112	Unknown	
UoN.bamGnut.24402	-3.07	0.01140	Pollen_Ole_e_I	
UoN.bamGnut.4147	-3.10	0.00205	DUF3474	
UoN.bamGnut.6146	-3.13	0.00442	zf-C2H2_jaz	Transcription
UoN.bamGnut.41377	-3.13	0.04328	Abhydrolase_5	
UoN.bamGnut.18591	-3.13	0.02100	BAH	
UoN.bamGnut.44251	-3.14	0.00716	zf-LSD1	Transcription
UoN.bamGnut.5343	-3.14	0.00732	Beta-fructofuranoisdae 38 (glycoside hydrolase family)	Hydrolysis of sucrose- Osmoprotectant
UoN.bamGnut.21138	-3.15	0.04691	Unknown	
UoN.bamGnut.24742	-3.15	0.02756	Phosphatidylinositol-4-phosphate 5-Kinase (PIP5K)	Induced by water stress and ABA
UoN.bamGnut.5347	-3.17	0.00716	Myb_DNA-binding	Transcription
UoN.bamGnut.2095	-3.21	0.01496	NA	
UoN.bamGnut.47473	-3.21	0.00580	Myb_DNA-binding	Transcription
UoN.bamGnut.9165	-3.21	0.03953	Exostosin	
UoN.bamGnut.4834	-3.22	0.00206	NA	
UoN.bamGnut.28815	-3.23	0.02576	PRK (phosphoribulokinase)	Related to calvin cycle/photosynthesis
UoN.bamGnut.52140	-3.26	0.00205	Unknown	

UoN.bamGnut.9895	-3.30	0.04150	DUF3411	
UoN.bamGnut.30540	-3.31	0.03959	Response_regulator	
UoN.bamGnut.46804	-3.37	0.04980	zf-CCCH	Transcription
UoN.bamGnut.22459	-3.38	0.00254	MTHFR	
UoN.bamGnut.24565	-3.38	0.00644	Cu_bind_like	
UoN.bamGnut.30565	-3.40	0.02140	Prefoldin	
UoN.bamGnut.53716	-3.41	0.03818	Leo1	
UoN.bamGnut.15245	-3.43	0.00589	PLAC8	
UoN.bamGnut.10260	-3.44	0.00400	Peptidase_S8	
UoN.bamGnut.26564	-3.44	0.00644	Glucose-6-phosphate dehydrogenase (G6PD_C)	Glycolysis/carbohydrate metabolism
UoN.bamGnut.26098	-3.44	0.00557	MIR164	
UoN.bamGnut.223	-3.44	0.03430	Roc	
UoN.bamGnut.16241	-3.46	0.00161	Beta-fructofuranoisdae 79 (glycoside hydrolase family)	Hydrolysis of sucrose- Osmoprotectant
UoN.bamGnut.53032	-3.46	0.01205	DUF1762	
UoN.bamGnut.25563	-3.46	0.00327	Glyco_transf_21	
UoN.bamGnut.9549	-3.47	0.02778	Phosducin	
UoN.bamGnut.24640	-3.49	0.00003	Tryp_alpha_amyl	
UoN.bamGnut.3717	-3.49	0.00937	LTP_2	
UoN.bamGnut.24073	-3.51	0.01423	PPR_1	
UoN.bamGnut.47452	-3.51	0.00394	Beta-fructofuranoisdae 79 (glycoside hydrolase family)	Hydrolysis of sucrose- Osmoprotectant
UoN.bamGnut.38186	-3.54	0.00368	B3	
UoN.bamGnut.25387	-3.57	0.00557	Emex10998	
UoN.bamGnut.36770	-3.57	0.02844	UCH	
UoN.bamGnut.22948	-3.58	0.00716	LRR_4	
UoN.bamGnut.26421	-3.60	0.01240	Calreticulin	
UoN.bamGnut.16353	-3.61	0.00313	Protein kinase	Signal transduction
UoN.bamGnut.23948	-3.64	0.01264	Myb_DNA-bind_6	Transcription
UoN.bamGnut.9452	-3.67	0.00072	NA	
UoN.bamGnut.16644	-3.68	0.04015	The pentatricopeptide repeat_2 (PPR)	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.31368	-3.68	0.03904	Protein kinase	Signal transduction
UoN.bamGnut.45237	-3.70	0.00166	Ribosomal_L27	
UoN.bamGnut.21051	-3.70	0.00105	NA	
UoN.bamGnut.31993	-3.71	0.00229	SEO_N	

UoN.bamGnut.3584	-3.71	0.01843	Unknown	
UoN.bamGnut.27311	-3.72	0.04429	WD40	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.49138	-3.72	0.00805	Cobalamin-independent synthase_2	Methionine synthesis
UoN.bamGnut.31471	-3.73	0.00047	Lipase_GDSL	
UoN.bamGnut.50834	-3.73	0.00112	HAND	
Down-regulated in TN				
UoN.bamGnut.43811	-2.11	0.04378	Myb_DNA-binding	Transcription
UoN.bamGnut.12920	-2.17	0.01867	CobN-Mg_chel	
UoN.bamGnut.21604	-2.20	0.04362	Unknown	
UoN.bamGnut.9473	-2.29	0.04209	Auxin_inducible	Hormone signaling
UoN.bamGnut.4042	-2.30	0.04787	HIT	
UoN.bamGnut.32275	-2.31	0.03077	TAXi_C	
UoN.bamGnut.4426	-2.33	0.04343	NA	
UoN.bamGnut.8763	-2.39	0.02863	NA	
UoN.bamGnut.19980	-2.40	0.04143	RdRP	
UoN.bamGnut.39330	-2.41	0.04378	HAD_2	
UoN.bamGnut.4644	-2.44	0.04479	Beta-fructofuranoisdae 16 (glycoside hydrolase family)	Hydrolysis of sucrose- Osmoprotectant
UoN.bamGnut.5246	-2.45	0.04467	HPP	
UoN.bamGnut.12921	-2.51	0.02737	CobN-Mg_chel	
UoN.bamGnut.24701	-2.52	0.03453	Unknown	
UoN.bamGnut.48474	-2.53	0.02954	Ribonuc_L-PSP	
UoN.bamGnut.13673	-2.55	0.02719	Unknown	
UoN.bamGnut.21326	-2.55	0.04285	HD-ZIP_N	
UoN.bamGnut.2019	-2.55	0.02695	p450	
UoN.bamGnut.48959	-2.56	0.04730	Response_regulator	
UoN.bamGnut.27675	-2.58	0.00316	p450	
UoN.bamGnut.50758	-2.58	0.04843	Rhodanese_C	
UoN.bamGnut.7143	-2.60	0.02350	Gal_Lectin	
UoN.bamGnut.16934	-2.60	0.02880	ENSPTRP00000042564	
UoN.bamGnut.32583	-2.61	0.03427	GYF	
UoN.bamGnut.31339	-2.62	0.00839	Lectin_legB	
UoN.bamGnut.18161	-2.63	0.04606	HEAT	
UoN.bamGnut.29834	-2.64	0.03183	NA	

UoN.bamGnut.8335	-2.66	0.01934	Epimerase	
UoN.bamGnut.47587	-2.67	0.01090	Unknown	
UoN.bamGnut.29592	-2.68	0.03941	Terpene_synth_C	
UoN.bamGnut.16423	-2.68	0.00966	F-box	
UoN.bamGnut.11846	-2.70	0.04857	adh_short	
UoN.bamGnut.43022	-2.72	0.02351	Thioredoxin_6	Antioxidant/ ROS scavenging
UoN.bamGnut.14769	-2.72	0.01655	Lalb_13759	
UoN.bamGnut.22911	-2.72	0.00877	Cupin_1	
UoN.bamGnut.21762	-2.73	0.04200	TauE	
UoN.bamGnut.19613	-2.73	0.02737	Pectinacetylesterase (PAE)	
UoN.bamGnut.20731	-2.73	0.02800	TRM	
UoN.bamGnut.26944	-2.74	0.00884	Cu_bind_like	
UoN.bamGnut.50834	-2.77	0.02102	HAND	
UoN.bamGnut.35096	-2.78	0.04183	Glyco_transf_28	
UoN.bamGnut.4147	-2.79	0.02657	DUF3474	
UoN.bamGnut.12186	-2.81	0.02573	Roc	
UoN.bamGnut.28995	-2.82	0.00393	Glycosyl_transferase_1	
UoN.bamGnut.7580	-2.83	0.02650	Unknown	
UoN.bamGnut.26409	-2.86	0.02531	zf-C3HC4_2	Transcription
UoN.bamGnut.8789	-2.89	0.00384	NA	
UoN.bamGnut.32696	-2.90	0.01934	NA	
UoN.bamGnut.13327	-2.90	0.03639	GASA	
UoN.bamGnut.19770	-2.90	0.04209	zf-rbx1	Transcription
UoN.bamGnut.1261	-2.91	0.02020	Cytokin-binding	Hormone signaling
UoN.bamGnut.16452	-2.91	0.01972	p450	
UoN.bamGnut.47558	-2.92	0.01409	Kinase-like	
UoN.bamGnut.12095	-2.93	0.00991	PPR_3	Repeats families in plants allow them to rapidly acclimatize to adverse conditions
UoN.bamGnut.15536	-2.93	0.00680	LBP_BPI_CETP_C	
UoN.bamGnut.17222	-2.94	0.02163	Light-harvesting complexes of green plants (Chloroa_b-bind)	Related to photosynthesis and plays a positive rolein guard cell signaling in response to ABA
UoN.bamGnut.26845	-2.94	0.00543	Unknown	
UoN.bamGnut.33703	-2.95	0.00348	Unknown	
UoN.bamGnut.11492	-2.96	0.00273	UCH_1	
UoN.bamGnut.28267	-2.96	0.02737	F-box	
UoN.bamGnut.28	-2.99	0.01070	Rhomboid	
UoN.bamGnut.17723	-2.99	0.00135	Unknown	

UoN.bamGnut.26823	-3.00	0.04857	Beta-fructofuranoisdae 1 (glycoside hydrolase family)	Hydrolysis of sucrose- Osmoprotectant
UoN.bamGnut.2493	-3.00	0.01474	HMA	
UoN.bamGnut.44427	-3.01	0.03717	G-patch_2	
UoN.bamGnut.29513	-3.01	0.02839	NA	
UoN.bamGnut.42548	-3.02	0.04857	DEAD	
UoN.bamGnut.2307	-3.03	0.02882	Hydrolase_3	
UoN.bamGnut.34580	-3.03	0.01457	PMR5N	
UoN.bamGnut.31273	-3.03	0.02695	Sakowv30030116m	
UoN.bamGnut.35761	-3.03	0.01545	DUF1731	
UoN.bamGnut.26420	-3.05	0.00840	Unknown	
UoN.bamGnut.42936	-3.05	0.01691	AAA	
UoN.bamGnut.6209	-3.05	0.01490	B2WMX7_PYRTR	
UoN.bamGnut.7343	-3.06	0.00079	Acid_phosphatase_B	Acid phosphatase is known to act under stress by maintaining a certain level of inorganic phosphate in plant cells
UoN.bamGnut.51729	-3.06	0.04677	GTP-bdg_N	
UoN.bamGnut.24640	-3.07	0.00288	Tryp_alpha_amyl	
UoN.bamGnut.4084	-3.07	0.01721	bHLH	Transcription
UoN.bamGnut.27744	-3.08	0.00195	Kelch_6	
UoN.bamGnut.30416	-3.08	0.01246	Chorismate_binding	
UoN.bamGnut.17800	-3.09	0.04380	Unknown	
UoN.bamGnut.16211	-3.09	0.02857	SQHop_cyclase_N	
UoN.bamGnut.9335	-3.10	0.00990	PP2C	Negative regulator of ABA signaling
UoN.bamGnut.44145	-3.10	0.04810	His_Phos_2	
UoN.bamGnut.7020	-3.11	0.04730	DUF1313	
UoN.bamGnut.5064	-3.13	0.01928	M7U9J7_BOTF1	
UoN.bamGnut.25268	-3.14	0.00079	LTP_2	
UoN.bamGnut.38031	-3.15	0.00910	RRM_5	
UoN.bamGnut.24742	-3.16	0.03591	Phosphatidylinositol-4-phosphate 5-Kinase (PIP5K)	Induced by water stress and ABA
UoN.bamGnut.8752	-3.17	0.00353	Myb_DNA-binding	Transcription
UoN.bamGnut.23664	-3.17	0.00509	PHUM466290-PA	
UoN.bamGnut.38252	-3.18	0.04183	Methyltransf_16	
UoN.bamGnut.52779	-3.18	0.03359	G2WHX8_YEASK	
UoN.bamGnut.11565	-3.19	0.01105	zf-FLZ	Transcription
UoN.bamGnut.44808	-3.19	0.00122	Unknown	

UoN.bamGnut.9070	-3.20	0.04976	NA	
UoN.bamGnut.39551	-3.20	0.02366	DEAD	
UoN.bamGnut.14953	-3.21	0.00119	Globin	
UoN.bamGnut.9630	-3.23	0.03322	ResIII	
UoN.bamGnut.30655	-3.24	0.00611	Peptidase_M16	

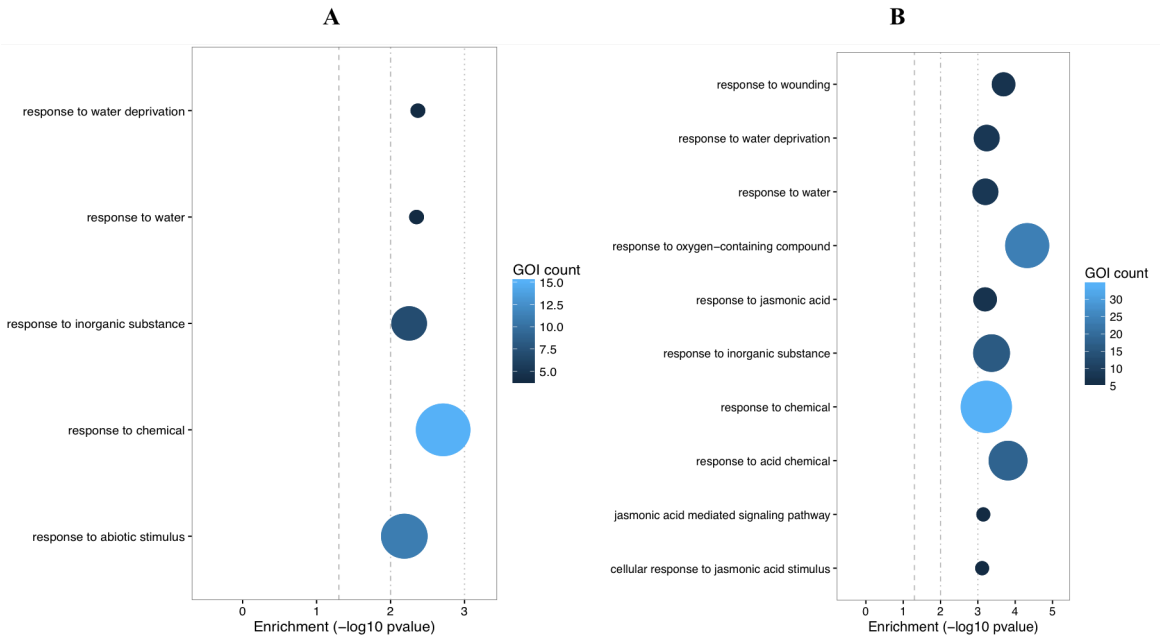

Figure 6.11: **Gene ontology classification of differentially expressed genes in drought stress under the 'biological process' category.** The top significant GO terms (having atleast 3 read counts) associated with differentially expressed transcripts in drought stress samples at 47 DAS. [A] and [B] refers to GO-terms assigned for DipC and TN, respectively. The pvalue cutoff was set at 0.01. The GO-terms assigned is from Arabidopsis gene ontology reciprocal blast hits. GOI = Number of genes of interest with the respective gene ontology (GO) term.

6.2.4.1 Hierarchical cluster analysis

To visualise global gene expression across all the samples, a heat map was generated using heatmap.2. Genes with FDR < 0.05 and a log FC ≥ 1 were used for clustering analysis. In both DipC and TN, control samples at 47 DAS were more clustered together with drought samples at 47 DAS in comparison to control samples at 33

DAS, as shown by dendograms (fig. 6.12). This shows that the differences in DipC and TN samples gene expression among different time points are greater than the treatment (Drought vs Control) within the same time point. The heat map shows genes differentially expressed between different experimental conditions (Control-47 vs Control-33 and Drought-47 vs Control-47), where blue represents genes with higher gene expression, and red represents the ones with lower expression.

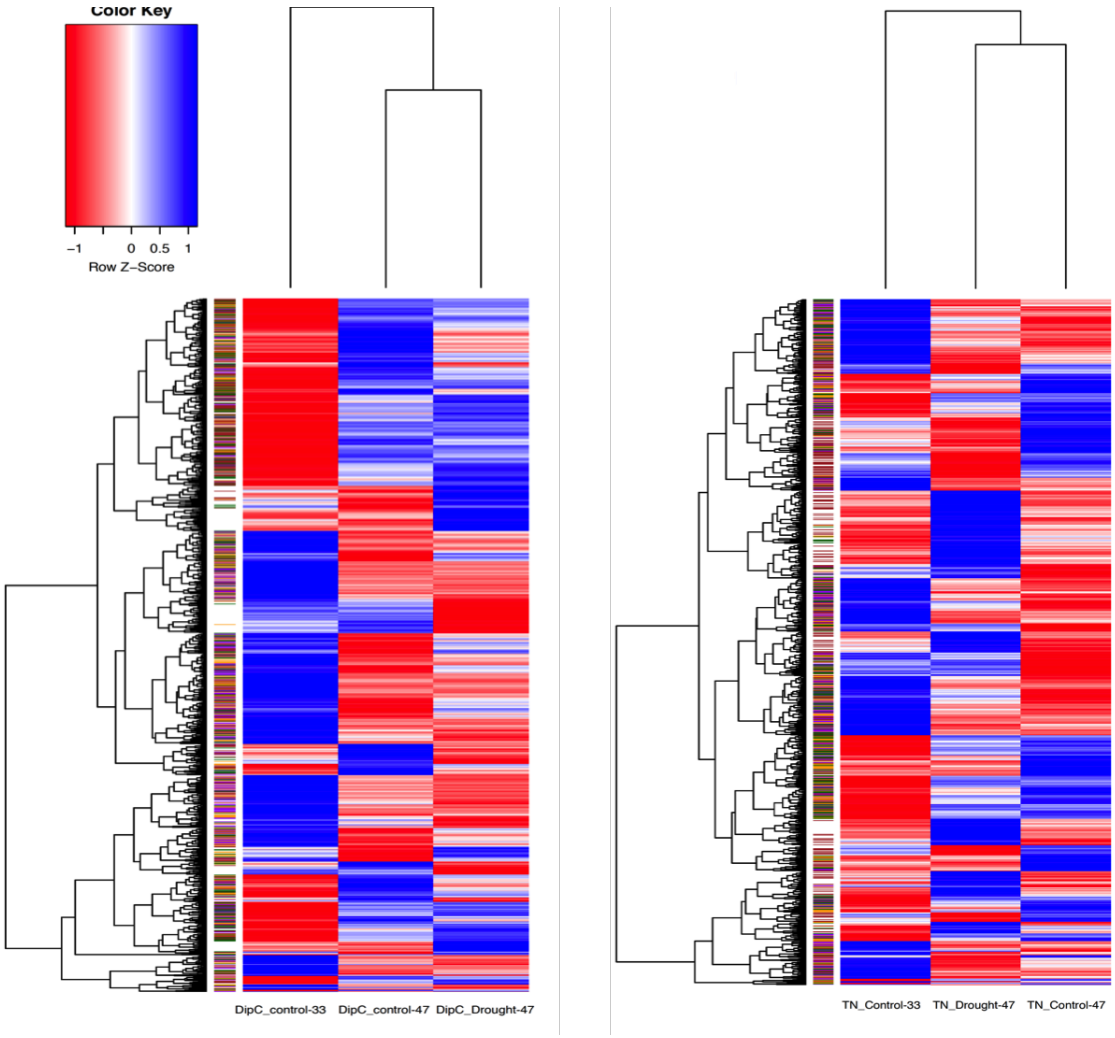

Figure 6.12: **Hierarchical clustering of global gene expression between different experimental conditions.** Heat map shows genes differentially expressed between control conditions at 33 and 47 DAS and drought condition at 47 DAS, where blue represents genes with higher gene expression, and red represents the ones with lower expression. Genes were filtered with FDR < 0.05 and a log FC ≥ 1, where all samples had read count > 1 in all biological replicates. Clustering of samples within each experimental condition denoted by the dendrogram.

6.2.5 Transcription factors associated with drought stress

The DE genes from both genotypes (DipC and TN) identified various Transcription(-related) Factors (TFs) at both vegetative and reproductive stages. Common TFs to both genotypes at vegetative stage are the down-regulation of *AP2/ERF*, Drought induced19 (*Di19*), zinc-binding and *RAD-like 6* (SANT/MYB TF) and up-regulation of *WRI1* (AP2-EREBP). *WRI1* TF is known to be involved in glycolysis metabolism [317]. In addition, other common TFs such as *zf-RING 2* and *MYB86* were up-regulated in DipC but down-regulated in TN (Table 6.3). Common TFs to both genotypes at reproductive stage are the up-regulation of *AUX_IAA*, Homeodomian (*HOX*) and *NAC3*. HOX and NAC family protein TFs are known to function in a wide variety of developmental processes and abiotic stress responses in plants [99]. On top of the common genes at vegetative stage, DipC shows up-regulation of 16 TFs (some are *NAC1*, *zf-CCHC*, *MYB46*, *zf-UDP*, *zf-B-box*, *zf-Dof*, *WRKY*, *zf-C2H2_jaz*, *MYB15*, *MYB6*, *zf-AN1*, *GRAS*, *zf-CCCH*, *zf-C3HC4_2*, *MYB* and *EMB1444* (bHLH-MYC) and down-regulation 6 others (*GRAS*, *EMB1444* (bHLH-MYC), *zf-AN1*, *GRAS*, *zf-CCCH* and *MYB*). TN, on the other hand, shows up-regulation of 37 TFs (some are *zf-RanBP*, *AUX_IAA*, *WRKY*, *HOX*, *zf-C2H2_jaz*, *zf-Nse*, *MYB36*, *WRKY33*, *NAC3*, *zf-C3HC4_3*, *MYB6*, *SRF-TF* (AGAMOUS-LIKE 12), *WRKY50* and *NAC100*) and down regulation of 41 others (some are *zf-C2H2_jaz*, *EMB1444*, *zf-RING_2*, *TINY2* (AP2-EREBP), *MYB6*, *zf-FLZ*, *zf-CCCH*, *NAC73*, *SHN3* (AP2-ERF), *bHLH*, *bZIP2* and *ERF1*(AP2)). At reproductive stage, DipC shows up-regulation of 6 TFs (*NAC*, *zf-RING_11*, *zf-RING_2*, *zf-C2H2_jaz* and *NAC*) and down-regulation of 11 others (*SHN3* (AP2-EREBP), *zf-C2H2_jaz*, *MYB46*, *BRX*, *zf-LSD1*, *bHLH*, *MYB*, *zf-Di19*, *MYB*, *MYB6* and *zf-CCCH*). While, TN at the reproductive stage shows up-regulation of 19 TFs (some are *NAC25*, *MYB2*, *zf-C2H2_6*, *WRKY28*, *bHLH*, *WRKY40*, *MYB78*, *zf-C2H2_6*, *PHD*, *MYB15* and *bZIP2*).

WRKY40 is known to be induced by ABA under drought stress [272]. Furthermore, common TFs to both developmental stages (vegetative and reproductive) in DipC are the down-regulation of *zf-Di19*, *zf-CCCH* and *MYB*. In addition, other common TFs in DipC such as *MYB46* and *zf-C2H2_jaz* were up-regulated at vegetative stage but down-regulated at reproductive stage. On the other hand, common TFs to both developmental stages in TN shows up-regulation of *AUX_IAA*, *NAC3* and *bHLH* and down-regulation of *MYB* and *zf-FLZ*. In addition, TF *zf-rbx1* was up-regulated at vegetative stage but down-regulated at reproductive stage in TN.

Separate co-expression networks were built for vegetative and reproductive stages in both DipC and TN (see Appendix 42 and 43 for DipC and Appendix 44 and 45 for TN), and the separate networks created for each genotype were merged in each stage to identify cases where co-expression between pairs of nodes is detected in both genotypes. In addition, a separate merged network was created to identify cases where co-expression between pairs of nodes is detected in both stages. The merged networks will give an idea of the direction of TFs linking to other genes. By looking at the number of links that each node has in the genotype-specific networks, it is possible to rank the potential importance of the different TFs (see Table 6.9 and 6.10). In the case of DipC at vegetative stage, *NAC1* stands out as being the TF with the most co-expressed genes with *zf-CCHC*, *MYB86*, *MYB46*, *zf-RING_2* and *zf-B_box* also looking relevant. For TN at vegetative stage, *zf-RanBP*, *zf-C2H2_jaz*, *AUX_IAA*, *WRKY*, *HOX*, *zf-C2H2_jaz*, *EMB1444*, *zf-Nse* and *MYB* (UoN.bamGnut.25299) seems important. On the other hand, important TFs in DipC at reproductive stage mostly consist of *SHN3* (AP2-EREBP), *zf-C2H2_jaz*, *MYB46*, *BRX* and *AUX_IAA*. Whereas for TN, *MYB* (UoN.bamGnut.8752), *NAC25*, *MYB2*, *zf-C2H2_6* and *WRKY28* stands out as relevant at reproductive stage. The TF *AUX_IAA* (UoN.bamGnut.39167) which was common to both genotype at reproductive stage also seems important which showed

the highest number of co-expression links connecting to other genes in the merged network (see Table 6.10). On the other hand, the merged network created in DipC between stages indicated importance of *MYB46* (UoN.bamGnut.5476) which showed the highest number of co-expression links connecting to other genes (Table 6.11). For the merged network created for TN between stages showed importance of *AUX_IAA* (UoN.bamGnut.39167) as this TF had the highest number of co-expression links connecting to other genes (Table 6.12). By looking at the results of the merged networks (between genotypes and also between stages), there is very little in common in terms of co-expression between pairs of nodes between genotypes as well as between stages. These observations are in accord with the results obtained from differential expression analysis from RNA-seq and cross species microarray (see chapter 4), where there was very little overlap observed between DipC and TN in terms of drought-responding genes. Hence the intersection is likely to be small as well. On the other hand, the small intersection seen between stages in both genotypes could be due to the fact that plants have gone past its vegetative stage into its floral stage.

Table 6.9: **Vertex degrees of differentially expressed Transcription factors in TN and DipC at vegetative stage**

TN				DipC			
Gene_ID	**Annotation**	**V°* whole**	**V°* for merged network**	**Gene_ID**	**Annotation**	**V°* whole**	**V°* for merged network**
UoN.bamGnut.53196	zf-RanBP	246	NA	UoN.bamGnut.51896	NAC1	238	NA
UoN.bamGnut.11042	zf-C2H2_jaz	186	NA	UoN.bamGnut.19689	zf-CCHC	201	NA
UoN.bamGnut.39167	AUX_IAA	160	NA	UoN.bamGnut.10038	MYB86	169	4
UoN.bamGnut.42217	WRKY	155	NA	UoN.bamGnut.5476	MYB46	199	NA
UoN.bamGnut.40849	HOX	138	2	UoN.bamGnut.25328	NAC	87	NA
UoN.bamGnut.51559	zf-C2H2_jaz	137	NA	UoN.bamGnut.9228	zf-RING_2	66	9
UoN.bamGnut.24800	bHLH-MYC_N (EMB1444)	126	NA	UoN.bamGnut.48677	zf-UDP	50	0
UoN.bamGnut.2786	zf-Nse	114	3	UoN.bamGnut.41951	zf-B_box	45	5
UoN.bamGnut.39367	MYB36	113	NA	UoN.bamGnut.4649	NAC	36	NA
UoN.bamGnut.7626	WRKY33	112	NA	UoN.bamGnut.39962	zf-Dof	36	0
UoN.bamGnut.43478	NAC3	109	NA	UoN.bamGnut.52340	zf-UDP	34	NA
UoN.bamGnut.38586	zf-C3HC4_3	100	NA	UoN.bamGnut.33128	zf-Dof	28	NA
UoN.bamGnut.25299	MYB	97	6	UoN.bamGnut.40504	WRI1 (AP2-EREBP)	26	0
UoN.bamGnut.52834	MYB6	93	NA	UoN.bamGnut.35944	WRKY	21	NA
UoN.bamGnut.8752	MYB	93	NA	UoN.bamGnut.14592	GRAS	16	NA
UoN.bamGnut.30177	SANT/MYB domain	89	NA	UoN.bamGnut.8660	AP2-ERF domain	14	2
UoN.bamGnut.12291	MYB	84	NA	UoN.bamGnut.6146	zf-C2H2_jaz	10	0
UoN.bamGnut.31501	zf-RING_2	82	NA	UoN.bamGnut.38935	MYB15	9	NA
UoN.bamGnut.45251	SRF-TF (AGAMOUS-LIKE 12)	79	NA	UoN.bamGnut.17348	RAD-like 6 (MYB TF)	9	0
UoN.bamGnut.5656	TINY2 (AP2-EREBP)	79	NA	UoN.bamGnut.44453	EMB1444	9	NA
UoN.bamGnut.17686	MYB6	77	2	UoN.bamGnut.26992	zf-Di19	8	0
UoN.bamGnut.17041	MYB61	76	NA	UoN.bamGnut.23947	MYB6	6	NA
UoN.bamGnut.25471	MYB	76	6	UoN.bamGnut.9193	zf-AN1	6	0
UoN.bamGnut.27580	MYB	74	NA	UoN.bamGnut.18435	GRAS	5	NA
UoN.bamGnut.27994	MYB	69	0	UoN.bamGnut.46804	zf-CCCH	5	NA
UoN.bamGnut.20995	RAD-like 6 (MYB TF)	63	0	UoN.bamGnut.26409	zf-C3HC4_2	4	NA

UoN.bamGnut.11565	zf-FLZ	60	0	UoN.bamGnut.47473	MYB	2	NA
UoN.bamGnut.41173	MYB	58	NA	UoN.bamGnut.43162	EMB1444	1	NA
UoN.bamGnut.2199	zf-CCCH	58	0				
UoN.bamGnut.24922	zf-UDP	57	NA				
UoN.bamGnut.21536	NAC73	57	4				
UoN.bamGnut.38558	NAC	54	NA				
UoN.bamGnut.50334	MYB	53	NA				
UoN.bamGnut.52114	zf-rbx1	53	NA				
UoN.bamGnut.19909	zf-UDP	49	NA				
UoN.bamGnut.25416	SHN3 (AP2-ERF domain)	49	NA				
UoN.bamGnut.37116	zf-AN1	45	NA				
UoN.bamGnut.14790	bHLH	43	0				
UoN.bamGnut.21448	zf-C3HC4_2	43	NA				
UoN.bamGnut.23387	bHLH	39	NA				
UoN.bamGnut.13321	HOX	39	NA				
UoN.bamGnut.39914	zf-rbx1	36	NA				
UoN.bamGnut.43965	SRF-TF (AGAMOUS-like 8)	34	NA				
UoN.bamGnut.44458	RAD-like 1 (SANT/MYB)	32	NA				
UoN.bamGnut.20007	AP2-ERF domain	31	NA				
UoN.bamGnut.27993	MYB	30	0				
UoN.bamGnut.40054	bHLH	30	NA				
UoN.bamGnut.24919	zf-DHHC	30	NA				
UoN.bamGnut.8661	bHLH	26	NA				
UoN.bamGnut.12335	zf-RING_2	26	NA				
UoN.bamGnut.36216	zf-UBR	25	NA				
UoN.bamGnut.3333	zf-rbx1	25	4				
UoN.bamGnut.17587	zf-HC5HC2H	25	0				
UoN.bamGnut.52036	MYB	24	NA				
UoN.bamGnut.8414	WRKY50	18	NA				
UoN.bamGnut.4430	AP2-EREBP	18	NA				
UoN.bamGnut.37586	NAC100	16	NA				
UoN.bamGnut.42550	bZIP2	15	NA				
UoN.bamGnut.9878	zf-C3HC4_2	15	NA				
UoN.bamGnut.13940	zf-C3HC4_4	14	0				
UoN.bamGnut.14726	bZIP1	14	NA				
UoN.bamGnut.6193	MEE3 (MYB TF)	13	0				
UoN.bamGnut.53377	zf-C3HC4_3	12	NA				
UoN.bamGnut.10856	zf-C3HC4_3	10	NA				
UoN.bamGnut.48161	zf-C3HC4_2	9	NA				

UoN.bamGnut.43998	zf-DHHC	9	NA				
UoN.bamGnut.13514	zf-RING_2	8	NA				
UoN.bamGnut.1277	ERF1(AP2)	8	NA				
UoN.bamGnut.11113	zf-RING_11	8	NA				
UoN.bamGnut.33083	AUX_IAA	7	NA				
UoN.bamGnut.28138	MYB	7	NA				
UoN.bamGnut.41882	zf-rbx1	5	NA				
UoN.bamGnut.24108	zf-A20	5	NA				
UoN.bamGnut.51302	AUX-IAA 14	4	NA				
UoN.bamGnut.21978	ZF-HD_dimer	4	NA				
UoN.bamGnut.12292	MYB	4	NA				
UoN.bamGnut.42189	bZIP2	3	NA				
UoN.bamGnut.23966	zf-Di19	2	NA				
Common TFs							
UoN.bamGnut.9228	zf-RING_2	86	9				
UoN.bamGnut.10038	MYB86	42	4				
UoN.bamGnut.8660	AP2-ERF domain	31	2				
UoN.bamGnut.40504	WRI1 (AP2-EREBP)	21	0				
UoN.bamGnut.26992	zf-Di19	7	0				
UoN.bamGnut.17348	RAD-like 6 (SANT/MYB TF)	7	0				

*(V° refers to the number of links of each TF node, in either the whole genotype-specific network, or merged network)

NA = Genes not found when networks were merged

Table 6.10: **Vertex degrees of differentially expressed Transcription factors in TN and DipC at reproductive stage**

TN				DipC			
Gene_ID	**Annotation**	**V°* whole**	**V°* for merged network**	**Gene_ID**	**Annotation**	**V°* whole**	**V°* for merged network**
UoN.bamGnut.8752	MYB	199	16	UoN.bamGnut.25416	SHN3 (AP2-EREBP)	624	2
UoN.bamGnut.39304	NAC25	99	NA	UoN.bamGnut.6146	zf-C2H2_jaz	537	NA
UoN.bamGnut.51895	MYB 2	67	NA	UoN.bamGnut.5476	MYB46	348	NA
UoN.bamGnut.38339	zf-C2H26	54	3	UoN.bamGnut.10310	BRX	333	NA
UoN.bamGnut.42702	WRKY28	48	2	UoN.bamGnut.39167	AUX_IAA	244	11
UoN.bamGnut.40247	NAC	35	NA	UoN.bamGnut.44251	zf-LSD1	89	NA
UoN.bamGnut.40340	bHLH	34	NA	UoN.bamGnut.37585	NAC	45	NA
UoN.bamGnut.52114	zf-rbx1	34	NA	UoN.bamGnut.9505	bHLH	39	NA
UoN.bamGnut.39437	bHLH	25	NA	UoN.bamGnut.43478	NAC 3	32	0
UoN.bamGnut.39011	MYB-like 102	23	0	UoN.bamGnut.47473	MYB	20	NA
UoN.bamGnut.31	NAC	22	NA	UoN.bamGnut.23622	NAC	16	NA
UoN.bamGnut.27611	zf-CCHC	20	NA	UoN.bamGnut.26992	zf-Di19	16	0
UoN.bamGnut.40054	bHLH	20	NA	UoN.bamGnut.18584	zf-RING11	11	0
UoN.bamGnut.20135	WRKY40	14	NA	UoN.bamGnut.7407	zf-RING_2	11	NA
UoN.bamGnut.43738	MYB78	14	16	UoN.bamGnut.5347	MYB	11	NA
UoN.bamGnut.40775	MYB	13	NA	UoN.bamGnut.40849	HOX	10	0
UoN.bamGnut.38850	zf-C2H26	12	0	UoN.bamGnut.28728	MYB 6	10	NA
UoN.bamGnut.53588	PHD	7	NA	UoN.bamGnut.8705	zf-C2H2_jaz	8	NA
UoN.bamGnut.26409	zf-C3HC42	7	NA	UoN.bamGnut.46804	zf-CCCH	4	NA
UoN.bamGnut.19770	zf-rbx1	7	NA	UoN.bamGnut.37586	NAC	3	0
UoN.bamGnut.14697	PHD	6	0				
UoN.bamGnut.38935	MYB15	5	NA				
UoN.bamGnut.47045	zf-rbx1	4	NA				
UoN.bamGnut.38859	NAC90	3	NA				
UoN.bamGnut.43811	MYB	2	NA				
UoN.bamGnut.11565	zf-FLZ	2	NA				

UoN.bamGnut.12377	bZIP2	2	NA				
Common TFs							
UoN.bamGnut.39167	AUX_IAA	207	11				
UoN.bamGnut.40849	HOX	179	0				
UoN.bamGnut.43478	NAC3	5	0				

*(V° refers to the number of links of each TF node, in either the whole genotype-specific network, or merged network)

NA = Genes not found when networks were merged

Table 6.11: **Vertex degrees of differentially expressed Transcription factors in DipC when networks were merged between stages**

Vegetative stage			**Reproductive stage**		
Gene_ID	**Annotation**	**V°* for merged network**	**Gene_ID**	**Annotation**	**V°* for merged network**
UoN.bamGnut.51896	NAC1	NA	UoN.bamGnut.25416	SHN3 (AP2-EREBP)	NA
UoN.bamGnut.19689	zf-CCHC	NA	UoN.bamGnut.6146	zf-C2H2_jaz	0
UoN.bamGnut.10038	MYB86	NA	UoN.bamGnut.5476	MYB46	18
UoN.bamGnut.25328	NAC	1	UoN.bamGnut.10310	BRX	NA
UoN.bamGnut.9228	zf-RING_2	1	UoN.bamGnut.39167	AUX_IAA	NA
UoN.bamGnut.48677	zf-UDP	0	UoN.bamGnut.44251	zf-LSD1	NA
UoN.bamGnut.41951	zf-B_box	NA	UoN.bamGnut.37585	NAC	NA
UoN.bamGnut.4649	NAC	NA	UoN.bamGnut.9505	bHLH	NA
UoN.bamGnut.39962	zf-Dof	NA	UoN.bamGnut.43478	NAC	NA
UoN.bamGnut.52340	zf-UDP	NA	UoN.bamGnut.47473	MYB	0
UoN.bamGnut.33128	zf-Dof	NA	UoN.bamGnut.23622	NAC	NA
UoN.bamGnut.40504	WRI1 (AP2-EREBP)	NA	UoN.bamGnut.26992	zf-Di19	0
UoN.bamGnut.35944	WRKY	NA	UoN.bamGnut.18584	zf-RING_11	0
UoN.bamGnut.14592	GRAS	NA	UoN.bamGnut.7407	zf-RING_2	NA
UoN.bamGnut.8660	AP2-ERF domain	0	UoN.bamGnut.5347	MYB	NA
UoN.bamGnut.38935	MYB15	NA	UoN.bamGnut.40849	HOX	0
UoN.bamGnut.17348	RAD-like 6 (MYB TF)	NA	UoN.bamGnut.28728	MYB6	NA
UoN.bamGnut.44453	bHLH-MYC_N	NA	UoN.bamGnut.8705	zf-C2H2_jaz	NA
UoN.bamGnut.23947	MYB6	NA	UoN.bamGnut.46804	zf-CCCH	0
UoN.bamGnut.9193	zf-AN1	NA	UoN.bamGnut.37586	NAC	NA
UoN.bamGnut.18435	GRAS	NA			
UoN.bamGnut.26409	zf-C3HC4_2	NA			
UoN.bamGnut.43162	EMB1444	NA			
Common TFs					
UoN.bamGnut.5476	MYB 46	18			
UoN.bamGnut.6146	zf-C2H2_jaz	0			
UoN.bamGnut.26992	zf-Di19	0			
UoN.bamGnut.46804	zf-CCCH	0			
UoN.bamGnut.47473	MYB	0			

*(V° refers to the number of links of each TF node, in the merged network)

NA = Genes not found when networks were merged

Table 6.12: **Vertex degrees of differentially expressed Transcription factors in TN when networks were merged between stages**

Vegetative stage			**Reproductive stage**		
Gene_ID	**Annotation**	**V°* for merged network**	**Gene_ID**	**Annotation**	**V°* for merged network**
UoN.bamGnut.53196	zf-RanBP	NA	UoN.bamGnut.39167	AUX_IAA	6
UoN.bamGnut.11042	zf-C2H2_jaz	NA	UoN.bamGnut.8752	MYB	0
UoN.bamGnut.42217	WRKY	NA	UoN.bamGnut.40849	HOX	NA
UoN.bamGnut.40849	HOX	4	UoN.bamGnut.39304	NAC25	NA
UoN.bamGnut.51559	zf-C2H2_jaz	NA	UoN.bamGnut.51895	MYB2	NA
UoN.bamGnut.24800	bHLH-MYC_N	NA	UoN.bamGnut.38339	zf-C2H2_6	NA
UoN.bamGnut.2786	zf-Nse	NA	UoN.bamGnut.42702	WRKY28	NA
UoN.bamGnut.39367	MYB36	NA	UoN.bamGnut.40247	NAC	NA
UoN.bamGnut.7626	WRKY33	NA	UoN.bamGnut.40340	bHLH	NA
UoN.bamGnut.38586	zf-C3HC43	NA	UoN.bamGnut.52114	zf-rbx1	1
UoN.bamGnut.25299	MYB	0	UoN.bamGnut.39437	bHLH	NA
UoN.bamGnut.52834	MYB6	NA	UoN.bamGnut.39011	MYB-like102	0
UoN.bamGnut.30177	SANT/MYB domain	NA	UoN.bamGnut.31	NAC	0
UoN.bamGnut.9228	zf-RING2	0	UoN.bamGnut.27611	zf-CCHC	0
UoN.bamGnut.12291	MYB	NA	UoN.bamGnut.40054	bHLH	0
UoN.bamGnut.31501	zf-RING2	3	UoN.bamGnut.20135	WRKY40	NA
UoN.bamGnut.45251	SRF-TF (AGAMOUS-LIKE 12)	0	UoN.bamGnut.43738	MYB78	NA
UoN.bamGnut.5656	TINY2 (AP2-EREBP)	NA	UoN.bamGnut.40775	Myb_CC _LHEQLE	NA
UoN.bamGnut.17686	MYB6	NA	UoN.bamGnut.38850	zf-C2H2_6	NA
UoN.bamGnut.17041	MYB61	NA	UoN.bamGnut.53588	PHD	NA
UoN.bamGnut.25471	MYB	0	UoN.bamGnut.26409	zf-C3HC4_2	NA
UoN.bamGnut.27580	MYB	NA	UoN.bamGnut.19770	zf-rbx1	NA
UoN.bamGnut.27994	MYB	0	UoN.bamGnut.14697	PHD	NA
UoN.bamGnut.20995	RAD-like 6 (MYB TF)	NA	UoN.bamGnut.38935	MYB15	NA
UoN.bamGnut.41173	MYB	NA	UoN.bamGnut.43478	NAC3	0
UoN.bamGnut.2199	zf-CCCH	NA	UoN.bamGnut.47045	zf-rbx1	NA
UoN.bamGnut.24922	zf-UDP	NA	UoN.bamGnut.38859	NAC90	NA
UoN.bamGnut.21536	NAC73	NA	UoN.bamGnut.43811	MYB	NA

UoN.bamGnut.38558	NAC	NA	UoN.bamGnut.11565	zf-FLZ	0
UoN.bamGnut.50334	MYB	NA	UoN.bamGnut.12377	bZIP2	NA
UoN.bamGnut.19909	zf-UDP	NA			
UoN.bamGnut.25416	SHN3 (AP2-ERF domain)	0			
UoN.bamGnut.37116	zf-AN1	NA			
UoN.bamGnut.14790	bHLH	NA			
UoN.bamGnut.21448	zf-C3HC4_2	NA			
UoN.bamGnut.10038	MYB 86	NA			
UoN.bamGnut.23387	bHLH	NA			
UoN.bamGnut.13321	HOX	NA			
UoN.bamGnut.39914	zf-rbx1	NA			
UoN.bamGnut.43965	SRF-TF (AGAMOUS-like 8)	NA			
UoN.bamGnut.44458	RAD-like 1 (SANT/MYB)	NA			
UoN.bamGnut.8660	AP2-ERF domain	NA			
UoN.bamGnut.20007	AP2-ERF domain	NA			
UoN.bamGnut.27993	MYB	NA			
UoN.bamGnut.24919	zf-DHHC	NA			
UoN.bamGnut.8661	bHLH	NA			
UoN.bamGnut.12335	zf-RING_2	NA			
UoN.bamGnut.36216	zf-UBR	NA			
UoN.bamGnut.3333	zf-rbx1	NA			
UoN.bamGnut.17587	zf-HC5HC2H	NA			
UoN.bamGnut.52036	MYB	1			
UoN.bamGnut.40504	WRI1 (AP2-EREBP)	NA			
UoN.bamGnut.8414	WRKY50	NA			
UoN.bamGnut.4430	AP2-EREBP	NA			
UoN.bamGnut.37586	NAC100	NA			
UoN.bamGnut.42550	bZIP 2	NA			
UoN.bamGnut.9878	zf-C3HC4_2	NA			
UoN.bamGnut.13940	zf-C3HC4_4	NA			
UoN.bamGnut.14726	bZIP1	0			
UoN.bamGnut.6193	MEE3 (MYB TF)	NA			
UoN.bamGnut.53377	zf-C3HC4_3	NA			

UoN.bamGnut.10856	zf-C3HC4_3	NA			
UoN.bamGnut.48161	zf-C3HC4_2	NA			
UoN.bamGnut.43998	zf-DHHC	NA			
UoN.bamGnut.13514	zf-RING_2	NA			
UoN.bamGnut.1277	ERF1(AP2)	NA			
UoN.bamGnut.11113	zf-RING_11	NA			
UoN.bamGnut.33083	AUX_IAA	0			
UoN.bamGnut.26992	zf-Di19	NA			
UoN.bamGnut.17348	RAD-like 6 (MYB TF)	NA			
UoN.bamGnut.28138	MYB	NA			
UoN.bamGnut.41882	zf-rbx1	NA			
UoN.bamGnut.24108	zf-A20	NA			
UoN.bamGnut.51302	AUX-IAA 14	NA			
UoN.bamGnut.21978	ZF-HD_dimer	NA			
UoN.bamGnut.12292	MYB	NA			
UoN.bamGnut.42189	bZIP2	NA			
UoN.bamGnut.23966	zf-Di19	NA			
Common TFs					
UoN.bamGnut.39167	AUX_IAA	6			
UoN.bamGnut.43478	NAC3	0			
UoN.bamGnut.8752	MYB	NA			
UoN.bamGnut.11565	zf-FLZ	0			
UoN.bamGnut.52114	zf-rbx1	1			
UoN.bamGnut.40054	bHLH	0			

*(V° refers to the number of links of each TF node, in the merged network)

NA = Genes not found when networks were merged

6.2.6 Comparison between DipC and TN for transcript abundance

A selection of the genes that were significantly different between DipC and TN in terms of their differential expression under drought stress were chosen to investigate further the potential cause for the observed differences and whether they are due to the genomic differences between the genotypes or due to the FDR cut-off set to identify gene expression between control and drought conditions, despite that patterns

of expression being similar. The genes were chosen from each genotype based on their fold-change significance and their potential association with drought response (based on gene ontology). The TMM normalised CPM values were used to identify the transcript abundance for each gene under each treatment condition. This analysis will aid in an initial understanding of any genomic or expression difference between the two genotypes.

In many cases, the CPM values showed significance differences in transcript abundance between DipC and TN under drought stress for genes; GST (UoN.bamGnut.682 and UoN.bamGnut.3282), ABA_WDS (UoN.bamGnut.42895 (GO:0006950), Alternative oxidase (UoN.bamGnut.43131 (GO:0009916) and UoN.bamGnut.15773) and Aspartate proteases (UoN.bamGnut.22301 and UoN.bamGnut.5777) (see figure 6.13and table 6.13 for CPM values in each genotype). Whereas, genes such as XET (UoN.bamGnut.48794 (GO:0004553) and UoN.bamGnut.19756 (GO:0004553)), Aspartate proteases (UoN.bamGnut.22301) and Peroxidase (UoN.bamGnut.20798 and UoN.bamGnut.52570 (GO:0004601)) showed significance difference in transcript abundance between DipC and TN under control conditions (Figure 6.13and table 6.13)

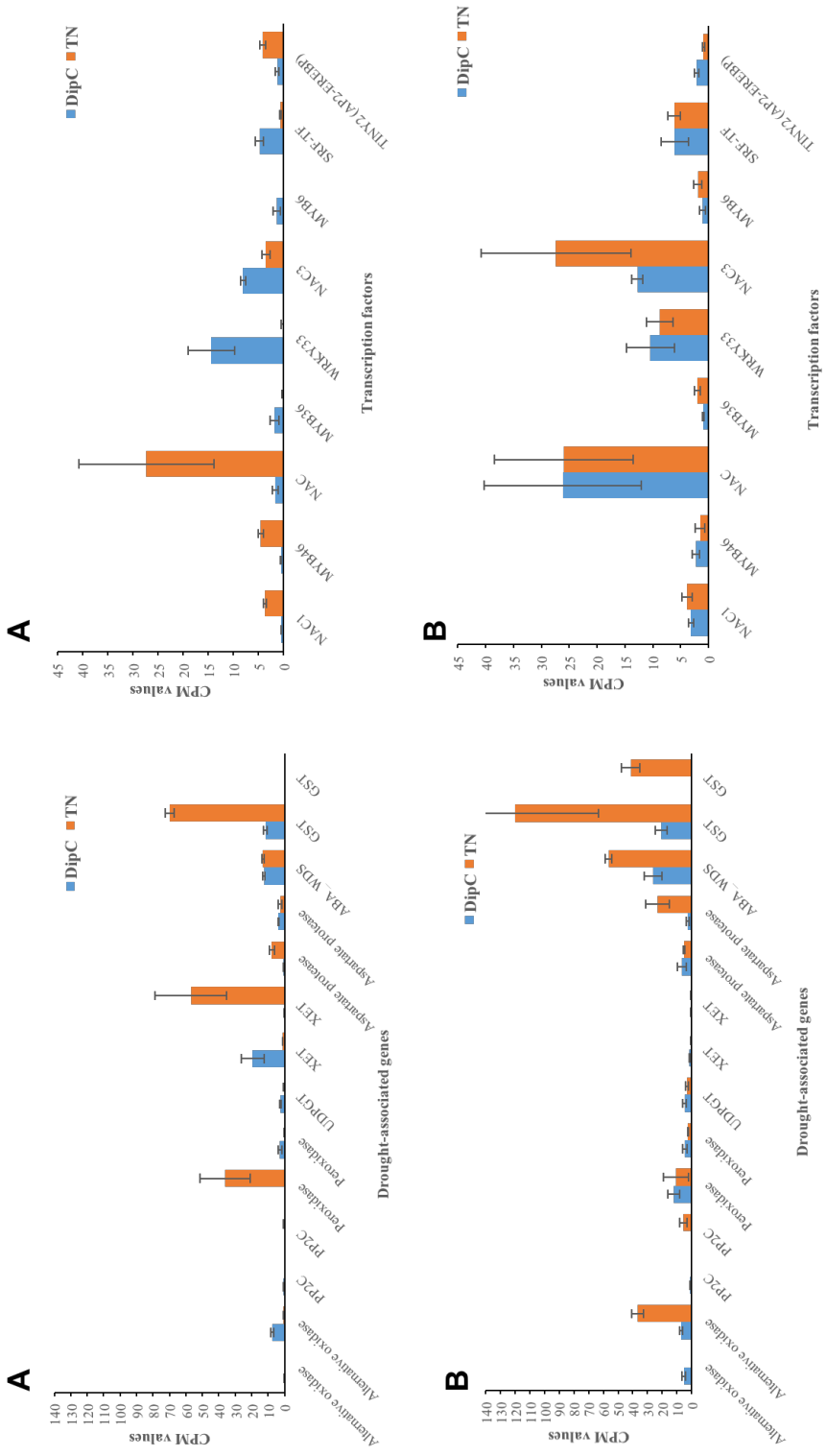

Figure 6.13: **Transcript abundance for differentially expressed genes identified in DipC and TN**. Columns [A] and [B] refer to transcript abundance for genes at control and drought condition, respectively. Error bars denote the standard error.

Table 6.13: **Counts per million (CPM) values for the drought-associated genes and transcription factors**

Gene_ID	Gene name	CPM values (Drought condition)		CPM values (Control condition)	
		DipC	TN	DipC	TN
UoN.bamGnut.15773	Alternative oxidase	5.35	0.11	0.04	0.04
UoN.bamGnut.43131	Alternative oxidase	7.12	36.90	7.36	0.71
UoN.bamGnut.29765	PP2C	1.02	0.00	0.80	0.00
UoN.bamGnut.37689	PP2C	0.00	5.59	0.00	0.40
UoN.bamGnut.20798	Peroxidase	12.25	10.66	0.00	36.13
UoN.bamGnut.52570	Peroxidase	4.63	2.39	2.86	0.20
UoN.bamGnut.43844	UDPGT	4.83	2.94	2.37	0.44
UoN.bamGnut.48794	XET	1.48	0.41	19.38	1.20
UoN.bamGnut.19756	XET	0.54	0.43	0.24	56.84
UoN.bamGnut.22301	Aspartate protease	6.54	5.15	0.72	7.78
UoN.bamGnut.5777	Aspartate protease	2.46	23.22	3.82	2.77
UoN.bamGnut.42895	ABA_WDS	26.14	56.40	12.53	13.18
UoN.bamGnut.3282	GST	20.78	119.93	11.70	69.85
UoN.bamGnut.682	GST	0.00	41.36	0.00	0.00
UoN.bamGnut.51896	NAC1	3.16	3.88	0.51	3.73
UoN.bamGnut.5476	MYB46	2.34	1.54	0.55	4.54
UoN.bamGnut.25328	NAC	26.16	25.94	1.70	27.28
UoN.bamGnut.39367	MYB36	0.96	2.04	1.83	0.23
UoN.bamGnut.7626	WRKY33	10.47	8.78	14.37	0.28
UoN.bamGnut.43478	NAC3	12.78	27.37	8.10	3.55
UoN.bamGnut.52834	MYB6	1.05	1.94	1.34	0.00
UoN.bamGnut.45251	SRF-TF	6.07	6.17	4.79	0.68
UoN.bamGnut.5656	TINY2 (AP2-EREBP)	2.11	0.94	1.25	4.15

Furthermore, some of the transcription factors (such as NAC1 (UoN.bamGnut .51896), MYB46 (UoN.bamGnut.5476), NAC (UoN.bamGnut.25328), MYB36 (UoN-.ba-mGnut.39367), WRKY33 (UoN.bamGnut.7626), NAC3 (UoN.bamGnut.43478), MYB6 (UoN.bamGnut.52834), SRF-TF (UoN.bamGnut.45251) and TINY2 (AP2-EREBP) (UoN.bamGnut.5656) that figured prominently in network analysis (based on vertex degree) in each genotype and were differentially expressed under drought stress were also chosen to investigate their transcript abundance under specific treatment conditions. The results showed that under drought conditions no significant

difference in transcript abundance for selected transcription factors were observed between the two genotypes (see figure 6.13 and table 6.13). Although, differences in gene expression under drought stress were observed between the two genotypes when differential expression analysis were performed using an FDR cut-off of < 0.05 and log FC ≥ 1. This indicates that these genes might be expressed in both genotypes under drought stress but owing to the cut-off set to identify differentially expressed genes, it was unable to be detected from gene expression analysis.

However, the differences observed in transcript abundance in the two genotypes under drought stress for the response genes indicates that the mechanism by which the two genotypes responding to drought might be different. However, validation through qRT-PCR is needed to further prove this hypothesis.

6.3 Discussion

6.3.1 Differential gene expression analysis

RNA-seq was used in this transcriptomic study of two genotypes of bambara groundnut (DipC and TN) under drought stress. RNA-seq was preferred over microarray in this study as it provides an unbiased assessment of all transcripts with a greater dynamic range, where quantification of expression levels would not be limited to a range of signal intensities as in microarray, even if there were a specific microarray available for bambara groundnut [318]. Furthermore, this study will also provide a platform to compare between RNA-seq and cross species microarray work done 4. Even though the experimental designs to study drought stress in bambara groundnut were different in both studies, the basic mechanisms on how bambara groundnut responds to drought would be expected to be relatively similar.

Sequencing was performed as 75bp paired end reads on a HiSeq4000 according to

Illumina specifications. The number of raw reads for all DipC (drought and control) and TN (drought and control) samples were ~ 25 million (Table 6.2). For comparing transcriptome across samples, all reads were combined into a single input in order to obtain a consolidated set of contigs (transcripts), followed by mapping back of the short reads for expression estimation. Once the reads are aligned back to the assembled transcriptome, statistical tests are applied to compare the counts of reads observed for each transcript across the different samples, and those transcripts observed to have significantly different representation by reads across samples are reported. Further analysis of the differentially expressed transcripts can reveal patterns of gene expression and yield insights into relationships among the investigated samples. This approach has been widely used for *de novo* transcriptomic studies [45].

To identify differentially expressed genes between treatments (control vs drought) and time points, the edgeR software was used. EdgeR is a robustly designed and well-documented software which accounts for biological variability of studies with only one or two replicate samples [235]. *De novo* transcriptome assembly was created using Trinity (version 2.2.0). Once the assembly was created, reads were mapped back to the assembled transcriptome using HISAT2 for expression estimation. StringTie was used to assemble read alignments into potential transcripts/contigs followed by merging/assembling these transcripts into a non-redundant set of transcripts and their associated gene-IDs across multiple RNA-Seq samples. Read counts were generated for genes using htseq-count to identify significantly differentially expressed genes using edgeR. EdgeR uses empirical Bayes estimation and exact test based on the negative binomial distribution with model-based scale normalization of sequence data, to identify statistically robust differentially expressed genes (DEGs), and this is reliable even for data of small sample sizes [235]. A two-step normalization method was applied to the count data; (i) First, to remove genes with low counts, genes must have at least

one count per million across all samples (ii) Secondly, TMM normalization was done to re-compute the library sizes for estimating relative RNA levels. TMM normalization is a preferred approach over RPKM, as TMM can account for a large dynamic range in RNA-seq data read count data and are able to control the false-positive rate while also maintaining the power to detect differentially expressed genes [319]. Filtering on RPKMs accounts for gene length and is found to be not effective at aligning the read count distributions across samples because attempting to correct for differences in gene length in a differential analysis has the effect of introducing a bias in the per-gene variances, in particular for lowly expressed genes [320]. FDR was set as a criterion for identifying significant differentially expressed genes over p-value as FDR calculates the percentage of false positives within significant results that was generated using p-values in multiple comparisons such as in RNA-seq. Therefore, this will reduce the incidence of rejecting a false hypothesis [321].

6.3.2 Gene ontology term enrichment annotation

Gene Ontology (GO) term enrichment annotation was used to get general insights into the biological processes, cellular components and molecular functions that were over-represented in DEGs. In particular between drought and control as well as between controls at different time points. To identify GO terms for bambara groundnut genes firstly, a reciprocal blast search was used to identify putative bambara groundnut orthologs of *Arabidopsis thaliana* and *Glycine max* genes. Once these putative orthologs were found, EnsemblPlants BioMart (http://plants.ensembl.org/biomart) was used to get the GO terms for the *Arabidopsis thaliana* and *Glycine max* genes. The hypergeometric test for over-representation of GO terms, was run on the differentially expressed genes at various conditions using GOstats for *Arabidopsis thaliana* and *Glycine max* separately. However it should be noted that as there were only 10283

and 11609 GO terms assigned out of 56,772 assembled bambara groundnut sequences when blasted against *Arabidopsis thaliana* and *Glycine max* respectively, exclusion of some DE genes may lead to under-representation or even total omission of biological processes and functions that could be of significance, in addition to the relative scarcity of information on bambara groundnut. It is possible that variations may arise between orthologous genes after the point at which the species diverged. Though it is most likely that orthologous genes maintained transcriptional responses associated with biological processes that are evolutionarily conserved between species [322]. GO terms identified in both genotypes were associated with *Arabidopsis thaliana* genes as it provided greater number of hits compared to *Glycine max*. This could mainly be because *Glycine max* is not being as comprehensive or extensively annotated as *Arabidopsis thaliana.* However the phylogenetic distance between *Glycine max* and bambara groundnut is smaller compared to *Arabidopsis thaliana* (20 MYA; [213]).

Significantly higher number of DEG observed under drought stress in TN compared to DipC at vegetative stage, which resulted in a greater enrichment of GO terms related to drought stress response in TN than in DipC. However, the drought stress imposed at reproductive stage showed DipC having higher number of GO terms associated with drought response than TN, even though the number of DEG were greater in TN compared to DipC. Interestingly, some GO terms enriched in DipC under control (irrigated) conditions at vegetative and reproductive are related to "secondary metabolic process" (GO:0019748), "hydrogen peroxide catabolic process" (GO:0042744), "response to external stimulus" (GO:0009605), "regulation of reactive oxygen species metabolic process" (GO:2000377), indicating its intrinsic tolerance to drought. These observations are in accord with the results obtained when high expression level of genes related drought stress were observed in control samples according to CPM values (Appendix 16-23). It has been observed in the drought study of sorghum (*Sorghum*

bicolor) where intrinsic drought tolerance of species was determined through enrichment of GO terms related to drought response in control conditions [297]. On the other hand, TN showed enrichment of some GO terms related drought response under control conditions at reproductive stage such as "response to abiotic stimulus" (GO:0009628) and "oxidation-reduction" (GO:0055114), indicating its intrinsic tolerance to drought. However it's surprising to see no GO terms were enriched for TN in control conditions related to drought at the vegetative stage, but again this could well be due to fact that some DEG did not have a GO term assigned to them when blasted against *Arabidopsis thaliana* genes as discussed previously.

Drought stress results in production of several reactive oxygen species (ROS) that causes oxidative stress [93, 312, 323]. Accumulation of antioxidants is a drought tolerant strategy against ROS employed by plants [93, 246, 296, 312, 313]. The antioxidant enzymes constitute the "first line of defense" against ROS and oxidative stress generated by different abiotic and biotic injuries in plants [323, 324]. In both genotypes, constitutive up-regulation of genes at vegetative stage associated with the enriched GO term "response to oxidative stress" (GO:0006979) "response to water deprivation" (GO:0009414), "response to hydrogen peroxide" (GO:0042542), response to oxidation-reduction process (GO:0055114) and "response to abiotic stimulus" (GO:0009628) led to increase drought tolerance trait in both genotypes, enabling speculation that the both genotypes had a constitutively higher basal expression of antioxidant genes at the vegetative stage. However, at the reproductive stage the genes associated with oxidative stress were down-regulated in DipC while TN showed continuous up-regulation of the genes related to oxidative stress (Appendix 36 and 37).

6.3.3 Function of genes differentially expressed under drought stress in relation to drought tolerance

Both genotypes were subjected to the same environmental conditions and to the same extent of drought stress by maintaining the field capacity (50-25%) for drought stressed plants and 75% field capacity for the control plants (see section 3.2.2). Although the intensity of drought stress applied was similar, the two genotypes responded differently; in TN a significantly higher number of DEGs was observed under drought stress than in DipC at both vegetative and reproductive stage.

Gene expression profiling provides a vast amount of transcriptional information that may help to understand molecular mechanism on how plants respond to drought stress [311, 325, 326]. There have been few drought studies carried out on bambara groundnut, but the molecular mechanisms of how it responds and adapts to drought is still under investigation. This study has carried out transcriptomic comparison in two genotypes of bambara groundnut; DipC and TN, in an attempt to identify potential genes conferring or associated with advantageous traits. Gene expression profiling of 36 samples (3 biological replicate per condition) was conducted. After analyzing the DEGs generated from edgeR, it was inferred that ABA synthesis, signal transduction, raffinose synthesis, accumulation of antioxidants, down-regulation of photosynthesis related genes, carbohydrate metabolism, cell-wall modification and transporters play important roles in bambara groundnut under drought stress.

Classically, there are three main responses to drought exhibited by plants, escape, avoidance and tolerance (and, in practice, often a combination of all three). Drought escape is described as the ability of plants to complete their growth cycle and reach maturity before drought-stress develops to damaging levels [84]. Drought avoidance is demonstrated by crop species, which are able to maintain high water potential in

the plant by minimising water loss and maximising water uptake under drought conditions. Mechanisms of avoidance include improved root traits, for greater extraction of soil moisture, decreased stomatal conductance, decreased radiation absorption and decreased leaf area for minimal water loss [87]. Drought tolerance allows plants to survive through water-use efficiency, i.e., performing all biological, molecular and cellular functions with minimal water. Plants with drought tolerance mechanisms are able to maintain their cell turgor through osmotic adjustment, which in turn will contribute to maintaining stomatal opening, leaf expansion and photosynthesis throughout the drought period [90].

It is known from other studies that plant hormones and other signal molecules are important in the drought response in plants [91, 311]. The most important hormone involved in drought response is ABA [91, 311]. ABA is a major molecule facilitating signal transduction during drought stress response. Synthesized or transported ABA is perceived by a receptor complex, which consists of PYR (PYRABACTIN RESISTANCE)/ PYL (PYR1-LIKE)/RCAR (REGULATORY COMPONENT OF ABA RESPONSE), PP2C (protein phosphatase 2C), and SnRK2 (sucrose non-fermenting 1-related protein kinase 2) [99]. In this study, some genes involved in ABA signal transduction such as genes in the PP2C and ABA_WDS family and some ABA induced family of transcription factors; MYB, WRKY, bZIP and NAC were up-regulated under drought stress. ABA_WDS is a family of plant proteins induced by water deficit stress (WDS) [327] or ABA stress and ripening [328]. ABA_WDS was induced in both DipC and TN under drought at the vegetative stage. In the absence of ABA, PP2Cs repress the ABA signaling pathway by dephosphorylation-triggered inactivation of SnRK2s. In the presence of ABA, ABA-bound PYL/PYR/RCARs recognize and bind to PP2Cs, thereby releasing SnRK2s from PP2C-dependent negative regulation. The activated SnRKs phosphorylate downstream proteins including AREB/ABF (ABA-

responsive cis-element binding protein/ABA-responsive cis-element binding factor) transcription factors [99]. The AREB/ABF transcription factors have a bZIP domain and four conserved domains containing SnRK2 phosphorylation sites. AREB/ABFs function as master transcriptional activators regulating ABRE-dependent gene expression in ABA signaling under drought stress conditions [99]. These results indicate ABA signaling is likely to play an important role in both genotypes in response to drought stress. In addition, there are some ABA-independent drought response pathways which includes genes in the HD-ZIP family, DREB2 and some NAC TF family genes [99]. DREB2 proteins are the members of AP2/ERF family of TFs [99]. In this study, several transcription factors belonging to AP2/ERF and NAC family were induced under drought stress in both DipC and TN at the vegetative stage indicating the components of ABA-independent pathways participating in the early growth stage process of drought adaptation in both genotypes.

Apart from ABA, among other plant hormones such as auxin and cytokinin binding proteins were differentially expressed under drought stress at vegetative stage in both genotypes. It is known that hormones do not act in isolation but are interrelated by synergistic or antagonistic cross-talk which enables them to modulate each other's biosynthesis or responses [329]. In this study, cytokinin binding protein was repressed under drought stress at the vegetative stage in DipC but up-regulated in TN. Cytokinin is an antagonist to ABA whose expression level decreases with the plants exposure to drought stress [329]. One role of auxin is to enhance the expression of LEA (late embryogenesis abundant) genes, which are correlated with the increased drought tolerance in plants [329]. In this study, the expression of auxin related genes were induced in both genotypes at vegetative and reproductive stage indicating their role in drought response.

Among the processes taking place downstream of this transcriptional regulatory

network, osmoprotectants were induced under drought stress in both genotypes at the vegetative stage including raffinose synthase, beta-fructofuranosidase, trehalose phosphatase and aspartate proteases. Aspartate proteases is known to regulate stomatal closure through ABA signaling in guard cells [330], hence leading to decreased water loss. However, at the reproductive stage fewer osmoprotectants were induced in response to drought stress in both genotypes suggesting a strong reaction to drought stress in the early stage of development. The role of osmoprotectants in the plant defense against drought stress is by contributing towards water uptake and maintenance and membrane protection [306]. Previous studies have reported high expression levels of raffinose synthase, trehalose phosphatase, beta-fructofuranosidase and aspartate proteases in response to drought stress [93, 99, 152, 155, 306, 308, 309]. From previous studies on bambara groundnut under drought stress, high expression level of proline content has been observed [90]. In this study, genes involved in proline synthesis were not significantly up-regulated under drought in both genotypes. Results from [90] imply that proline may play little part in osmotic adjustment in bambara groundnut but due to the different level of drought stress applied in this study it has not been found to be induced. Raffinose synthase was highly expressed under drought stress in DipC and TN at both developmental stages indicating it may have major impacts on osmotic adjustment in bambara groundnut under drought stress. Furthermore, drought stress causes the production of reactive oxygen species (ROS), which leads to cell damage and oxidative stress. Antioxidants enzymes are a major components that plays a vital role in removing ROS [93]. Oxidative damage in the plant tissue is alleviated by action of enzymes such as superoxide dismutase, peroxidase, ascorbate peroxidase, catalase and glutathione S-transferases [93]. In this study, several antioxidants including genes in the peroxidase, glutaredoxin, GST, glutathione S-transferases and ferritin family were highly induced under drought stress in both genotypes at vegetative stage indicating

their role in ROS scavenging at early stage of development in bambara groundnut. However, the antioxidants differentially expressed in DipC at the reproductive stage were highly repressed indicating the sensitivity of DipC in response to drought stress at flowering stage, previously only characterised by physiological data (see chapter 3.2) TN, on the other hand was not affected by drought stress and continues to produce antioxidants as high expression of genes related to oxidative stress was observed even at reproductive stage indicating its tolerance to drought stress (Appendix 37 and Table 6.7).

Furthermore, most of the genes related to photosynthesis and Calvin cycle were significantly down-regulated, whereas genes related to glycolysis were up-regulated in both genotypes at vegetative stage (Appendix 27-28 and Table 6.4 and 6.5). This results could explain the observations obtained from the physiology data where drought plants in both genotypes had reduced photosynthesis compared to control plants (see figure 5.2 in chapter 3.2). Similar results have been observed in physic nut (*Jatropha curcas L.*) when subjected to early drought stress [306] and other previous studies [93, 288]. Drought causes changes in photosynthetic pigments and components, damaged photosynthetic apparatus and reduced activities of Calvin cycle enzymes [93, 285, 331, 332], which are important causes of reduced crop yield. However in this study, up-regulation of carbonic anhydrase (Pro_CA) in both genotypes at vegetative stage was observed which catalyses reversible hydration of CO_2 and helps the plants to compensate for the lack of water and CO_2 in under stress conditions [287]. In a study conducted by [287] who showed that higher activity of carbonic anhydrase in tomato (*Solanum lycopersicum*) helped the plant to retain higher photosynthetic capacity, higher water use efficiency (WUE), and higher water potential without cellular damage. In this study, even though there was a reduction in photosynthesis from the physiological point of view, plants still managed to photosynthesise (see fig. 5.2 in

chapter 5). This may be due to the fact of high expression level of carbonic anhydrase observed in both genotypes under drought stress at vegetative stage, indicating the plant drought tolerance at early stage of plant development. However, reduced expression of carbonic anhydrase was observed in both genotypes at reproductive stage which indicates the sensitivity of both genotypes to photosynthesise under drought stress at reproductive stage.

Furthermore, genes associated with cell wall maintenance and synthesis including genes in the UDP- Glucosyl Transferase enzyme (UDPGT) family, cellulose synthase and fascilin were strongly up-regulated in both genotypes at vegetative and reproductive stage. The UDPGT coding gene was found to be highly induced under drought stress in both genotypes. This enzyme is known to be involved in the formation of β-glucans, which are thought to be involved in cell wall formation [326]. In addition, other genes such as LEA were highly induced under drought stress in both genotypes at the vegetative and reproductive stage. LEA genes plays role in cellular protection against drought stress in plants [99]. Overexpressing of LEA proteins leads to enhanced drought tolerance in rice [333].

The results indicate that drought treatment affected more TN genes than DipC at both developmental stages. However, the above observations show that the two genotypes appear to be very similar in terms of genotype and general transcriptional behavior. This might be expected as DipC and TN are highly likely to be adapted to the similar drought conditions of their countries of origin, Botswana and Mali respectively. However, when the sets of differentially expressed genes are compared at both developmental stages, there is very little overlap between the two genotypes (see table 6.3 and 6.6). These results are in accord with the microarray cross species study done on the genotypes under drought stress (see chapter 4). The two genotypes with very similar genomes have adapted to achieve the same drought response traits (transcrip-

tional, cell-wall modification, photosynthesis, transporters, hormone signaling, signal transduction, osmoprotection, oxidative stress) through largely different sets of effector genes. Furthermore, both genotypes have shown greater number of genes related to drought response traits at vegetative stage compared to reproductive stage. This study also showed that many potential drought-responsive genes are expressed even under irrigated conditions in both genotypes, perhaps priming the plants for drier times ahead, previously observed in the other study (see chapter 4).

Several transcription factors that seem likely to play a role in the bambara groundnut drought response. Common to both genotypes at vegetative stage are *zf-RING_2*, *MYB86*, *AP2-ERF* domain, *WRI1* (AP2-EREBP), *RAD-like 6* (SANT/MYB TF) and *zf-Di19*. *WRI1* TF is known to be involved in glycolysis metabolism [317] and *MYB86* is involved in negative regulation of transcription [334]. On the other hand, common TF to both genotypes at reproductive stage are *AUX_IAA*, *HOX* and *NAC3*. Overexpressing the *NAC3* gene in *Arabadopsis thaliana* revealed that several stress inducible genes were up-regulated in transgenic plants, hence the plants showed significantly greater drought tolerance in physiological terms [335]. On top of the common TFs at vegetative stage, DipC showed bigger response with changes to *NAC1*. It plays a role in regulating protein phosphatase18 gene which modulates drought and oxidative stress tolerance through abscisic acid-independent reactive oxygen species scavenging in rice (*oryza sativa*) [336] and is the most highly linked TF node in the coexpression networks. For DipC, the network also reveals the importance of *zf-CCHC* and *MYB46*. TF *MYB46* is a key player in the transcriptional network involved in the secondary wall biosynthesis in *Arabadopsis thaliana* [337]. TN, on the other hand, revealed a strong role of *EMB1444*, *AUX_IAA*, *WRKY*, *HOX*, *zf-RanBP*, *WRKY33*, *MYB36*, *zf-C2H2_jaz*, *zf-C2H2_jaz* and *zf-Nse* at the vegetative stage. *MYB36* plays an important role in casparian strip formation which is vital in regulating water and

nutrient use efficiencies and enhanced resistance to abiotic stresses [338]. *WRKY33* is induced in response to salicylic acid or paraquat herbicide which generates activated oxygen species in exposed cells and it is known to activate several stress-related downstream genes, increase germination rates, and promoted root growth in *Arabidopsis thaliana* under drought stress [339]. On the other hand, at the reproductive stage, DipC showed the importance of *SHN3* (AP2-EREBP), *zf-C2H2_jaz*, *MYB46*, *BRX* and *AUX_IAA*. *SHN3* are key activator for cuticular wax biosynthesis [340] and *BRX* TF is known to regulate the extent of cell proliferation and elongation in the growth zone of the root in plants [341, 342]. TN, on the other hand, showed importance to *MYB*, *NAC25*, *MYB2* at reproductive stage. *NAC25* was induced in the shoots of *Arabidopsis thaliana* in response to drought stress [298] and *MYB2* functions as transcriptional activators in ABA inducible gene expression under drought stress in plants [343]. Furthermore, *MYB46* showed importance in DipC at both developmental stages as it was one of highly linked TF node in the co-expression network, whereas for TN, *AUX_IAA* (UoN.bamGnut.39167) showed importance at both stages.

6.4 Conclusion and future perspectives

RNA-seq has been proven to be a robust method for transcriptomic study in investigating complex traits in two genotypes of bambara groundnut. This transcriptomic study has revealed genes of drought tolerance in both genotypes. The first indication of drought response in the DipC genotype was shown by high expression level of "secondary metabolic process" (GO:0019748), "hydrogen peroxide catabolic process" (GO:0042744), "response to external stimulus" (GO:0009605) and "regulation of reactive oxygen species metabolic process" (GO:2000377) activities under irrigated (control) conditions. Whereas for TN, high expression level of "oxidation-reduction" (GO:0055114) and "response to abiotic stimulus" (GO:0009628) activity under irri-

gated conditions was observed. Clearly, if the plant has already activated part of the drought response under control conditions, it could have multiple effects which may lead them to greater root growth and other avoidance traits for later dry periods, hence escaping the worst drought conditions. In addition, differential expression of several genes associated with drought response traits under drought conditions in both genotypes was observed as well. However, in terms of differential expression, although the transcription factors and drought-response genes in both genotypes were largely different, they may achieve the same effect in terms of survival under drought conditions as genes differentially expressed in both genotypes were largely coding for similar drought response traits. Differential expression of genes related to photosynthesis and stomatal closure in both genotypes could explain observed reduction in photosynthesis and stomatal conductance from the physiology results (see fig. 5.2 and 5.3in chapter 3.2). Differential expression of some well known drought-related transcriptions factors in DipC (especially *NAC1* and *MYB46*) and TN (especially *MYB36*, *MYB2*, *NAC25* and *AUX_IAA*) was also observed under drought stress.

To ensure survival under minimal water supply, managing the effects of drought stress through the cultivation of the most appropriate genotypes together with adjustment of agronomic practices (sowing time, plant density and soil management) is necessary. This is to ensure that the sensitive growth stages in plants do not occur at the time when the likelihood of drought stress is most severe. In this study, the gene expression data has provided clues about gene function, but it does not reveal what exactly these genes are doing inside a cell. Genetics could provide a solution to this problem because mutants that lack a particular gene may quickly reveal the function of the protein that it encodes. To facilitate such studies of gene function, the coding sequence of a gene and its regulatory regions can be engineered (either by overexpressing or gene knockout of the gene of interest) to change the functional

properties of the protein product [344]. This will also prove if the identified gene is associated with the trait of interest.

Chapter 7

Identification of cross species syntenic locations of drought-related genes

Chapter 4 and 6 identified several candidate drought-related genes from Xspecies microarray and RNA-sequencing (RNA-seq) approaches. This chapter aims to identify the syntenic locations of these genes from a cross-species approach due to the lack of a bambara groundnut genome sequence. So far, the well annotated common bean physical genome (*Phaseolus vulgaris*) has been shown to have good conserved marker order (genetic versus physical location) based on marker order alignment using two bambara groundnut genetic maps derived from two different crosses [345]. A total of 15 and 77 genes were chosen from XSpecies microarray and RNA-seq experiments, respectively based on the fold change significance, network analysis (based on vertex degree) and their potential association with drought response (based on gene ontology)).

Chai et al. [24], studied the effect of drought stress on the morpho-physiological characteristics for the F_5 segregating population derived from TN x DipC cross. This study was followed by identifying quantitative trait loci (QTL) for drought-related traits (including stomatal density, leaf carbon (Delta C^{13}) isotope analysis (CID), stomatal conductance, relative water content, and leaf (Delta N^{15}) isotope analysis

(d15N)) on an Gene Expression Marker-based (GEM) genetic map based on a Xspecies microarray approach [238]

Segregating markers were developed from the F_5 segregating population using gene expression levels in leaf tissue from individual lines of the cross after a six weeks mild drought treatment. An initial linkage analysis at logarithm of the odds (LOD) score > 3 for grouping generated 13 linkage groups (LG) with 218 GEMs, spanning 982.7 cM of the bambara groundnut genome. A total of six QTLs (five significant and one putative) were mapped on LG1, LG2, LG4 and LG5 in the drought treatment population for drought-related traits (stomatal conductance, stomatal density and CID). With the lack of whole genome sequence for bambara groundnut, common bean, belonging to the same legume family of Fabaceae as bambara groundnut was used to identify the syntenic locations of these drought treatment genes. Common bean genome was chosen as it is well annotated and having a close evolutionary relationship among these two legumes which the divergence time between *Phaseolus* and *Vigna* has been estimated to be 5.5 million years ago (MYA) [213]. The putative conserved synteny blocks identified across legume genomes will potentially help in the identification of the location of genes underlying QTL involved in controlling drought traits in bambara groundnut, hence could build a reasonable platform for marker assisted selection (MAS).

7.1 Results and Discussions

7.1.1 Syntenic locations of bambara groundnut genes of interest in common bean genome

A total of 15 genes detected using the XSpecies microarray used in Experiment 1 and 77 from RNA-seq from Experiment 2 were mapped to the common bean genome. The selected genes were chosen because of their potential association with drought response

and significance in fold-change and network analysis (based on vertex degrees). The syntenic locations of these genes in the common bean genome are listed in Table 7.1.

Table 7.1: **Drought-related gene locations in common bean**

Chromosome number	**RNA-seq/Microarray probe ID**	**Location (Mbp)**	**Common bean gene ID**	**Annotation**
Chromosome 1	UoN.bamGnut.17348	1,192,760	Phvul.001G014300	RAD-like 6 (MYB TF)
	Gma.16733.1.S1_at	4,312,300	Phvul.001G042100	WRKY40
	UoN.bamGnut.37689	10,508,320	Phvul.001G075400	PP2C
	UoN.bamGnut.54259	34,549,000	Phvul.001G123000	Beta-fructofuranosidase 31
	GmaAffx.93094.1.S1_s_at	44,153,800	Phvul.001G177800	PAL 1
	UoN.bamGnut.46640	46,879,760	Phvul.001G203500	Glutaredoxin
	UoN.bamGnut.43646	47,035,660	Phvul.001G205900	Tryptophan-rich sensory protein (TspO)
	UoN.bamGnut.51895	47,851,080	Phvul.001G215200	MYB 2
	UoN.bamGnut.25299	48,165,000	Phvul.001G219000	MYB
Chromosome 2	UoN.bamGnut.11042	6,166,040	Phvul.002G058900	zf-C2H2_jaz
	UoN.bamGnut.9228	21,315,150	Phvul.002G105800	zf-RING_2
	UoN.bamGnut.43131	25,631,780	Phvul.002G127100	Alternative oxidase
	UoN.bamGnut.15773	25,631,780	Phvul.002G127100	Alternative oxidase
	GmaAffx.60283.1.S1_at	30,055,100	Phvul.002G158800	brassinosteroid-responsive RING-H2
	UoN.bamGnut.42702	34,413,860	Phvul.002G196800	WRKY 28
Chromosome 3	UoN.bamGnut.6146	791,560	Phvul.003G008100	zf-C2H2_jaz
	UoN.bamGnut.53196	1,762,800	Phvul.003G019900	zf-RanBP
	UoN.bamGnut.31501	32,327,820	Phvul.003G132600	zf-RING_2
	UoN.bamGnut.52570	33,941,380	Phvul.003G143400	Peroxidase
	UoN.bamGnut.10038	41,844,080	Phvul.003G203900	MYB 86
	Gma.1546.1.S1_a_at	43,439,240	Phvul.003G217900	Beta-fructofuranosidase
	UoN.bamGnut.31777	44,799,400	Phvul.003G227800	Ras of Complex, (DAPkinase domain)
Chromosome 4	UoN.bamGnut.19756	8,350,860	Phvul.004G062400	Xyloglucan endo-transglycosylase
	UoN.bamGnut.3314	25,700,400	Phvul.004G094600	ABA_WDS
	UoN.bamGnut.48031	44,048,400	Phvul.004G158800	Dehydrin

Chromosome 5	UoN.bamGnut.8752	1,605,220	Phvul.005G018500	MYB
	UoN.bamGnut.52834	23,464,740	Phvul.005G087400	MYB DNA_binding_6
	UoN.bamGnut.5476	37,515,100	Phvul.005G147500	MYB 46
	UoN.bamGnut.38337	39,834,920	Phvul.005G174800	LEA-4
Chromosome 6	GmaAffx.67871.1.S1_at	22,612,200	Phvul.006G110100	UBC-2
	UoN.bamGnut.38586	28,885,500	Phvul.006G178800	zf-C3HC4_3
	UoN.bamGnut.39304	29,706,960	Phvul.006G188900	NAC 25 (NAM)
	UoN.bamGnut.27002	30,674,140	Phvul.006G202600	Carbonic anhydrase (Pro_CA)
Chromosome 7	UoN.bamGnut.42895	7,717,240	Phvul.007G080400	ABA_WDS
	UoN.bamGnut.20798	8,159,240	Phvul.007G082700	Peroxidase
	UoN.bamGnut.30177	11,669,700	Phvul.007G102900	SANT/MYB domain
	GmaAffx.92679.1.S1_s_at	41,297,500	Phvul.007G176700	ATAUX2-11
	UoN.bamGnut.29777	48,049,920	Phvul.007G240600	Superoxide dismutase (Sod_Fe_C)
	UoN.bamGnut.8660	48,154,280	Phvul.007G241800	AP2-ERF domain
	UoN.bamGnut.2786	50,605,580	Phvul.007G267800	zf-Nse
Chromosome 8	UoN.bamGnut.38339	1,271,580	Phvul.008G014600	zf-C2H2_6
	GmaAffx.45249.1.S1_at	1,956,300	Phvul.008G022800	CONSTANS-like 1
	GmaAffx.84566.2.S1_at	3,542,560	Phvul.008G041500	MYB60
	GmaAffx.86517.1.S1_at	9,172,900	Phvul.008G090000	AGL83
	UoN.bamGnut.36919	9,621,500	Phvul.008G093700	Ferritin
	GmaAffx.33796.3.S1_at	58,041,460	Phvul.008G270400	Zinc-finger like C2H2 (IIIA)
Chromosome 9	UoN.bamGnut.41951	2,247,080	Phvul.009G014100	zf-B_box
	UoN.bamGnut.25416	14,412,400	Phvul.009G093600	SHN3 (AP2-EREBP)
	UoN.bamGnut.5656	16,431,700	Phvul.009G109600	TINY2 (AP2-EREBP)
	UoN.bamGnut.12291	17,795,300	Phvul.009G119900	MYB
	UoN.bamGnut.51896	22,688,560	Phvul.009G156300	NAC domain 1 (NAM)
	UoN.bamGnut.48677	30,314,820	Phvul.009G205100	zf-UDP
	UoN.bamGnut.44892	36,176,680	Phvul.009G248700	translationally controlled tumor protein

Chromosome 10	UoN.bamGnut.40849	38,454,640	Phvul.010G117200	HOX
	UoN.bamGnut.45698	41,578,810	Phvul.010G144300	Acid phosphatase
	UoN.bamGnut.39167	41,833,100	Phvul.010G147200	AUX_IAA
Chromosome 11	UoN.bamGnut.45251	436,560	Phvul.011G005800	SRF-TF (AGAMOUS-LIKE 12)
	UoN.bamGnut.19689	2,264,360	Phvul.011G028100	zf-CCHC
	Gma.17248.1.A1_at	4,882,660	Phvul.011G056900	JMJD5
	GmaAffx.9286.1.S1_s_at	5,372,820	Phvul.011G062100	MYB
	Gma.6670.1.S1_at	6,195,260	Phvul.011G070600	PRR-7
	UoN.bamGnut.39367	8,103,220	Phvul.011G084500	MYB 36
	UoN.bamGnut.47465	12,854,200	Phvul.011G107700	Beta-fructofuranosidase 14

7.1.2 Comparing the locations of QTL for CID and stomatal density/leaf area with physical positions in the common bean genome and bambara groundnut gene conserved synteny positions

A total of six QTLs (five significant and one putative for stomatal conductance, CID and stomatal density drought traits) were identified in drought treatment population of F_5 segregating population derived from TN x DipC cross for drought-related traits using GEM map [238]. Among these, significant QTL for stomatal density/leaf area and leaf carbon 13 isotope analysis (CID) were selected and their identified syntenic blocks to common bean (Pv) chromosomes are shown in Figure 7.1.

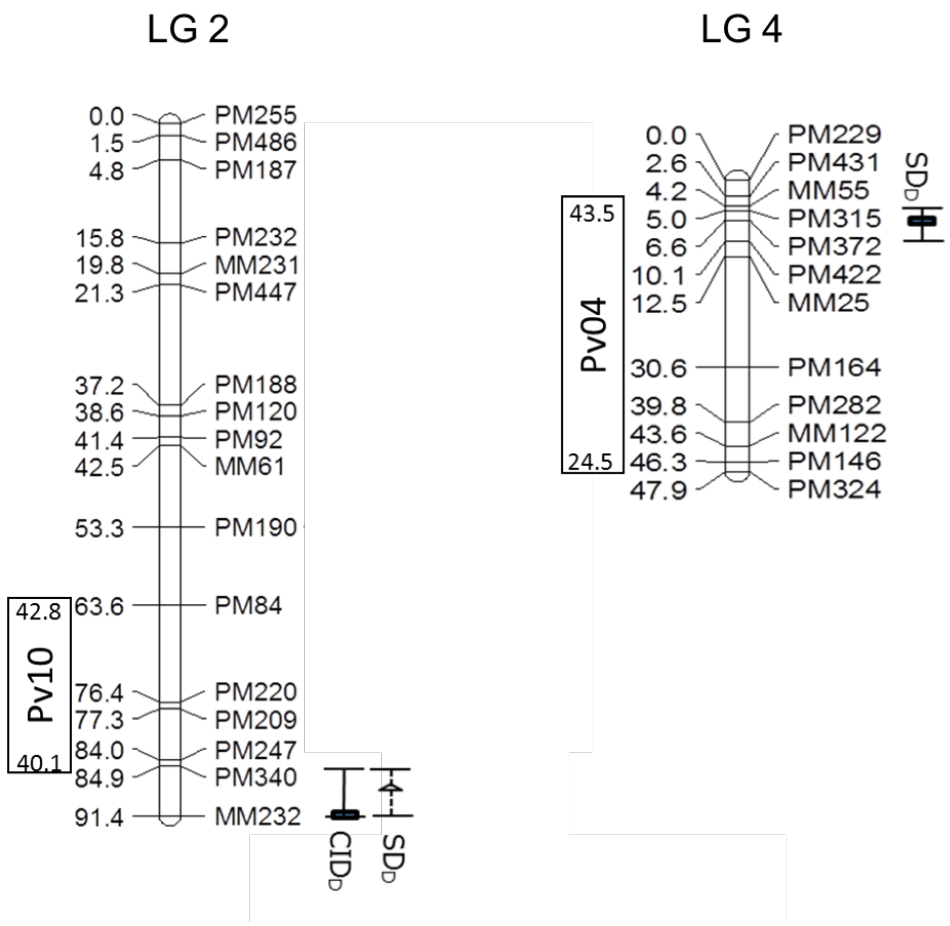

Figure 7.1: **Identified syntenic blocks to common bean chromosomes of stomatal density/leaf area and CID QTLs.** Significant stomatal density and CID QTL on LG2 and LG4 of GEM map. The syntenic blocks of common bean chromosomal locations are given in Mbp.

Both CID and stomatal density QTL in response to drought had syntenic blocks with Pv04 and Pv10 chromosomes. Comparing with the syntenic location of genes of interest in Table 1, it can be observed that all three genes identified from RNA-seq; HOX (UoN.bamGnut.40849; Homeobox family transcription factor); AUX_IAA (AUX_IAA family transcription factor; UoN.bamGnut.39167) and acid phosphatase (HAD super- family) (UoN.bamGnut.45698) are located at the neighboring regions

of the CID and stomatal density drought QTL of Tiga Nicuru x DipC F_5 population. In addition a gene identified in the transcriptomic analysis that was found near to the flanking markers of the QTLs associated with stomatal density/leaf area (87.9 cM) trait at Linkage Group 4 was a an homologue of dehydrin gene family (UoN.bamGnut.48031), although likely to be outside of the immediate confidence interval for the QTL.

The HOX and AUX_IAA transcription factors were induced under drought stress in both genotypes (DipC and TN). In TN, AUX_IAA was induced when drought stress was imposed at the vegetative stage (log fold change of 2.90) and the reproductive stage (log fold change of 2.38). Whereas, for DipC, it was up-regulated only at the reproductive stage (log fold change of 2.32). On the other hand, HOX was up-regulated under drought stress in TN at the vegetative stage (log fold change if 4.62) and the reproductive stage (log fold change of 2.32). DipC, on the other hand, showed high expression of HOX at the reproductive stage (log fold change of 4.62).

HOX and AUX_IAA family of protein transcription factors are known to function in a wide variety of developmental processes and abiotic stress responses in plants [99]. In addition, acid phosphatase was also highly expressed in both genotypes at the vegetative stage (log fold change of 4.08 and 4.89 for DipC and TN, respectively). [346] reported high induction of acid phosphatase in pigweed leaves under drought stress which enabled the plant to resist drought stress by enhancing the activity of protective enzymes such as superoxide dismutase (SOD), peroxide dismutase (POD) and hydrogen peroxidase (CAT). On the other hand, the homologue of dehydrin gene family was highly expressed under drought stress in TN at the vegetative stage (log fold change of 8.11) and the reproductive stage (log fold change of 5.22). However, it was not found to be differentially expressed in DipC. Dehydrin is a multi-family of proteins present in plants that is produced in response to drought stress [93]. Dehydrin

are identified as the genetic basis of drought tolerance in plants which includes cellular protection during osmotic stress and response to ABA [93, 347].

Table 7.2: Physical location of drought-related genes near the QTL identified for drought traits

Linkage group	Drought trait	Position (cM)	LOD score	Flanking markers	Nearest marker	Physical location of flanking markers in common bean (Mbp)	Common bean chromosome number	Drought-related genes location near the flanking markers (Mbp)	Gene name
2	Stomatal density/leaf area (D)	87.894	2.54[P]	PM340 (84.894cM) - MM232 (91.441 cM)	MM232 (91.441 cM)	40,156,820	10	38,454,640 - 41,833,100	HOX, Acid phosphatase and AUX_IAA
	CID	91.441	3.46	PM340 (84.894cM)	MM232 (91.4cM)	40,156,820			
4	Stomatal density/leaf area (D)	6.649	3.52	PM315 (4.967 cM) - PM422 (10.126 cM)	PM372 (6.649cM)	42,894,080-41,145,938	4	44,048,400	Dehydrin

p: putative QTLs where LOD score is lower than 3.

7.2 Conclusion

The preliminary results overlaying the location of QTLs related to drought traits (stomatal density/leaf area and CID) onto a sequenced genome of common bean assisted in identifying whether any of the candidate drought-related genes identified in Chapter 4 and 6 might lie within the confidence intervals of the QTL and could represent candidate genes for further investigation. However, only 4 genes could be found near to the QTLs underlying the drought traits. The reason for this could be because of the limited significant QTLs generated from GEM map for drought-related traits. Nevertheless, the initial results suggesting that some of the locations of genes identified in XSpecies microarray and RNA-seq experiments could underly QTL involved in controlling drought traits in bambara groundnut has produced a route to explore further.

Chapter 8

General Conclusion

8.1 Challenges and potential approach for food security

Currently only 20 plant species comprise 90% of the world's food calories [18]. These include rice, maize and wheat [348, 349]. Climate change, along with a predicted increase in pests and diseases present a potential threat that could drastically effect all crop's growth and development. Hence the need for crop diversification is important in order to maintain and, ideally, increase the food supply, while improving the nutritional content of food available to the world population as well as to avoid dependence on a limited number of plant species for global food and nutritional security. Along with climate change, the loss of crop diversity is also related to the intensification of agriculture and the growing of cash (commodity) crops [350]. It has been estimated that about 75% of the original varieties of crop species have been lost since 1900 [351]. Since the launch of the Green Revolution in the 1960s, cultivation of single, high-yielding varieties have been preferred over traditional landraces and an unintended consequence of this agricultural practice has been the narrowing of the genetic, both within species and in terms of numbers of cultivated species, raising the chance of more major loss events from pests and diseases in the future [12].

The heavy dependence of global food security on major crops is a concern for the future for food supply, as yield gains from these major crops may not be enough to sustain the estimated nine billion people on the planet by 2050 [17]. It is predicted that crop yields must increase by around 70% to feed the estimated 9 billion people [352], although reducing food waste at all stages of the supply chain is clearly one area where much can be achieved. Currently there are around 868 million people suffering from hunger and malnutrition which includes, 35% malnourished children and over 2 billion people with micronutrient deficiencies [353–355].

Regarding these issues, underutilised crops could be part of a solution for a more resilient and diversified agricultural system [19, 61, 62]. However, there are many obstacles that inhibit comprehensive research and development in underutilised crops. These include; lack of large-scale characterized collections of germplasm and problems with access [220], competition with commodity crops, lack of financial support from national governments and international breeding companies and insufficient market demand for unknown crops.

The application of conventional and molecular breeding using biotechnology is an option for crop improvement to select plants containing advantageous traits such as drought tolerance, resistance to pests, diseases, water-logging and eroded soils. With the availability of advanced expression analysis techniques such as next generation sequencing (NGS) and microarrays, it has become possible to carry out extensive gene expression analysis to provide a foundation to understand gene expression, biological responses and tolerance traits [194]. In addition, with the help of genomic and transcriptomic approaches, sequence data and trait knowledge derived from major and model plant species will help in the development of new molecular markers for advantages traits and gene discovery in underutilised crops, leading to crop improvement. With the availability of genomic resources and the completion of reference genome se-

quences of legume crops, such as *Medicago truncatula* [78], common bean [79], soybean [80] and cowpea [81], it is now possible to dissect information and transfer genomic and transcriptomic data to other underutilised legume crops such as bambara groundnut.

The present study aimed to investigate and evaluate the transcriptome of bambara groundnut under drought stress using microarray XSpecies and RNA-sequencing (RNA-seq) approaches by utilising data, resources and approaches derived from major crops and model plants. Identified genes and expression patterns from this study will assist in understanding mechanisms underlying the drought tolerance in bambara groundnut which enables it to grow under semi-arid conditions. Furthermore, localisation of drought-related genes (identified from microarray XSpecies and RNA-seq approaches) underlying quantitative trait loci (QTL) involved in controlling drought traits could build a foundation for marker-assisted selection (MAS) for breeding purposes.

8.2 Effect of drought stress in bambara groundnut

To perform the XSpecies approach, a close association between studied crop species and model/major plant species is critical. Sequence divergence between species will result in inefficient hybridisation of certain transcripts to the probes that would lead to the production of background noise which could be an obstacle in data analysis or even lead to the complete loss of signal. The XSpecies microarray approach was applied in bambara groundnut using the soybean affymetrix chip. Despite the sequences available for soybean not being as comprehensive or as extensive annotated as *Arabidopsis thaliana* and rice, the phylogenetic distance between the soybean and bambara groundnut is smaller (20 MYA [213, 222]. In addition, soybean belongs to the same legume family "Fabaceae" as bambara groundnut and the phylogenetic distance between bambara groundnut and soybean is the smallest among the legume species

whose microarray affymetrix chip was available at the time the experimental data was generated, such as *Medicago truncatula* and Lotus [213]. However, the duplication of the soybean genome since evolutionary divergence of the two species (2n = 2x = 22 for bambara groundnut compared with 2n = 2x = 40 for soybean) creates complications.

In addition to the analysis of plant at a single development stage (pod development) using the XSpecies approach, gene expression was extended to cover two other development stages (vegetative and reproductive) with a more severe drought stress imposed in Experiment 2 using an RNA-seq approach to compare between parental genotypes DipC and TN, which are both expected to show some adaption to drought, based on their origins. High-throughput sequencing of pooled cDNA populations created expression profiles for thousands of genes under water deficit conditions for comparison within controls. This is first study reporting the transcriptomic changes in bambara groundnut under drought stress using RNA-sequencing (RNA-seq) approach, in an attempt to identify potential genes conferring or associated with advantageous traits. Gene expression profiling of 36 samples (3 biological replicate per condition) was conducted. Sequencing was performed as 75bp paired end reads on a HiSeq4000 according to Illumina specifications. The number of raw reads for all DipC (drought and control) and TN (drought and control) samples were over 25 million.

Gene expression analysis from both approaches showed that various aspects of metabolism are affected by drought stress. Hence, there is no single response to a trait as complex as drought tolerance. Both investigation of mild and a relatively severe drought stress for the same two bambara groundnut genotypes adapted to similar environmental conditions provided evidence that the two genotypes used different sets of genes to achieve the same drought response traits (including, ABA synthesis, hormone signaling, osmotic adjustment, accumulation of antioxidants, lignin synthesis, down-regulation of photosynthesis related genes, carbohydrate metabolism, cell-wall

modification and transporters). Hence, both genotypes may have adapted in different ways to enable them to grow in the semi-arid conditions, suggesting that there may be more than a single way to achieve resilience in the face of drought stress. This hypothesis was supported by the differences observed in transcript abundance between the two genotypes under drought stress in a selection of genes associated with drought response such as GST (UoN.bamGnut.682 and UoN.bamGnut.3282), ABA_WDS (UoN.bamGnut.42895 (GO:0006950), Alternative oxidase (UoN.bamGnut.43131 (GO-:0009916) and UoN.bamGnut.15773) and Aspartate proteases (UoN.bamGnut.22301 and UoN.bamGnut.5777), indicating the mechanism by which the two genotypes responding to drought stress might be different. However, validation through qRT-PCR is needed to further validate this hypothesis.

Classically, there are three main responses to drought exhibited by plants, escape, avoidance and tolerance (and, in practice, often a combination of all three)[93, 286, 356]. When gene expression profiles were compared between vegetative and reproductive stage for drought experiment 2 (RNA-seq), greater number of genes associated with drought-response were expressed at the vegetative stage than at the reproductive stage, indicating stronger drought response at the early stage of development in both genotypes, at least in terms of expression of known gene sets. These include plant growth hormones (cytokinin, auxin and ABA) that plays roles in modulating the plant responses towards drought [91, 99, 285, 357], high expression of genes associated with antioxidant defense system (including gene family of peroxidase, GST, ferritin, thioredoxin and glutaredoxin), which plays a vital role in ROS scavenging under drought stress [93, 296, 312]. In addition transcripts encoding some enzymes related to glycolysis were highly induced under drought stress in DipC (GDP-fucose-o-fucosyltransferase (UoN.bamGnut.46713), glyceraldehyde 3-phosphate dehydrogenase (UoN.bamGnut.21011), pyruvate kinase (UoN.bamGnut.5098), glucose-6-

phosphate dehydrogenase (UoN.bamGnut.26564)) and TN (fructose-bisphosphate aldolase (UoN.bamGnut.28748), glyceraldehyde 3-phosphate dehydrogenase (UoN.bamGnut.39128), GDP-fucose-o-fucosyltransferase ((UoN.bamGnut.25866)) at the vegetative stage, suggesting that during drought stress at the vegetative stage, the flow of sugars through the glycolysis pathway is enhanced, possibly for the production of reducing energy, such as NAD(P) H, in the absence of photosynthesis [306, 358, 359]. As would perhaps be expected, down-regulation of genes associated with photosynthesis were observed at both developmental stages including high expression of some other drought induced genes such as LEA, heat shock protein, transporters (ABC and sugar transporters), protein kinases (serine/threonine) and calmodulin-related genes [93, 99, 249, 311, 315]. However, genes playing roles in plant water-deficit response such as Responsive to Dehydration (RD) and Early Responsive to Dehydration (ERD) [99, 143, 360] could not be identified in the bambara groundnut transcriptome using RNA-seq approach. Likewise, the genes of the signal transduction cascade leading to the activation of transcription involving Mitogen-activated protein kinase (MAPKs) were not detected under drought stress in both XSpecies microarray and RNA-seq approaches.

In addition, both drought experiments (mild and severe) also showed that the two genotypes expressed what are classically considered to be 'drought-response' genes even under the control (irrigation) condition. High expression of a gene under control condition using XSpecies microarray analysis was identified through the highly ranked intensity values. While TMM normalised average counts per million (CPM) > 50 across all biological replicates were used to define relatively high expression of a gene within each experimental condition in the RNA-seq experiment. Genes highly expressed under control condition from the XSpecies microarray experiments mostly include homologues of much of the ABA synthesis and response network, the

DREB1 transcription factor, Early-Response to Dehydration proteins, four osmoprotectant genes, two drought-response genes influencing photosynthesis, and 21 other probe sets corresponding to drought-related proteins of unknown function (Appendix 4-6) [93, 99, 285, 288, 308, 361]. Whereas, the RNA-seq experiment showed high expression of little over 45 genes models with annotations related to ABA signaling, and drought responses in all four control conditions (i.e. control conditions at 19 and 33 DAS at vegetative stage and control conditions at 33 and 47 DAS at reproductive stage) (Appendix 16-23). These include homologues of many of the genes involved in the ABA signaling pathway (such as the PP2C gene and family of transcription factors including bZIP, and MYB) [99, 295, 362], osmoprotectants (such as aspartate proteases and beta-fructofuranosidase) [93, 99, 308], several antioxidants (belonging to peroxidase, glutaredoxin, thioredoxin, ferritin and glutathione S-transferases (GST) gene families) [93, 285, 312, 323] and various drought induced genes such as late embryogenesis abundant (LEA), heat shock proteins (HSP), myo-inositol-1-phosphate synthase (MIPS) and dehydrins gene family [93, 174, 175, 271, 347, 358, 363, 364]. To further support these results, some of the highly enriched GO-terms associated with drought response were observed when comparing between control conditions (e.g. Control-33 DAS versus Control-19 DAS and Control-47 DAS versus Control-33 DAS) for differential expression in DipC and TN. For DipC, enriched GO-terms assigned under the biological process category includes 'hydrogen peroxide catabolic process' (GO:0042744) and 'response to external stimulus' (GO:0009605) at the vegetative stage and 'regulation of reactive oxygen species metabolic process' (GO:2000377) at the reproductive stage. Enriched GO-terms assigned under the biological process category for TN includes 'response to abiotic stimulus' (GO:0009628) and 'oxidation-reduction' (GO:0055114) at the reproductive stage. While no GO-terms associated with drought response were observed for TN at the vegetative stage. It could be

speculated that high expression of drought-response genes even under control conditions in both genotypes may lead to greater root growth and other avoidance traits which prime the plant for future dry periods, hence preparing for drought conditions. This hypothesis was supported from the results obtained when both genotypes were subjected to mild drought stress in experiment 1 (XSpecies microarray). The reason for this could well be that the plant is permanently in a drought 'ready' state and this could lead to greater root growth and other avoidance traits, which are permanently induced, leading to an physical and biochemical architecture better able to handle drought when it happens. However, relatively profound effects of drought stress were observed when both genotypes were subjected to more severe drought stress and analysed using the RNA-seq approach. Several family of transcription factors reported to play a role in plant water-deficit response such as members of the MYB [365], NAC [366], Homeodomain (HOX) [93, 367], AP2-ERF [99, 368] and zinc fingers [93, 99, 140, 147, 369] were differentially expressed in both genotypes under drought stress. On top of the common transcription factors found in both genotypes from drought experiment 1 (XSpecies microarray), DipC showed the importance (based on the vertex degrees from network analysis) of WRKY40 (Gma.16733.1.S1_at) [272], while TN showed the importance of CONSTANS-LIKE 1 (GmaAffx.45249.1.S1_at) [255] and MYB60 (GmaAffx.84566.2.S1_at) [275, 370]. On the other hand, drought experiment 2 (RNA-seq) showed the importance of NAC1 (UoN.bamGnut.51896) [336] and MYB46 (UoN.bamGnut.5476) [337] in DipC, while TN showed the importance to MYB36 (UoN.bamGnut.39367) [338], MYB2 (UoN.bamGnut.51895) [142, 343], AUX-IAA (UoN.bamGnut.39167) [93, 371, 372] and NAC25 (UoN.bamGnut.39304) [298].

When gene expression profiles were compared between XSpecies microarray and RNA-seq approaches, it was observed that the latter identified a greater number of differentially expressed genes. However, the XSpecies hybridisation to the soybean

microarray has been shown to be effective in detecting differentially expressed genes. Despite the low number of differentially expressed genes detected from the XSpecies microarray approach here, at least four (PAL1 (GmaAffx.93094.1.S1_s_at), Beta-fructofuranosidase (Gma.1546.1.S1_a_at), COMT (Gma.6501.1.A1_at) and UBC-2 (GmaAffx.67871.1.S1_at)) promising candidate genes for drought tolerance in bambara groundnut were identified and validated using qRT-PCR (Khan *et al.*, paper submitted).

A total of 15 and 77 genes were chosen from XSpecies microarray and RNA-seq experiments, respectively, (based on the fold-change significance, network analysis (based on vertex degree) and their potential association with a reported drought response (based on gene ontology)) to identify their approximate chromosomal location in bambara groundnut using a cross-species approach. With the lack of whole genome sequence for bambara groundnut, common bean (*Phaseolus vulgaris*), belonging to the same legume family of Fabaceae as bambara groundnut was used to identify the approximate syntenic locations of the selected potential candidate genes for drought tolerance. The putative conserved synteny blocks identified across the legume genomes will potentially help in the identification of the genes within the confidence intervals underlying QTL involved in controlling drought traits in bambara groundnut. A total of 4 genes (HOX (UoN.bamGnut.40849; Homeobox family transcription factor); AUX_IAA (AUX_IAA family transcription factor; UoN.bamGnut.39167), acid phosphatase (HAD super-family) (UoN.bamGnut.45698) and dehydrin (UoN.bamGnut.48-031) were found to be near or within the confidence intervals of the QTLs underlying the drought traits (stomatal density/leaf area and CID). Further validation in various genomic backgrounds as well as multi-locational field trails of the derived crosses between the parental lines are necessary to confirm their contribution to the drought traits. Focusing on generating high density and integrated genetic linkage maps, us-

ing dominant DArT, SNPs and GEMs with linkages between orthologous genes in the bambara groundnut genetic and common bean physical genome sequence would improve these initial results as one piece of information for the identification of candidate gene locations in bambara groundnut.

The gene expression data has provided clues about gene function, but it does not reveal what exactly these genes are doing inside a cell. Genetics could provide a solution to this problem because mutants that lack a particular gene may quickly reveal the function of the protein that it encodes. To facilitate such studies of gene function, the effects of the coding sequence of a gene and its regulatory regions can be engineered (either by overexpressing or gene knockout of the gene of interest) to change the functional properties, timing and level of expression of the protein product [344]. This will also provide evidence whether the identified genes are associated with the trait of interest, although always bearing in mind that drought resilience is a complex trait.

While the work presented focussed on gene expression, approaches were undertaken to try to link these data to physiological processes in bambara groundnut in response to drought stress. The first visible sign of altered physiology in both genotypes was through the reduction of stomatal conductance, thus exhibiting strategies to minimise water loss through stomatal closure. With decreased stomatal conductance, reduction in photosynthesis and transpiration were observed as well. These observed physiological responses under drought stress in both genotypes were backed up by high expression of genes related to stomatal closure via ABA signaling. These include Translationally controlled tumor protein [373] (UoN.bamGnut.44892), Protein tyrosine phosphatase [316] (UoN.bamGnut.27176) and Aspartate proteases [330] (UoN.bamGnut.22301), which were up-regulated in DipC. TN, on the other hand, showed up-regulation of Protein tyrosine phosphatase [316] (UoN.bamGnut.27176) and

Aspartate proteases [330] (UoN.bamGnut.5777). Photosynthesis related genes encoding the different protein subunits of the photosystem II reaction center pigment protein complexes (PS II-RC), which is the core of photosystem II and functions as the light reaction center [374], showed decreased expression under drought stress in both genotypes. These include Light harvesting complex (LHC) [297] (UoN.bamGnut.2813), which was down-regulated in TN. Whereas, DipC showed down-regulation of homologues belonging to the PRK [3, 261, 375] (UoN.bamGnut.28815) and PsbP [374, 376] (UoN.bamGnut.6437) family. However, further validation is required to prove if the identified genes are associated with these traits. Furthermore, in terms of physiological comparison of two bambara groundnut genotypes and peanut in relation to drought response, both plant species have showed similar declining trends, with significant reductions in stomatal conductance, photosynthesis and transpiration were observed, which indicates their ability to respond to water limiting conditions through very similar drought mechanisms.

The complexity of drought tolerance, however, requires a more holistic view of biological processes. As this study only deals with the leaf tissues, roots as the primary sensors and transmitters of water scarcity should be considered to understand the whole mechanism governing drought response in bambara groundnut. Secondly, gene expression allows for inference, but does not necessarily reflect the actual metabolism such as post-transcriptional modification or substrate availability/affinity that may influence the impact of expressed genes or translated proteins, respectively. Therefore, transcriptomics should be complemented with other disciplines such as proteomics and metabolomics in order to integrate transcriptomic data into more consistent picture of the physiology of bambara groundnut under drought conditions.

8.3 Implications of the study and future research opportunities

In this study, bambara groundnut was used as an exemplar crop species, to provide a framework on how genomic and transcriptomic methodologies developed for major crops can be applied to bambara groundnut. This will help in better understanding the genetics governing important agronomic traits in bambara groundnut. Furthermore, this study evaluated the effect of drought on changes in gene expression in two genotypes of bambara groundnut (DipC and TN) subjected to drought stress at different developmental stages of plant growth. The results obtained in the present study will help to provide a platform for genetic improvement in bambara groundnut through breeding for varieties with desired traits and will assist in understanding the mechanisms underlying the drought tolerance in bambara groundnut which enables it to grow under semi-arid conditions.

Several strategies has been implemented to produce drought-tolerant varieties in many species using the knowledge of the responses of plants to drought stress and the mechanisms involved [93, 194, 196, 197, 268, 377]. The same strategies can be applied in bambara groundnut to produce enhanced drought tolerant genotypes. The candidate genes associated with drought tolerance can be used as genic molecular markers (GMMs) and integrated into genetic/QTL maps [378]. In cases where the candidate genes identified may be associated with QTLs for drought tolerance traits, genetical genomics approaches can be applied which involves quantitative analysis of transcript profiling of the candidate genes providing the expression QTLs (eQTLs) for drought tolerance-related traits [196]. Expression QTL is an elegant way of combining genetic markers and gene expression profiling [379]. Considering expression levels as quantitative traits, they can be mapped in a segregating population. This will give

an indication of what portion of the variation in gene expression is attributed to the gene itself (cis-acting factors) and to what extent other genomic locations (trans-acting factors) influence gene expression [25]. With this approach, it is possible to detect candidate regulatory genes (influencing expression of other genes) which may have been missed by gene expression profiling due to their low expression levels. Hence, this will help provide additional insight into the regulatory network governing drought trait as has been reported in poplar [380]and rice [363]. Identification of QTL associated with drought tolerance in both genotypes will serve a platform for marker assisted selection of desirable ideotypes. The analysis of gene expression from this study has provided a rich source of biological information, which allows breeders to understand the molecular basis drought response in bambara groundnut, leading to the identification of new targets for manipulating drought response [381]. In addition, transcriptome re-sequencing is a way of genome complexity reduction for discovering SNPs in specific genes of interest. Identification of SNPs in the candidate genes involved in drought response identified from this study will be of particular interest for breeders. The identification SNPs can be used for map construction, map saturation, genome-wide diversity studies and association mapping. This and other technical revolutions (such as genomic selection and marker-assisted selection) will provide genome-wide molecular tools for breeders (large collections of markers, high density genetic markers and new experimental populations) that can be incorporated into existing breeding methods [381].

Another strategy would be to clone the drought responsive genes and responsive elements associated with drought-tolerant QTL which could suggest that these genes may represent the molecular basis of the identified components of drought tolerance through Association Genetic and other positional cloning approaches [93]. However, the enhanced expression of drought related genes in some species is frequently associ-

ated with retarded growth and thus may limit its practical applications. To be able to prove that the transgenic plant is more tolerant to drought stress than wild-type, one would require a rigorous evaluation of the physiological performance as well as water status of transformed plants [361]. Nevertheless, it's clear that the combination of marker assisted selection and transgenics may allow a rapid way of further improving drought tolerance in bambara groundnut. Figure 8.1 illustrate some of the possible ways in which this study could helpful in further research for drought tolerance in bambara groundnut.

Genome-wide association studies (GWAS) is another approach to study causal relationship between genetic polymorphism within a species and the phenotypic differences observed [382]. However due to the insufficient population size, this was not performed in this study. While this thesis presents the opportunity of mapping the candidate drought-associated genes underlying QTL controlling drought traits in biparental cross, it is restricted in allelic diversity and have limited genomic resolution. Using GWAS will overcome several limitations of traditional gene mapping by providing higher resolution (often to the gene level) and using samples from previously well-studied populations in which commonly occurring genetic variations can be associated with phenotypic variation [382]. However, GWAS has limitations, mainly due to their assumption that common genetic variation plays a large role in explaining the heritable variation of common trait. Alternative strategies suggested involve genotyping array-based GWA studies [382, 383]. At this point in time, a GWAS panel of 350 lines is being co-developed with the International Institute of Tropical Agriculture (Ibadan, Nigeria), but is not yet complete.

Underutilised crops such as bambara groundnut having the ability to produce more consistent yield under water-limiting conditions, on poor agricultural land or with low inputs for subsistence farmers compared to other more favoured species, could be a

potential source to increase the food supply as well as to avoid dependence on a limited number of plant species for global food and nutritional security. The importance of drought tolerance for minor crops may relate more to low input, rain-fed, agriculture. Although both XSpecies microarray and RNA-seq approaches possesses advantages and disadvantages, this present study showed that the combined approach is a sensible strategy that could allow molecular mechanisms underlying traits of interest to be studied at the RNA level. If extensive transcriptomic studies and breeding for a crop species with advantageous traits such as drought resilience can be developed, some of the current issues could be resolved, such as over-reliance on staple food crops and development of species with resilience and exceptional traits.

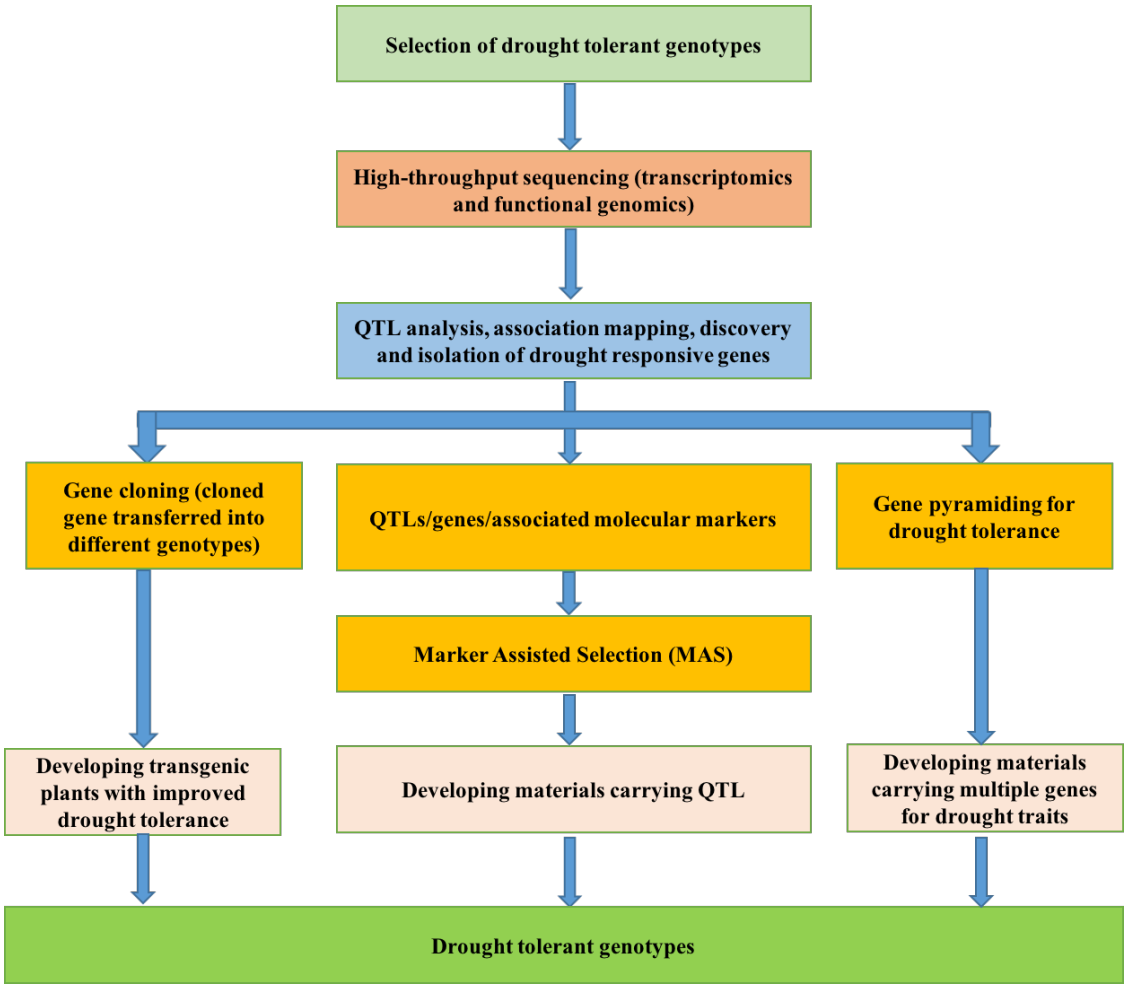

Figure 8.1: **Future opportunities for drought tolerance improvement in bambara groundnut.** The genotypes which are known to survive in areas of water-deficit conditions are selected. To analyse the genes governing drought response in the selected genotypes, gene expression study is performed. Using this information QTL analysis and gene mapping are conducted. For gene cloning, identified genes or major QTL are analysed in detail using a large size population. A cloned gene is transferred into widely adapted varieties. Molecular markers which are linked to the gene or QTL are used for marker-assisted selection. Likewise, marker-assisted selection is used for developing the materials of gene pyramiding.

References

[1] NRC (National Research Council), "Lost Crops of Africa," tech. rep., Washington DC, 1996.

[2] N. S. Ahmad, *Genetic Analysis of Plant Morphology in Bambara Groundnut [Vigna subterranea (L.) Verdc.].* Phd thesis, The University of Nottingham, 2012.

[3] F. Stadler, *Analysis of differential gene expression under water-deficit stress and genetic diversity in bambara groundnut [Vigna subterranea (L .) Verdc .] using novel high-throughput technologies.* Phd thesis, Technical University of Munich, 2009.

[4] K. C. Mohale, A. K. Belane, and F. D. Dakora, "Symbiotic N nutrition, C assimilation, and plant water use efficiency in Bambara groundnut (Vigna subterranea L. Verdc) grown in farmers' fields in South Africa, measured using 15N and 13C natural abundance," *Biology and Fertility of Soils*, vol. 50, no. 2, pp. 307–319, 2013.

[5] S. N. Azam-Ali, A. Sesay, S. Karikari, F. J. Massawe, J. Aguilar-Manjarrez, M. Bannayan, and K. Hampson, "ASSESSING THE POTENTIAL OF AN UNDERUTILIZED CROP. A CASE STUDY USING BAMBARA GROUNDNUT," *Experimental Agriculture*, vol. 37, 2001.

[6] J. Heller, *Bambara groundnut: Vigna subterranea (L.) Verdc. promoting the*

conservation and use of under-utilized and neglected crops. Harare, Zimbabwe: IPGRI, 1997.

[7] F. Begemann, *No Ecogeographic differentiation of Bambara groundnut (Vigna subterranea) in the collection of the International Institute of Tropical Agriculture (IITA). PhD*. PhD thesis, University of Giessen, Germany, 1988.

[8] B. C. Oyiga, M. I. Uguru, and C. B. Aruah, "Pollen behaviour and fertililization impairment in Bambara groundnut (Vigna subterrenea [L.] Verdc.)," *Journal of Plant Breeding and Crop Science*, vol. 2, no. 1, pp. 12–23, 2010.

[9] A. Linnemann and S. Azam-ALI, *Bambara groundnut (Vigna subterranea). In J.T. Williams (ed.), Underutilised Crops. Pulses and Vegetables,*. London: Chapman and Hall: Williams JT, ed. Underutilized crops, series II. Pulses and vegetables, 1993.

[10] E. Doku and S. Karikari, "Operational selection in wild bambara groundnuts," *Ghana journal of science*, vol. 11, no. 1, pp. 47–56, 1971.

[11] R. S. Pasquet, S. Schwedes, and P. Gepts, "Isozyme Diversity in Bambara Groundnut," *Crop Science*, vol. 39, pp. 1228–1236, 1999.

[12] H. H. Chai, *DEVELOPING NEW APPROACHES FOR TRANSCRIPTOMICS AND GENOMICS. USING MAJOR RESOURCES DEVELOPED IN MODEL SPECIES FOR RESEARCH IN CROP SPECIES*. PhD thesis, The University of Nottingham, 2014.

[13] E. R. Forni-Martins, "New Chromosome Number in the Genus Vigna Savi (Leguminosae-Papilionoideae)," *Bulletin du Jardin botanique national de Belgique / Bulletin van de National Plantentuin van Belgi&xeb;*, vol. 56, no. 1/2, pp. 129–133, 1986.

[14] J. N. Berchie, "Evaluation of five bambara groundnut (Vigna subterranea (L.) Verdc.) landraces to heat and drought stress at Tono-Navrongo, Upper East Region of Ghana," *African Journal of Agricultural Reseearch*, vol. 7, no. 2, pp. 250–256, 2012.

[15] S. Basu, S. Mayes, M. Davey, J. A. Roberts, S. N. Azam-Ali, R. Mithen, and R. S. Pasquet, "Inheritance of 'domestication' traits in bambara groundnut (Vigna subterranea (L.) Verdc.)," *Euphytica*, vol. 157, no. 1-2, pp. 59–68, 2007.

[16] I. Makanda, P. Tongoona, R. Madamba1, D. Icishahayo, and J. Derera, "Evaluation of bambara groundnut varieties for off-season production in Zimbabwe," *African Crop Science Journal*, vol. 16, no. 3, pp. 175–183, 2009.

[17] H. C. J. Godfray, J. R. Beddington, I. R. Crute, L. Haddad, D. Lawrence, J. F. Muir, J. Pretty, S. Robinson, S. M. Thomas, and C. Toulmin, "Food security: the challenge of feeding 9 billion people.," *Science (New York, N.Y.)*, vol. 327, no. 5967, pp. 812–818, 2010.

[18] F. Massawe, S. Mayes, and A. Cheng, "Crop Diversity : An Unexploited Treasure Trove for Food Security," *Trends in Plant Science*, vol. 21, no. 5, pp. 1–4, 2016.

[19] S. Mayes, F. J. Massawe, P. G. Alderson, J. A. Roberts, S. N. Azam-Ali, and M. Hermann, "The potential for underutilized crops to improve security of food production," *Journal of Experimental Botany*, vol. 63, no. 3, pp. 1075–1079, 2012.

[20] A. M. Mazahib, M. O. Nuha, I. S. Salawa, and E. E. Babiker, "Some nutritional attributes of bambara groundnut as influenced by domestic processing," *International Food Research Journal*, vol. 20, no. 3, pp. 1165–1171, 2013.

[21] F. J. Massawe, S. S. Mwale, and J. a. Roberts, "Breeding in bambara groundnut (Vigna subterranea (L .) Verdc .): strategic considerations," *African Journal of Biotechnology*, vol. 4, no. 6, pp. 463–471, 2005.

[22] J. Burgess, "Country Pasture Forage Resource Profiles-Botswana.," tech. rep., 2006.

[23] M. Pedercini, H. Kanamaru, and S. Derwisch, "Potential impacts of climate change on food security in Mali. Natural Resources Management and Environment Department," tech. rep., Rome, 2012.

[24] H. H. Chai, F. Massawe, and S. Mayes, "Effects of mild drought stress on the morpho-physiological characteristics of a bambara groundnut segregating population," *Euphytica*, vol. 208, no. 2, 2015.

[25] F. Stadler, *Analysis of differential gene expression under water-deficit stress and genetic diversity in bambara groundnut [Vigna subterranea (L .) Verdc .] using novel high-throughput technologies*. PhD thesis, TECHNICAL UNIVERSITY OF MUNICH, 2009.

[26] A. Nantel, "Transcription Profiling of Candida albicans Cells Undergoing the Yeast-to-Hyphal Transition," *Molecular Biology of the Cell*, vol. 13, no. 10, pp. 3452–3465, 2002.

[27] C. Bar-Or, H. Czosnek, and H. Koltai, "Cross-species microarray hybridizations: a developing tool for studying species diversity," *Trends in Genetics*, vol. 23, no. 4, pp. 200–207, 2007.

[28] R. Garg, R. K. Patel, A. K. Tyagi, and M. Jain, "De novo assembly of chickpea transcriptome using short reads for gene discovery and marker identification," *DNA Research*, vol. 18, no. 1, pp. 53–63, 2011.

[29] J. Pevsner, *Bioinformatics and functional genomics*. No. 610, International, Barnett Discovery, Accelerating Drug, 2002.

[30] L. Pariset, G. Chillemi, S. Bongiorni, V. R. Spica, and A. Valentini, "Microarrays and high-throughput transcriptomic analysis in species with incomplete availability of genomic sequences," *New Biotechnology*, vol. 25, no. 5, pp. 272–279, 2009.

[31] J. P. Hammond, M. R. Broadley, D. J. Craigon, J. Higgins, Z. F. Emmerson, H. J. Townsend, P. J. White, and S. T. May, "Using genomic DNA-based probe-selection to improve the sensitivity of high-density oligonucleotide arrays when applied to heterologous species.," *Plant methods*, vol. 1, no. 1, p. 10, 2005.

[32] Y. Lu, P. Huggins, and Z. Bar-Joseph, "Cross species analysis of microarray expression data," *Bioinformatics*, vol. 25, no. 12, pp. 1476–1483, 2009.

[33] J. A. S. Rinaudo and J. L. Gerin, "Cross-species hybridization: Characterization of gene expression in woodchuck liver using human membrane arrays," *Journal of Medical Virology*, vol. 74, no. 2, pp. 300–313, 2004.

[34] J. C. Yee, K. F. Wlaschin, S. H. Chuah, P. M. Nissom, and W.-S. Hu, "Quality assessment of cross-species hybridization of CHO transcriptome on a mouse DNA oligo microarray," *Biotechnology and Bioengineering*, vol. 101, no. 6, pp. 1359–1365, 2008.

[35] P. Bagnaresi, A. Moschella, O. Beretta, F. Vitulli, P. Ranalli, and P. Perata, "Heterologous microarray experiments allow the identification of the early events associated with potato tuber cold sweetening.," *BMC Genomics*, vol. 9, p. 176, 2008.

[36] M. W. Davey, J. Keulemans, B. Vanholme, N. Graham, S. T. May, and R. Swennen, "Assessing the use of heterologous oligonucleotide microarrays for transcriptomics in a non-model species; Application to the study of drought stress in Musa," *Acta Horticulturae*, vol. 839, pp. 607–614, 2009.

[37] S. Das, P. R. Bhat, C. Sudhakar, J. D. Ehlers, S. Wanamaker, P. a. Roberts, X. Cui, and T. J. Close, "Detection and validation of single feature polymorphisms in cowpea (Vigna unguiculata L. Walp) using a soybean genome array.," *BMC Genomics*, vol. 9, p. 107, 2008.

[38] C. Wu, R. Carta, and L. Zhang, "Sequence dependence of cross-hybridization on short oligo microarrays," *Nucleic Acids Research*, vol. 33, no. 9, pp. 1–7, 2005.

[39] F. S. Pathways, M. Vahey, T. Assessment, D. O. Arrays, M. Bitzer, D. Liang, A. Massimi, S. Kneitz, E. Piek, and E. P. Bottinger, "GeneSpring TM : Tools for Analyzing Data Microarray Expression," vol. 229, pp. 227–229, 2001.

[40] N. S. Graham, S. T. May, Z. C. T. R. Daniel, Z. F. Emmerson, J. M. Brameld, and T. Parr, "Use of the Affymetrix Human GeneChip array and genomic DNA hybridisation probe selection to study ovine transcriptomes.," *Animal : an international journal of animal bioscience*, vol. 5, no. 6, pp. 861–6, 2011.

[41] B. J. Haas and M. C. Zody, "Advancing RNA-Seq analysis," *Nat Biotech*, vol. 28, no. 5, pp. 421–423, 2010.

[42] X. Huang, X.-G. Chen, and P. A. Armbruster, "Comparative performance of transcriptome assembly methods for non-model organisms," *BMC Genomics*, vol. 17, no. 1, p. 523, 2016.

[43] H. Yang, R. Ratnapriya, T. Cogliati, J. Kim, and A. Swaroop, "Vision from next generation sequencing: multi-dimensional genome-wide analysis for producing

gene regulatory networks underlying retinal development, aging and disease.," *Prog Retin Eye Res*, vol. 46, pp. 1–30, 2015.

[44] K. Elmer and A. Meyer, "No Adaptation in the age of ecological genomics: insights from parallelism and convergence.," *Trends Ecol Evol*, vol. 26, no. 6, pp. 298–306, 2011.

[45] B. J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman, and A. Regev, "De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis," *Nature Protocols*, vol. 8, no. 8, pp. 1494–1512, 2013.

[46] M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev, "Full-length transcriptome assembly from RNA-Seq data without a reference genome.," *Nature biotechnology*, vol. 29, no. 7, pp. 644–52, 2011.

[47] L. H. Legault and C. N. Dewey, "Inference of alternative splicing from RNA-Seq data with probabilistic splice graphs," *Bioinformatics*, vol. 29, no. 18, pp. 2300–2310, 2013.

[48] S. Zhao, W. P. Fung-Leung, A. Bittner, K. Ngo, and X. Liu, "Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells," *PLoS ONE*, vol. 9, no. 1, 2014.

[49] J. Ruan, A. Dean, and W. Zhang, "A general co-expression network-based ap-

proach to gene expression analysis: Comparison and applications," *BMC Systems Biology*, vol. 4, 2010.

[50] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 25, pp. 14863–8, 1998.

[51] A. Ghazalpour, S. Doss, B. Zhang, S. Wang, C. Plaisier, R. Castellanos, A. Brozell, E. E. Schadt, T. A. Drake, A. J. Lusis, and S. Horvath, "Integrating genetic and network analysis to characterize genes related to mouse weight," *PLoS Genetics*, vol. 2, no. 8, pp. 1182–1192, 2006.

[52] D. J. Weston, L. E. Gunter, A. Rogers, and S. D. Wullschleger, "Connecting genes, coexpression modules, and molecular signatures to environmental stress phenotypes in plants.," *BMC systems biology*, vol. 2, no. 1, p. 16, 2008.

[53] L. L. Elo, H. Järvenpää, M. Orešič, R. Lahesmaa, and T. Aittokallio, "Systematic construction of gene coexpression networks with applications to human T helper cell differentiation process," *Bioinformatics*, vol. 23, no. 16, pp. 2096–2103, 2007.

[54] A. Aggarwal, D. Li Guo, Y. Hoshida, S. Tsan Yuen, K.-M. Chu, S. So, A. Boussioutas, X. Chen, D. Bowtell, H. Aburatani, S. Y. Leung, and P. Tan, "Topological and Functional Discovery in a Gene Coexpression Meta-Network of Gastric Cancer," *Cancer Research*, vol. 66, no. 1, pp. 232 LP – 241, 2006.

[55] A. P. Presson, E. M. Sobel, J. C. Papp, C. J. Suarez, T. Whistler, M. S. Rajeevan, S. D. Vernon, and S. Horvath, "Integrated Weighted Gene Co-expression

Network Analysis with an Application to Chronic Fatigue Syndrome," *BMC Systems Biology*, vol. 2, no. 1, p. 95, 2008.

[56] H. K. Lee, A. K. Hsu, J. Sajdak, J. Qin, and P. Pavlidis, "Coexpression Analysis of Human Genes Across Many Microarray Data Sets," *Genome Research*, vol. 14, no. 6, pp. 1085–1094, 2004.

[57] J. Li, H. Wei, and P. X. Zhao, "DeGNServer : Deciphering Genome-Scale Gene Networks through High Performance Reverse Engineering Analysis," vol. 2013, no. Mi, 2013.

[58] M. E. Smoot, K. Ono, J. Ruscheinski, P. L. Wang, and T. Ideker, "Cytoscape 2.8: New features for data integration and network visualization," *Bioinformatics*, vol. 27, no. 3, pp. 431–432, 2011.

[59] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.

[60] J. Ellson, E. Gansner, L. Koutsofios, S. C. North, and G. Woodhull, "Graphvizâ Open Source Graph Drawing Tools BT - Graph Drawing: 9th International Symposium, GD 2001 Vienna, Austria, September 23â26, 2001 Revised Papers," pp. 483–484, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[61] I. Jaenicke, H., Höschle-Zeledon, *Strategic Framework for Underutilized Plant Species Research and Development, with Special Reference to Asia and the Pacific, and to Sub-Saharan Africa.* Rome: Colombo and Global Facilitation Unit for Underutilized Species, 2006.

[62] J. T. Williams and N. Haq, *Global Research on Underutilised Crops an Assessment of Current Activities and Proposals*. 2002.

[63] K. T. Moe, S. W. Kwon, and Y. J. Park, "Trends in genomics and molecular marker systems for the development of some underutilized crops," *Genes and Genomics*, pp. 1–16, 2012.

[64] S. Jain and S. Gupta, *Biotechnology of neglected and underutilized crops*. Dordrecht.: Springer Netherlands, 2013.

[65] C. Khoury, B. Laliberté, and L. Guarino, "Trends in ex situ conservation of plant genetic resources: A review of global crop and regional conservation strategies," *Genetic Resources and Crop Evolution*, vol. 57, no. 4, pp. 625–639, 2010.

[66] A. William, D. S. A, N. George, O. R. P. J, O. Olawale, M. Sean, F. Aryo, and A. H. R, "Adoption of Bambara groundnut production and its effects on farmers ' welfare in Northern Ghana," vol. 11, no. 7, pp. 583–594, 2016.

[67] C. Mkandawire, "Review of bambara groundnut (Vigna subterranea (L.) Verdc.) production in Sub-Sahara Africa," *Agricultural Journal 2*, pp. 464–470, 2007.

[68] R. J. Hillocks, C. Bennett, and O. M. Mponda, "Bambara Nut : a Review of Utilisation , Market Potential and Crop Improvement," *African Crop Science Journal*, vol. 20, no. 1, pp. 1–16, 2012.

[69] S. Aliyu, F. Massawe, and S. Mayes, "Genetic diversity and population structure of Bambara groundnut (Vigna subterranea (L.) Verdc.): synopsis of the past two decades of analysis and implications for crop improvement programmes," *Genetic Resources and Crop Evolution*, vol. 63, no. 6, pp. 925–943, 2016.

[70] O. M. Bamshaiye, J. A. Adegbola, and E. I. Bamishaiye, "Bambara groundnut :

an Under-Utilized Nut in Africa," *Advances in Agricultural Biotechnology*, vol. 1, pp. 60–72, 2011.

[71] A. A. Olaleye, E. I. Adeyeye, and A. J. Adesina, "Chemical composition of bambara groundnut (V . subterranea L . Verdc) seed parts," vol. 48, no. 3, pp. 167–178, 2013.

[72] M. Coudert, "Market openings in West Africa for cowpea and bambara groundnut," in *International Trade Forum 20*, pp. 14–29, 1984.

[73] Y. Y. Murevanhema and V. a. Jideani, "Potential of Bambara Groundnut (Vigna subterranea (L.) Verdc) Milk as a Probiotic Beverage-A Review," *Critical Reviews in Food Science and Nutrition*, vol. 53, pp. 954–967, 2013.

[74] M. A. Y. Abdualrahman, A. O. Ali, E. A. Elkhalifa, and A. E. Sulieman, "Effect of bambara groundnut flour (Vigna subterranea (L.) Verdc.) supplementation on chemical, physical, nutritional and sensory evaluation of wheat bread.," *Pakistan journal of biological sciences*, vol. 15, no. 17, pp. 845–9, 2012.

[75] A. Goli, "Bambara groundnut: Bibliographical review," in *Bambara groundnut (Vignia subterranean), Proceedings of the workshop on conservation and improvement of bambara groundnut (Vignia subterranean (L.) Verdc.)*, (Harare, Zimbabwe), pp. 4–10, 1995.

[76] I. Barimalaa and S. Anoghalu, "Effect of Processing on Certain Antinutrients in Bambara Groundnut (Vigna subterranea) Cotyledons," *Journal of the Science of Food and Agriculture*, vol. 73, pp. 186–188, 1997.

[77] O. S. Ijarotimi and Ruth E.T, "Comparison of nutritional composition and anti-ânutrient status of fermented, germinated and roasted bambara groundnut seeds (vigna subterranea)," *British Food Journal*, vol. 111, pp. 376–386, 2009.

[78] N. D. Young and M. Udvardi, "Translating Medicago truncatula genomics to crop legumes," *Current Opinion in Plant Biology*, vol. 12, no. 2, pp. 193–201, 2009.

[79] J. Schmutz, P. E. McClean, S. Mamidi, G. A. Wu, S. B. Cannon, J. Grimwood, J. Jenkins, S. Shu, Q. Song, C. Chavarro, M. Torres-Torres, V. Geffroy, S. M. Moghaddam, D. Gao, B. Abernathy, K. Barry, M. Blair, M. a. Brick, M. Chovatia, P. Gepts, D. M. Goodstein, M. Gonzales, U. Hellsten, D. L. Hyten, G. Jia, J. D. Kelly, D. Kudrna, R. Lee, M. M. S. Richard, P. N. Miklas, J. M. Osorno, J. Rodrigues, V. Thareau, C. a. Urrea, M. Wang, Y. Yu, M. Zhang, R. a. Wing, P. B. Cregan, D. S. Rokhsar, and S. a. Jackson, "A reference genome for common bean and genome-wide analysis of dual domestications.," *Nature genetics*, vol. 46, no. 7, pp. 707–13, 2014.

[80] J. Schmutz, S. B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D. L. Hyten, Q. Song, J. J. Thelen, J. Cheng, D. Xu, U. Hellsten, G. D. May, Y. Yu, T. Sakurai, T. Umezawa, M. K. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, B. Abernathy, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.-C. Zhang, K. Shinozaki, H. T. Nguyen, R. A. Wing, P. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R. C. Shoemaker, and S. A. Jackson, "Genome sequence of the palaeopolyploid soybean.," *Nature*, vol. 463, no. 7278, pp. 178–83, 2010.

[81] M. Muñoz-Amatriaín, H. Mirebrahim, P. Xu, S. I. Wanamaker, M. Luo, H. Alhakami, M. Alpert, I. Atokple, B. J. Batieno, O. Boukar, S. Bozdag, N. Cisse, I. Drabo, J. D. Ehlers, A. Farmer, C. Fatokun, Y. Q. Gu, Y.-N. Guo, B.-L. Huynh, S. A. Jackson, F. Kusi, C. T. Lawley, M. R. Lucas, Y. Ma, M. P. Timko,

J. Wu, F. You, P. A. Roberts, S. Lonardi, and T. J. Close, "Genome resources for climate-resilient cowpea, an essential crop for food security," 2016.

[82] N. Ward and G. Moreno-Hagelsieb, "Quickly finding orthologs as reciprocal best hits with BLAT, LAST, and UBLAST: How much do we miss?," *PLoS ONE*, vol. 9, no. 7, pp. 1–6, 2014.

[83] National Drought Mitigation Center, "Defining Drought: Overview."

[84] N. J. Kooyers, "The evolution of drought escape and avoidance in natural herbaceous populations," *Plant Science*, vol. 234, pp. 155–162, 2015.

[85] M. Ludlow, "Strategies of response to water stress.," in *Structural and functional responses to environmental stresses: Water shortage*, (The Hague), pp. 269–282, SPB Academic publishing, 1989.

[86] P. Gaur, L. Krishnamurthy, and J. Kashiwagi, "Improving drought avoidance root traits in Chickpea (Cicer arietinum. L.). Current status of research at ICRISAT," *Plant Production Science*, vol. 11, pp. 3–11, 2008.

[87] A. Harb, A. Krishnan, M. Ambavaram, and A. Pereira, "Molecular and physiological analysis of drought stress in Arabidopsis reveals early responses leading to acclimation in plant growth.," *Plant Physiology*, vol. 154, pp. 1254– 1271, 2010.

[88] E. Ocampo and R. Robles, "Drought tolerance in Mungbean. I. Osmotic adjustment in drought stressed Mungbean," *The Philippine Journal of Crop Science*, vol. 25, pp. 1–5, 2000.

[89] G. Subbarao, Y. Chauhan, and C. Johansen, "Patterns of osmotic adjustments in pigeonpea â its importance as a mechanism of drought resistance," *European Journal of Agronomy*, vol. 12, pp. 239–249, 2000.

[90] S. Collinson, E. Clawson, S. N. Azam-Ali, and C. Black, "Effects of soil moisture defecits on the water relation of bambara groundnut (Vigna subtarranea L. Verdc)," *Journal of Experimental Botany*, vol. 48, no. 309, pp. 877–884, 1997.

[91] K. Nakashima and K. Yamaguchi-Shinozaki, "ABA signaling in stress-response and seed development," *Plant Cell Reports*, 2013.

[92] D. Man, Y.-x. Bao, and L.-b. Han, "Drought Tolerance Associated with Proline and Hormone Metabolism in Two Tall Fescue Cultivars," vol. 46, no. 7, pp. 1027–1032, 2011.

[93] M. Farooq, A. Wahid, N. Kobayashi, D. Fujita, and S. M. A. Basra, "Plant drought stress: Effects, mechanisms and management," *Sustainable Agriculture*, pp. 153–188, 2009.

[94] S. S. Mwale, S. N. Azam-Ali, and F. J. Massawe, "Growth and development of bambara groundnut (Vigna subterranea) in response to soil moisture. 1. Dry matter and yield," *European Journal of Agronomy*, vol. 26, pp. 345–353, 2007.

[95] N. Gupta, S. Gupta, and A. Kumar, "Effect of Water Stress on Physiological Attributes and their Relationship with Growth and Yield of Wheat Cultivars at Different Stages," *Journal of Agronomy and Crop Science*, vol. 186, pp. 55–62, 2001.

[96] F. Liu, C. Jensen, and M. Andersen, "Hydraulic and chemical signals in the control of leaf expansion and stomatal conductance in soybean exposed to drought stress," *Functional Plant Biology*, vol. 30, 2003.

[97] S. Eslami, G. Gill, and G. McDonald, "Effect of water stress during seed development on morphometric characteristics and dormancy of wild radish (Raphanus raphanistrum L.) seeds.," *International Journal of Plant Production*, 2010.

[98] K. Shinozaki and K. Yamaguchi-Shinozaki, "Gene networks involved in drought stress response and tolerance," *Journal of Experimental Botany*, vol. 58, no. 2, pp. 221–227, 2006.

[99] D. Todaka, K. Shinozaki, and K. Yamaguchi-Shinozaki, "Recent advances in the dissection of drought-stress regulatory networks and strategies for development of drought-tolerant transgenic rice plants.," *Frontiers in plant science*, vol. 6, p. 84, 2015.

[100] D. Collino, J. Dardanelli, R. Sereno, and R. Racca, "Physiological responses of argentine peanut varieties to water stress.: Light interception, radiation use efficiency and partitioning of assimilates," *Field Crops Research*, vol. 70, pp. 177–184, 2001.

[101] A. Anyia and Herzog, "Water-use efficiency, leaf area and leaf gas exchange of cowpeas under mid-season drought," *European Journal of Agronomy*, vol. 20, pp. 327–339, 2004.

[102] P. Singh, "Influence of water deficits on phenology, growth and dry matter allocation in chickpea (Cicer arietinum)," *Field Crops Research*, vol. 28, pp. 1–15, 1991.

[103] G. Okcu, M. Kaya, and M. Atak, "Effects of salt and drought stresses on germination and seedling growth of pea (Pisum sativum L.).," *Turk J Agric*, vol. 29, pp. 237–242, 2005.

[104] M. Blair, C. Galeano, E. Tovar, M. Muñoz Torres, A. Castrillón, S. Beebe, and I. Rao, "Development of a Mesoamerican intra-genepool genetic map for quantitative trait loci detection in a drought tolerant Ã susceptible common

bean (Phaseolus vulgaris L.) cross," *Molecular Breeding*, vol. 29, pp. 71–88, 2010.

[105] H. Teran and S. Singh, "Comparison of Sources and Lines Selected for Drought Resistance in Common Bean," *Crop Science*, vol. 42, 2002.

[106] S. Singh, H. Teran, and J. Gutiérrez, "Registration of SEA 5 and SEA 13 Drought Tolerant Dry Bean Germplasm," *Crop Science*, vol. 41, 2001.

[107] H. Budak, M. Kantar, and K. Yucebilgili Kurtoglu, "Drought Tolerance in Modern and Wild Wheat," *The Scientific World Journal*, pp. 1–16, 2013.

[108] A. M. Poets, Z. Fang, M. T. Clegg, and P. L. Morrell, "Barley landraces are characterized by geographically heterogeneous genomic origins," *Genome Biology*, pp. 1–11, 2015.

[109] J. R. Lasky, H. D. Upadhyaya, P. Ramu, S. Deshpande, C. T. Hash, J. Bonnette, T. E. Juenger, K. Hyma, C. Acharya, S. E. Mitchell, E. S. Buckler, Z. Brenton, S. Kresovich, and G. P. Morris, "Genome-environment associations in sorghum landraces predict adaptive traits," pp. 1–13, 2015.

[110] X. Wei, X. Zhu, J. Yu, L. Wang, Y. Zhang, D. Li, R. Zhou, and X. Zhang, "Identification of Sesame Genomic Variations from Genome Comparison of Landrace and Variety," *Frontiers in Plant Science*, vol. 7, pp. 1–12, 2016.

[111] B. Valliyodan, Dan Qiu, G. Patil, P. Zeng, J. Huang, L. Dai, C. Chen, Y. Li, T. Joshi, L. Song, T. D. Vuong, T. A. Musket, D. Xu, J. G. Shannon, C. Shifeng, X. Liu, and H. T. Nguyen, "Landscape of genomic diversity and trait discovery in soybean," *Scientific Reports*, vol. 6, p. 23598, 2016.

[112] D. Dejan, Z. Miroslav, K. Vesna, P. Dragan, and Š.-M. Gordana, "Comparison of responses to drought stress of 100 wheat accessions and landraces to identify

opportunities for improving wheat drought resistance," *plant breeding*, vol. 131, no. 3, 2012.

[113] T. Mabhaudhi, a. T. Modi, and Y. G. Beletse, "Growth, phenological and yield responses of a bambara groundnut (Vigna subterranea L. Verdc) landrace to imposed water stress: IL. Rain shelter conditions," *Water SA*, vol. 39, no. 2, pp. 191–198, 2013.

[114] X. Y. Huang, D. Y. Chao, J. P. Gao, M. Z. Zhu, M. Shi, and H. X. Lin, "A previously unknown zinc finger protein, DST, regulates drought and salt tolerance in rice via stomatal aperture control," *Genes and Development*, vol. 23, no. 15, pp. 1805–1817, 2009.

[115] M. Benešová, D. Holá, L. Fischer, P. Jedelský, F. Hnilička, and N. Wilhelmová, "The Physiology and Proteomics of Drought Tolerance in Maize: Early Stomatal Closure as a Cause of Lower Tolerance to Short-Term Dehydration?.," *PloS ONE*, vol. 7, 2012.

[116] R. Vurayai, V. Emongor, B. Moseki, and B. M. R. Vurayai, V. Emongor, "Physiological Responses of Bambara Groundnut (Vigna subterranea L. Verdc) to Short Periods of Water Stress During Different Developmental Stages," *Asian Journal of Agricultural Sciences*, vol. 3, no. 1, pp. 37–43, 2011.

[117] M. Kasunga, Q. Liu, S. Miura, K. Yamaguchi-Shinozaki, and K. Shinozaki, "Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor," *Nature Biotechnology*, vol. 17, 1999.

[118] Q. Liu, M. Kasuga, Y. Sakuma, H. Abe, S. Miura, K. Yamaguchi-Shinozaki, and K. Shinozaki, "Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction

pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis.," *The Plant cell*, vol. 10, no. 8, pp. 1391–1406, 1998.

[119] J. G. Dubouzet, Y. Sakuma, Y. Ito, M. Kasuga, E. G. Dubouzet, S. Miura, M. Seki, K. Shinozaki, and K. Yamaguchi-Shinozaki, "OsDREB genes in rice, Oryza sativa L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression," *Plant Journal*, vol. 33, no. 4, pp. 751–763, 2003.

[120] T.-H. Hsieh, J.-t. Lee, Y.-y. Charng, and M.-T. Chan, "Tomato plants ectopically expressing Arabidopsis CBF1 show enhanced resistance to water deficit stress.," *Plant physiology*, vol. 130, no. 2, pp. 618–626, 2002.

[121] S. Oh, C. Kwon, D. Choi, S. Song, and J. Kim, "Expression of barley HvCBF4 enhances tolerance to abiotic stress in transgenic rice," *Plant Biotechnology Journal*, vol. 5, pp. 646–656, 2007.

[122] Q. Wang, Y. Guan, Y. Wu, H. Chen, F. Chen, and C. Chu, "Overexpression of a rice OsDREB1F gene increases salt, drought, and low temperature tolerance in both Arabidopsis and rice," *Plant Mol Biol*, vol. 67, pp. 589–602, 2008.

[123] J. Chen, X. Meng, Y. Zhang, M. Xia, and X. Wang, "Over-expression of OsDREB genes lead to enhanced drought tolerance in rice," *Biotechnology Letters*, no. 30, pp. 2191–2198, 2008.

[124] M. Cui, W. Zhang, Q. Zhang, Z. Xu, Z. Zhu, F. Duan, and R. Wu, "Induced over-expression of the transcription factor OsDREB2A improves drought tolerance in rice," *Plant Physiology and Biochemistry*, vol. 49, no. 12, pp. 1384–1391, 2011.

[125] A. Karaba, S. Dixit, R. Greco, A. Aharoni, K. R. Trijatmiko, N. Marsch-Martinez, A. Krishnan, K. N. Nataraja, M. Udayakumar, and A. Pereira, "Im-

provement of water use efficiency in rice by expression of HARDY, an Arabidopsis drought and salt tolerance gene.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 39, pp. 15270–5, 2007.

[126] S. Morran, O. Eini, T. Pyvovarenko, B. Parent, and R. Singh, "Improvement of stress tolerance of wheat and barley by modulation of expression of DREB/CBF factors," *Plant Biotechnology Journal*, vol. 9, pp. 230–249, 2011.

[127] V. Haake, D. Cook, J. L. Riechmann, O. Pineda, M. F. Thomashow, and J. Z. Zhang, "Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis.," *Plant physiology*, vol. 130, no. 2, pp. 639–648, 2002.

[128] Y. Xiang, N. Tang, H. Du, H. Ye, and L. Xiong, "Characterization of OsbZIP23 as a Key Player of the Basic Leucine Zipper Transcription Factor Family for Conferring Abscisic Acid Sensitivity and Salinity and Drought Tolerance in Rice," *Plant Physiology*, vol. 148, no. 4, pp. 1938–1952, 2008.

[129] N. Tang, H. Zhang, X. Li, J. Xiao, and L. Xiong, "Constitutive activation of transcription factor OsbZIP46 improves drought tolerance in rice.," *Plant physiology*, vol. 158, no. 4, pp. 1755–68, 2012.

[130] G. Lu, C. Gao, X. Zheng, and B. Han, "Identification of OsbZIP72 as a positive regulator of ABA response and drought tolerance in rice," *Planta*, vol. 229, pp. 605–615, 2008.

[131] S. Orellana, M. Yañez, A. Espinoza, I. Verdugo, E. González, S. Ruiz-Lara, and J. a. Casaretto, "The transcription factor SlAREB1 confers drought, salt stress tolerance and regulates biotic and abiotic stress-related genes in tomato," *Plant, Cell and Environment*, vol. 33, no. 12, pp. 2191–2208, 2010.

[132] J. Kang, "Arabidopsis Basic Leucine Zipper Proteins That Mediate Stress-Responsive Abscisic Acid Signaling," *THE PLANT CELL ONLINE*, vol. 14, pp. 343–357, 2002.

[133] H. Hu, M. Dai, J. Yao, B. Xiao, and X. Li, "Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 12987–12992, 2006.

[134] M. Redillas, J. Jeong, Y. Kim, H. Jung, and S. Bang, "The overexpression of OsNAC9 alters the root architecture of rice plants enhancing drought resistance and grain yield under field conditions," *Plant Biotechnology Journal*, vol. 10, pp. 792–805, 2012.

[135] J. S. Jeong, Y. S. Kim, K. H. Baek, H. Jung, S. H. Ha, Y. Do Choi, M. Kim, C. Reuzeau, and J. K. Kim, "Root-Specific Expression of OsNAC10 Improves Drought Tolerance and Grain Yield in Rice under Field Drought Conditions," *Plant Physiology*, vol. 153, no. 1, pp. 185–197, 2010.

[136] J. S. Jeong, Y. S. Kim, M. C. F. R. Redillas, G. Jang, H. Jung, S. W. Bang, Y. D. Choi, S. H. Ha, C. Reuzeau, and J. K. Kim, "OsNAC5 overexpression enlarges root diameter in rice plants leading to enhanced drought tolerance and increased grain yield in the field," *Plant Biotechnology Journal*, vol. 11, no. 1, pp. 101–114, 2013.

[137] K. Nakashima, L.-S. P. Tran, D. Van Nguyen, M. Fujita, K. Maruyama, D. Todaka, Y. Ito, N. Hayashi, K. Shinozaki, and K. Yamaguchi-Shinozaki, "Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice," *The Plant Journal*, vol. 51, no. 4, pp. 617–630, 2007.

[138] A. S. I. Saad, X. Li, H. P. Li, T. Huang, C. S. Gao, M. W. Guo, W. Cheng, G. Y. Zhao, and Y. C. Liao, "A rice stress-responsive NAC gene enhances tolerance of transgenic wheat to drought and salt stresses," *Plant Science*, vol. 203-204, pp. 33–40, 2013.

[139] G.-P. Xue, H. M. Way, T. Richardson, J. Drenth, P. a. Joyce, and C. L. McIntyre, "Overexpression of TaNAC69 leads to enhanced transcript levels of stress up-regulated genes and dehydration tolerance in bread wheat.," *Molecular Plant*, vol. 4, no. 4, pp. 697–712, 2011.

[140] D.-Q. Xu, J. Huang, S.-Q. Guo, X. Yang, Y.-M. Bao, H.-J. Tang, and H.-S. Zhang, "Overexpression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (Oryza sativa L.).," *FEBS letters*, vol. 582, no. 7, pp. 1037–1043, 2008.

[141] B. Z. Xiao, X. Chen, C. B. Xiang, N. Tang, Q. F. Zhang, and L. Z. Xiong, "Evaluation of seven function-known candidate genes for their effects on improving drought resistance of transgenic rice under field conditions," *Molecular Plant*, vol. 2, no. 1, pp. 73–83, 2009.

[142] A. Yang, X. Dai, and W.-H. Zhang, "A R2R3-type MYB gene, OsMYB2, is involved in salt, cold, and dehydration tolerance in rice," *Journal of Experimental Botany*, vol. 63, no. 7, pp. 2541–2556, 2012.

[143] Z. Zhang, X. Liu, X. Wang, M. Zhou, X. Zhou, X. Ye, and X. Wei, "An R2R3 MYB transcription factor in wheat, TaPIMP1, mediates host resistance to Bipolaris sorokiniana and drought stresses through regulation of defense- and stress-related genes," *New Phytologist*, vol. 196, pp. 1155–1170, 2012.

[144] D. Shin, S.-J. Moon, S. Han, B.-G. Kim, S. R. Park, S.-K. Lee, H.-J. Yoon, H. E.

Lee, H.-B. Kwon, D. Baek, B. Y. Yi, and M.-O. Byun, "Expression of StMYB1R-1, a novel potato single MYB-like domain transcription factor, increases drought tolerance.," *Plant Physiology*, vol. 155, no. 1, pp. 421–432, 2011.

[145] X. Wu, Y. Shiroto, S. Kishitani, Y. Ito, and K. Toriyama, "Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing OsWRKY11 under the control of HSP101 promoter," *Plant Cell Reports*, vol. 28, no. 1, pp. 21–30, 2009.

[146] H. Shen, C. Liu, Y. Zhang, X. Meng, X. Zhou, C. Chu, and X. Wang, "OsWRKY30 is activated by MAP kinases to confer drought tolerance in rice," *Plant Molecular Biology*, vol. 80, no. 3, pp. 241–253, 2012.

[147] S. Sugano, H. Kaminaka, Z. Rybka, R. Catala, J. Salinas, K. Matsui, M. Ohme-Takagi, and H. Takatsuji, "Stress-responsive zinc finger gene ZPT2-3 plays a role in drought tolerance in petunia," *Plant Journal*, vol. 36, no. 6, pp. 830–841, 2003.

[148] B. Zhu, J. Su, M. Chang, D. P. S. Verma, Y.-L. Fan, and R. Wu, "Overexpression of a Δ1-pyrroline-5-carboxylate synthetase gene and analysis of tolerance to water- and salt-stress in transgenic rice," *Plant Science*, vol. 139, no. 1, pp. 41–48, 1998.

[149] E. A. H. Pilon-Smits, N. Terry, T. Sears, and K. van Dun, "Enhanced drought resistance in fructan-producing sugar beet," *Plant Physiology and Biochemistry*, vol. 37, no. 4, pp. 313–317, 1999.

[150] K. Holmström, E. Mäntylä, B. Welin, and A. Mandal, "Drought tolerance in tobacco," 1996.

[151] E. Sheveleva, W. Chmara, H. J. Bohnert, and R. C. Jensen, "lncreased Salt and Drought Tolerance by D-Ononitol Production in Transgenic Nicotiana tabacum L.," *Plant Physiology*, vol. 115, no. 1 997, pp. 1211–1219, 1997.

[152] A. K. Garg, J.-K. Kim, T. G. Owens, A. P. Ranwala, Y. D. Choi, L. V. Kochian, and R. J. Wu, "Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 25, pp. 15898–903, 2002.

[153] I.-c. Jang, S.-j. Oh, J.-s. Seo, W.-b. Choi, S. I. Song, C. H. Kim, Y. S. Kim, H.-s. Seo, Y. D. Choi, B. H. Nahm, and J.-k. Kim, "Expression of a Bifunctional Fusion of the," *Society*, vol. 131, pp. 516–524, 2003.

[154] T. Hirayama, C. Ohto, T. Mizoguchi, and K. Shinozaki, "A gene encoding a phosphatidylinositol-specific phospholipase C is induced by dehydration and salt stress in Arabidopsis thaliana.," *Proceedings of the National Academy of Sciences*, vol. 92, no. 9, pp. 3903–3907, 1995.

[155] H.-W. Li, B.-S. Zang, X.-W. Deng, and X.-P. Wang, "Overexpression of the trehalose-6-phosphate synthase gene OsTPS1 enhances abiotic stress tolerance in rice.," *Planta*, vol. 234, no. 5, pp. 1007–18, 2011.

[156] T. Abebe, A. C. Guenzi, B. Martin, and J. C. Cushman, "Tolerance of mannitol-accumulating transgenic wheat to water stress and salinity.," *Plant physiology*, vol. 131, no. 4, pp. 1748–55, 2003.

[157] J. You, H. Hu, and L. Xiong, "An ornithine δ-aminotransferase gene OsOAT confers drought and oxidative stress tolerance in rice," *Plant Science*, vol. 197, pp. 59–69, 2012.

[158] B. D. McKersie, S. R. Bowley, E. Harjanto, and O. Leprince, "Water-Deficit Tolerance and Field Performance of Transgenic Alfalfa Overexpressing Superoxide Dismutase.," *Plant physiology*, vol. 111, no. 4, pp. 1177–1181, 1996.

[159] A. Oberschall, M. Deák, K. Török, L. Sass, I. Vass, I. Kovács, A. Fehér, D. Dudits, and G. V. Horváth, "A novel aldose/aldehyde reductase protects transgenic plants against lipid peroxidation under chemical and drought stresses," *Plant Journal*, vol. 24, pp. 437–446, 2000.

[160] R. Sunkar, D. Bartels, and H.-H. Kirch, "Overexpression of a stress-inducible aldehyde dehydrogenase gene from Arabidopsis thaliana in transgenic plants improves stress tolerance.," *The Plant journal : for cell and molecular biology*, vol. 35, no. 4, pp. 452–464, 2003.

[161] G. Hamid Badawi, Y. Yamauchi, E. Shimada, R. Sasaki, N. Kawano, K. Tanaka, and K. Tanaka, "Enhanced tolerance to salt stress and water deficit by overexpressing superoxide dismutase in tobacco (Nicotiana tabacum) chloroplasts," *Plant Science*, vol. 166, no. 4, pp. 919–928, 2004.

[162] Y. Saijo, S. Hata, J. Kyozuka, K. Shimamoto, and K. Izui, "Over-expression of a single Ca2+-dependent protein kinase confers both cold and salt/drought tolerance on rice plants," *Plant Journal*, vol. 23, no. 3, pp. 319–327, 2000.

[163] J. Ning, X. Li, L. M. Hicks, and L. Xiong, "A Raf-like MAPKKK gene DSM1 mediates drought resistance through reactive oxygen species scavenging in rice.," *Plant physiology*, vol. 152, no. 2, pp. 876–890, 2010.

[164] S.-Q. Ouyang, Y.-F. Liu, P. Liu, G. Lei, S.-J. He, B. Ma, W.-K. Zhang, J.-S. Zhang, and S.-Y. Chen, "Receptor-like kinase OsSIK1 improves drought and salt

stress tolerance in rice (Oryza sativa) plants.," *The Plant journal : for cell and molecular biology*, vol. 62, no. 2, pp. 316–29, 2010.

[165] Y. Ning, C. Jantasuriyarat, Q. Zhao, H. Zhang, S. Chen, J. Liu, L. Liu, S. Tang, C. H. Park, X. Wang, X. Liu, L. Dai, Q. Xie, and G.-L. Wang, "The SINA E3 Ligase OsDIS1 Negatively Regulates Drought Response in Rice.," *Plant physiology*, vol. 157, no. 1, pp. 242–255, 2011.

[166] G.-G. Park, J.-J. Park, J. Yoon, S.-N. Yu, and G. An, "A RING finger E3 ligase gene, Oryza sativa Delayed Seed Germination 1 (OsDSG1), controls seed germination and stress responses in rice.," *Plant molecular biology*, vol. 74, no. 4-5, pp. 467–78, 2010.

[167] T. Gao, Y. Wu, Y. Zhang, L. Liu, Y. Ning, D. Wang, H. Tong, S. Chen, C. Chu, and Q. Xie, "OsSDIR1 overexpression greatly improves drought tolerance in transgenic rice," *Plant Molecular Biology*, vol. 76, no. 1-2, pp. 145–156, 2011.

[168] H. Bae, S. K. Kim, S. K. Cho, B. G. Kang, and W. T. Kim, "Overexpression of OsRDCP1, a rice RING domain-containing E3 ubiquitin ligase, increased tolerance to drought stress in rice (Oryza sativa L.)," *Plant Science*, vol. 180, no. 6, pp. 775–782, 2011.

[169] L. P. Manavalan, X. Chen, J. Clarke, J. Salmeron, and H. T. Nguyen, "RNAi-mediated disruption of squalene synthase improves drought tolerance and yield in rice," *Journal of Experimental Botany*, vol. 63, no. 1, pp. 163–175, 2012.

[170] X. Hou, K. Xie, J. Yao, Z. Qi, and L. Xiong, "A homolog of human ski-interacting protein in rice positively regulates cell viability and stress tolerance.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 15, pp. 6410–5, 2009.

[171] S.-Y. Jiang, R. Bhalla, R. Ramamoorthy, H.-F. Luan, P. N. Venkatesh, M. Cai, and S. Ramachandran, "Over-expression of OSRIP18 increases drought and salt tolerance in transgenic rice plants.," *Transgenic research*, vol. 21, no. 4, pp. 785–95, 2012.

[172] H. Du, N. Wang, F. Cui, X. Li, J. Xiao, and L. Xiong, "Characterization of the beta-carotene hydroxylase gene DSM2 conferring drought and oxidative stress resistance by increasing xanthophylls and abscisic acid synthesis in rice.," *Plant Physiology*, vol. 154, no. 3, pp. 1304–18, 2010.

[173] Z. PELEG, T. FAHIMA, T. KRUGMAN, S. ABBO, D. YAKIR, A. B. KOROL, and Y. SARANGA, "Genomic dissection of drought resistance in durum wheatâÃâwild emmer wheat recombinant inbreed line population," *Plant, Cell & Environment*, vol. 32, no. 7, pp. 758–779, 2009.

[174] B. Xiao, Y. Huang, N. Tang, and L. Xiong, "Over-expression of a LEA gene in rice improves drought resistance under the field conditions," *Theoretical and Applied Genetics*, vol. 115, no. 1, pp. 35–46, 2007.

[175] J. Duan and W. Cai, "OsLEA3-2, an abiotic stress induced gene of rice plays a key role in salt and drought tolerance.," *PloS ONE*, vol. 7, no. 9, p. e45117, 2012.

[176] R. C. Babu, J. Zhang, a. Blum, T. H. D. Ho, R. Wu, and H. T. Nguyen, "HVA1, a LEA gene from barley confers dehydration tolerance in transgenic rice (Oryza sativa L.) via cell membrane protection," *Plant Science*, vol. 166, no. 4, pp. 855–862, 2004.

[177] A. Wei, C. He, B. Li, N. Li, and J. Zhang, "The pyramid of transgenes TsVP

and BetA effectively enhances the drought tolerance of maize plants," *Plant Biotechnology Journal*, vol. 9, no. 2, pp. 216–229, 2011.

[178] P. S. Schnable, D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T. A. Graves, P. Minx, A. D. Reily, L. Courtney, S. S. Kruchowski, C. Tomlinson, C. Strong, K. Delehaunty, C. Fronick, B. Courtney, S. M. Rock, E. Belter, F. Du, K. Kim, R. M. Abbott, M. Cotton, A. Levy, P. Marchetto, K. Ochoa, S. M. Jackson, B. Gillam, W. Chen, L. Yan, J. Higginbotham, M. Cardenas, J. Waligorski, E. Applebaum, L. Phelps, J. Falcone, K. Kanchi, T. Thane, A. Scimone, N. Thane, J. Henke, T. Wang, J. Ruppert, N. Shah, K. Rotter, J. Hodges, E. Ingenthron, M. Cordes, S. Kohlberg, J. Sgro, B. Delgado, K. Mead, A. Chinwalla, S. Leonard, K. Crouse, K. Collura, D. Kudrna, J. Currie, R. He, A. Angelova, S. Rajasekar, T. Mueller, R. Lomeli, G. Scara, A. Ko, K. Delaney, M. Wissotski, G. Lopez, D. Campos, M. Braidotti, E. Ashley, W. Golser, H. Kim, S. Lee, J. Lin, Z. Dujmic, W. Kim, J. Talag, A. Zuccolo, C. Fan, A. Sebastian, M. Kramer, L. Spiegel, L. Nascimento, T. Zutavern, B. Miller, C. Ambroise, S. Muller, W. Spooner, A. Narechania, L. Ren, S. Wei, S. Kumari, B. Faga, M. J. Levy, L. McMahan, P. Van Buren, M. W. Vaughn, K. Ying, C.-T. Yeh, S. J. Emrich, Y. Jia, A. Kalyanaraman, A.-P. Hsia, W. B. Barbazuk, R. S. Baucom, T. P. Brutnell, N. C. Carpita, C. Chaparro, J.-M. Chia, J.-M. Deragon, J. C. Estill, Y. Fu, J. A. Jeddeloh, Y. Han, H. Lee, P. Li, D. R. Lisch, S. Liu, Z. Liu, D. H. Nagel, M. C. McCann, P. SanMiguel, A. M. Myers, D. Nettleton, J. Nguyen, B. W. Penning, L. Ponnala, K. L. Schneider, D. C. Schwartz, A. Sharma, C. Soderlund, N. M. Springer, Q. Sun, H. Wang, M. Waterman, R. Westerman, T. K. Wolfgruber, L. Yang, Y. Yu, L. Zhang, S. Zhou, Q. Zhu, J. L. Bennetzen, R. K. Dawe, J. Jiang, N. Jiang, G. G. Presting, S. R. Wessler, S. Aluru, R. A. Martienssen, S. W. Clifton, W. R. McCombie,

R. A. Wing, and R. K. Wilson, "The B73 maize genome: complexity, diversity, and dynamics.," *Science (New York, N.Y.)*, vol. 326, no. 5956, pp. 1112–5, 2009.

[179] M. L. Metzker, "Sequencing technologies - the next generation.," *Nature reviews. Genetics*, vol. 11, no. 1, pp. 31–46, 2010.

[180] M. W. Bevan and C. Uauy, "Genomics reveals new landscapes for crop improvement," *Genome Biology*, pp. 4–6, 2013.

[181] J. Shendure and E. Lieberman Aiden, "The expanding scope of DNA sequencing.," *Nature biotechnology*, vol. 30, no. 11, pp. 1084–94, 2012.

[182] P. Thoquet, M. Gherardi, E. P. Journet, A. Kereszt, J.-M. Ané, J.-M. Prospéri, T. Huguet, M. Ghérardi, and J.-M. Prosperi, "The molecular genetic linkage map of the model legume Medicago truncatula: an essential tool for comparative legume genomics and the isolation of agronomically important genes.," *BMC Plant Biology*, vol. 2, p. 1, 2002.

[183] R. Freyre, P. W. Skroch, V. Geffroy, A. F. Adam-Blondon, A. Shirmohamadali, W. C. Johnson, V. Llaca, R. O. Nodari, P. A. Pereira, S. M. Tsai, J. Tohme, M. Dron, J. Nienhuis, C. E. Vallejos, and P. Gepts, "Towards an integrated linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP maps," *Theoretical and Applied Genetics*, vol. 97, no. 5-6, pp. 847–856, 1998.

[184] Q. J. Song, L. F. Marek, R. C. Shoemaker, K. G. Lark, V. C. Concibido, X. Delannay, J. E. Specht, and P. B. Cregan, "A new integrated genetic linkage map of the soybean.," *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, vol. 109, pp. 122–128, 2004.

[185] C. M. Menendez, a. E. Hall, and P. Gept, "Genetic linkage map of cowpea (Vigna unguiculata) developed from a cross between twi inbred,domesticated lines," *Theoretical and Applied Genetics*, vol. 95, pp. 1210–1217, 1997.

[186] S. R. McCouch, L. Teytelman, Y. Xu, K. B. Lobos, K. Clare, M. Walton, B. Fu, R. Maghirang, Z. Li, Y. Xing, Q. Zhang, I. Kono, M. Yano, R. Fjellstrom, G. DeClerck, D. Schneider, S. Cartinhour, D. Ware, and L. Stein, "Development and mapping of 2240 new SSR markers for rice (Oryza sativa L.)," *DNA Research*, vol. 9, no. 6, pp. 199–207, 2002.

[187] M. Sledge, I. Ray, and G. Jiag, "An expressed sequence tag SSR map of tetraploid alfalfa (Medicago sativa L.)," *Theor Appl Genet*, vol. 111, pp. 980–992, 2005.

[188] M. Gutierrez, M. Vaz Patto, T. Huguet, J. Cubero, M. Moreno, and A. Torres, "Cross-species amplification of Medicago truncatula microsatellites across three major pulse crops," *Theor Appl Genet*, vol. 110, pp. 1210–1217, 2005.

[189] Y. Zhang, M. Sledge, and J. Bouton, "Genome mapping of white clover (Trifolium repens L.) and comparative analysis within the Trifolieae using cross-species SSR markers," *Theor Appl Genet*, vol. 114, pp. 1367–1378, 2007.

[190] C. Ameline-Torregrosa, M. Cazaux, D. Danesh, F. Chardon, S. B. Cannon, M. T. Esquerre-Tugaye, B. Dumas, N. D. Young, D. A. Samac, T. Huguet, C. Jacquet, M.-T. Esquerré-Tugayé, B. Dumas, N. D. Young, D. A. Samac, T. Huguet, and C. Jacquet, "Genetic Dissection of Resistance to Anthracnose and Powdery Mildew in Medicago truncatula," *Molecular Plant-Microbe Interactions*, vol. 21, no. 1, pp. 61–69, 2007.

[191] S. Yang, M. Gao, C. Xu, J. Gao, S. Deshpande, S. Lin, B. A. Roe, and H. Zhu,

"Alfalfa benefits from Medicago truncatula: the RCT1 gene from M. truncatula confers broad-spectrum resistance to anthracnose in alfalfa.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 34, pp. 12164–9, 2008.

[192] L. G. Kamphuis, J. Lichtenzveig, R. P. Oliver, and S. R. Ellwood, "Two alternative recessive quantitative trait loci influence resistance to spring black stem and leaf spot in Medicago truncatula.," *BMC Plant Biology*, vol. 8, p. 30, 2008.

[193] R. J. Nelson, R. L. Naylor, and M. M. Jahn, "The Role of Genomics Research in Improvement of Orphan Crops," *Crop Science*, vol. 44, no. 6, p. 1901, 2004.

[194] B. P. M. Swamy and A. Kumar, "Genomics-based precision breeding approaches to improve drought tolerance in rice.," *Biotechnology advances*, vol. 31, no. 8, pp. 1308–18, 2013.

[195] H. Shashidhar, A. Kanbar, M. Toorchi, G. Raveendra, and P. Kundur, "Breeding for Drought Resistance Using Whole Plant Architecture â Conventional and Molecular Approach," in *Plant Breeding from Laboratories to Fields* (S. Andersen, ed.), (Vienna, Austria), 2013.

[196] R. K. Varshney, A. Graner, and M. E. Sorrells, "Genomics-assisted breeding for crop improvement.," *Trends in Plant Science*, vol. 10, no. 12, pp. 621–30, 2005.

[197] M. Thudi, P. M. Gaur, L. Krishnamurthy, R. R. Mir, H. Kudapa, A. Fikre, P. Kimurto, S. Tripathi, K. R. Soren, R. Mulwa, C. Bharadwaj, S. Datta, S. K. Chaturvedi, and R. K. Varshney, "Genomics-assisted breeding for drought tolerance in chickpea," *Functional Plant Biology*, vol. 41, no. 11, pp. 1178–1190, 2014.

[198] T. H. E. Meuwissen, B. J. Hayes, and M. E. Goddard, "Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps," *Genetics*, vol. 157, no. 4, pp. 1819–1829, 2001.

[199] P. L. Morrell, E. S. Buckler, and J. Ross-Ibarra, "Crop genomics: advances and applications.," *Nature reviews. Genetics*, vol. 13, no. 2, pp. 85–96, 2011.

[200] S. R. Eathington, T. M. Crosbie, M. D. Edwards, R. S. Reiter, and J. K. Bull, "Molecular Markers in a Commercial Breeding Program," *Crop Science*, vol. 47, no. Supplement_3, pp. S–154, 2007.

[201] P. Kendabie, S. Jørgensen, F. Massawe, S. Azam-Ali, and S. Mayes, "Day-length effects on growth and seed production efficiency in Bambara groundnut (Vigna subterranea L.)," in *3rd International conference on neglected and underutilized species: For a food-secure Africa, Accra, Ghana (pp. 7)* (R. Hall, P. Rudebjer, and S. Padulosi, eds.), Book of Abstract Rome: Bioversity International, 2013.

[202] N. S. Ahmad, S. M. Basu, E. S. Redjeki, E. Murchie, F. Massawe, S. Azam-Ali, a. Kilian, and S. Mayes, "Developing genetic mapping and marker-assisted breeding techniques in Bambara groundnut (Vigna subterranea L.)," *Acta Horticulturae*, vol. 979, pp. 437–450, 2013.

[203] H. Chai, L. H, G. H, B. S, F. Massawe, S. Graham, and S. Mayes, "Developing XSpecies approaches for genomics and transcriptomics. Using Resources developed in major species for research in Bambara," in *2nd International symposium on underutilized species: Crops for the future â Beyond food security*, (Kuala Lumpur), pp. 773–778, Acta Horticulture, 2013.

[204] R. Beena, M. Sheshshayee, J. Madhura, T. Prasad, and M. Udayakumar, "Development of SSR Markers and Genetic Variability in Physiological Traits in

Bambara Groundnut (Vigna subterranea L. Verdc)," in *Prospects in Bioscience: Addressing the Issues*, pp. 229–242, 2012.

[205] O. O. Molosiwa, S. Aliyu, F. Stadler, K. Mayes, F. Massawe, A. Kilian, and S. Mayes, "SSR marker development, genetic diversity and population structure analysis of Bambara groundnut [Vigna subterranea (L.) Verdc.] landraces," *Genetic Resources and Crop Evolution*, vol. 62, no. 8, pp. 1225–1243, 2015.

[206] B. Olukolu, S. Mayes, F. Stadler, Q. Ng, I. Fawole, D. Dominique, S. N. Azam-Ali, A. G. Abbott, and C. Kole, "Genetic diversity in Bambara groundnut (Vigna subterranea (L.) Verdc.) as revealed by phenotypic descriptors and DArT marker analysis," *Genetic Resources and Crop Evolution*, vol. 59, no. 3, pp. 347–358, 2012.

[207] F. J. Massawe, M. Dickinson, J. a. Roberts, and S. N. Azam-Ali, "(L.) Verdc) landraces revealed by AFLP markers," *Genome*, vol. 45, no. 6, pp. 1175–1180, 2002.

[208] S. Basu, J. Roberts, M. Davey, S. Azam-Ali, and R. Mithen, "Towards genetic linkage mapping in Bambara groundnut," in *Proceedings of the international bambara groundnut symposium in Gaborone, Botswana*, pp. 211–222, 2003.

[209] N. Ahmad, E. Redjeki, W. Kuan, S. Aliyu, K. Mayes, F. Massawe, A. Kilian, and S. Mayes, "Construction of a genetic linkage map and QTL analysis in bambara groundnut," *Genome*, vol. 59, 2016.

[210] H. Hu and L. Xiong, "Genetic Engineering and Breeding of Drought-Resistant Crops," *Annual Review of Plant Biology*, vol. 65, no. 1, pp. 715–741, 2014.

[211] B. A. AkpÄ±*nar*, *S. J.Lucas*, *andH. Budak*, "*GenomicsApproachesforCropImprovementaga*
$-9, 2013$.

[212] A. P. Dhanapal, "Genomics of crop plant genetic resources," *Advances in Bioscience and Biotechnology*, vol. 03, no. 04, pp. 378–385, 2012.

[213] S. B. Cannon, G. D. May, and S. A. Jackson, "Three Sequenced Legume Genomes and Many Crop Species: Rich Opportunities for Translational Genomics," *Plant Physiology*, vol. 151, no. 3, pp. 970–977, 2009.

[214] I. Armstead, I. Donnison, S. Aubry, J. Harper, S. Hörtensteiner, C. James, J. Mani, M. Moffet, H. Ougham, L. Roberts, A. Thomas, N. Weeden, H. Thomas, and I. King, "Cross-species identification of Mendel's I locus.," *Science (New York, N.Y.)*, vol. 315, no. 5808, p. 73, 2007.

[215] S. Mayes, S. M. Basu, O. Molosiwa, E. S. Redjeki, N. S. Ahmad, F. Khan, S. Zehra, S. Noah, K. Mayes, J. Roberts, F. Stadler, F. Massawe, A. Kilian, and S. Azam-Ali, "Molecular Analysis of Bambara Groundnut , an Underutilised African Legume Crop as Part of the BAMLINK Project â What Lessons Can We Learn?," *Proc. 2nd Int. Symp. on Underutilized Plants Species "Crops for the Future - Beyond Food Security"*, pp. 451–458, 2013.

[216] V. S. Bonthala, K. Mayes, J. Moreton, M. Blythe, V. Wright, S. May, F. Massawe, S. Mayes, and J. Twycross, "Identification of Gene Modules Associated with Low Temperatures Response in Bambara groundnut by Network-based Analysis," *PLoS ONE*, pp. 1–18, 2015.

[217] E. R. Mardis, "The impact of next-generation sequencing technology on genetics.," *Trends in Genetics*, vol. 24, pp. 133–41, mar 2008.

[218] L. Yan, D. Fu, C. Li, A. Blechl, G. Tranquilli, M. Bonafede, A. Sanchez, M. Valarik, S. Yasuda, and J. Dubcovsky, "The wheat and barley vernalization gene VRN3 is

an orthologue of FT," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 19581–19586, dec 2006.

[219] J. A. Higgins, P. C. Bailey, and D. A. Laurie, "Comparative genomics of flowering time pathways using Brachypodium distachyon as a model for the temperate grasses.," *PloS one*, vol. 5, p. e10065, jan 2010.

[220] R. Bhattacharjee, "Harnessing biotechnology for conservation and utilization of genetic diversity in orphan crops," *African Technology Development Forum*, vol. 6, no. 3/4, pp. 24–31, 2009.

[221] J. Ye, G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. L. Madden, "Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction," *BMC Bioinformatics*, vol. 13, no. 1, p. 134, 2012.

[222] Y. J. Kang, S. K. Kim, M. Y. Kim, P. Lestari, K. H. Kim, B.-K. Ha, T. H. Jun, W. J. Hwang, T. Lee, J. Lee, S. Shim, M. Y. Yoon, Y. E. Jang, K. S. Han, P. Taeprayoon, N. Yoon, P. Somta, P. Tanya, K. S. Kim, J.-G. Gwag, J.-K. Moon, Y.-H. Lee, B.-S. Park, A. Bombarely, J. J. Doyle, S. a. Jackson, R. Schafleitner, P. Srinives, R. K. Varshney, and S.-H. Lee, "Genome sequence of mungbean and insights into evolution within Vigna species.," *Nature communications*, vol. 5, p. 5443, 2014.

[223] R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed, "Exploration, normalization, and summaries of high density oligonucleotide array probe level data," *Biostatistics*, vol. 4, no. 2, pp. 249–264, 2003.

[224] S. Maere, K. Heymans, and M. Kuiper, "BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks," *Bioinformatics*, vol. 21, no. 16, pp. 3448–3449, 2005.

[225] R. Hu, C. Fan, and L. Hongyu, "Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR," *BMC Molecular Biology*, vol. 10, no. 93, 2009.

[226] M. Jain, A. Nijhawan, A. K. Tyagi, and J. P. Khurana, "Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR," *Biochemical and Biophysical Research Communications*, vol. 345, no. 2, pp. 646–651, 2006.

[227] S. P. Long, P. K. Farage, and R. L. Garcia, "Measurement of leaf and canopy photosynthetic C0 2 exchange in the field 1," *Journal of Experimental Botany*, vol. 47, no. 304, pp. 1629–1642, 1996.

[228] J. Clausen and T. Kozlowski, "USE OF THE RELATIVE TURGIDITY TECHNIQUE FOR MEASUREMENT OF WATER STRESSES IN GYMNOSPERM LEAVES," *Canadian Journal of Botany*, 1965.

[229] H. D. Barrs and P. E. Weatherley, "A Re-Examination of the Relative Turgidity Technique for Estimating Water Deficits in Leaves," *Australian Journal of Biological Sciences*, vol. 15, no. 3, pp. 413–428, 1962.

[230] S. Lamble, E. Batty, M. Attar, D. Buck, R. Bowden, G. Lunter, D. Crook, B. El-Fahmawi, and P. Piazza, "Improved workflows for high throughput library preparation using the transposome-based Nextera system.," *BMC Biotechnology*, vol. 13, p. 104, 2013.

[231] M. Martin, "Cutadapt removes adapter sequences from high-throughput sequencing reads," *EMBnet.journal*, vol. 17, no. 1, p. 10, 2011.

[232] R. Schmieder and R. Edwards, "Quality control and preprocessing of metagenomic datasets," *Bioinformatics*, vol. 27, no. 6, pp. 863–864, 2011.

[233] M. Pertea, G. M. Pertea, C. M. Antonescu, T.-C. Chang, J. T. Mendell, and S. L. Salzberg, "StringTie enables improved reconstruction of a transcriptome from RNA-seq reads," *Nat Biotech*, vol. 33, no. 3, pp. 290–295, 2015.

[234] S. Anders, P. T. Pyl, and W. Huber, "HTSeq-A Python framework to work with high-throughput sequencing data," *Bioinformatics*, vol. 31, no. 2, pp. 166–169, 2015.

[235] M. D. Robinson, D. J. McCarthy, and G. K. Smyth, "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data," *Bioinformatics*, vol. 26, no. 1, pp. 139–140, 2010.

[236] M. Robinson and A. Oshlack, "A scaling normalization method for differential expression analysis of RNA-seq data," *Genome Biology*, vol. 11, no. 3, p. R25, 2010.

[237] S. Falcon and R. Gentleman, "Using GOstats to test gene lists for GO term association," *Bioinformatics*, vol. 23, no. 2, pp. 257–258, 2007.

[238] H. Chai, W. Ho, N. Graham, S. May, F. Massawe, and S. Mayes, "A Cross-Species Gene Expression Marker-Based Genetic Map and QTL Analysis in Bambara Groundnut," *Genes*, vol. 8, no. 3, p. 84, 2017.

[239] Y. C. Li, F. R. Meng, C. Y. Zhang, N. Zhang, M. S. Sun, J. P. Ren, H. B. Niu, X. Wang, and J. Yin, "Comparative analysis of water stress-responsive transcriptomes in drought-susceptible and -tolerant wheat (Triticum aestivum L.)," *Journal of Plant Biology*, vol. 55, no. 5, pp. 349–360, 2012.

[240] I. Majláth, E. Darko, B. Palla, Z. Nagy, T. Janda, and G. Szalai, "Reduced light and moderate water deficiency sustain nitrogen assimilation and sucrose degradation at low temperature in durum wheat," *Journal of Plant Physiology*, vol. 191, pp. 149–158, 2016.

[241] M. Vaittinen, D. Kaminska, P. Kkel, M. Eskelinen, M. Kolehmainen, J. Pihlajam??ki, M. Uusitupa, and L. Pulkkinen, "Downregulation of CPPED1 expression improves glucose metabolism in vitro in adipocytes," *Diabetes*, vol. 62, no. 11, pp. 3747–3750, 2013.

[242] Y.-i. Tsukada, J. Fang, H. Erdjument-Bromage, M. Warren, C. Borchers, P. Tempst, and Y. Zhang, "No Histone demethylation by a family of JmjC domain-containing proteins," *Nature*, vol. 439, pp. 811–816, 2006.

[243] Y. Ding, N. Liu, L. Virlouvet, J.-J. Riethoven, M. Fromm, and Z. Avramova, "Four distinct types of dehydration stress memory genes in Arabidopsis thaliana.," *BMC plant biology*, vol. 13, no. 1, p. 229, 2013.

[244] T.-w. Kim, M. Michnlewicz, D. Begmann, and Z.-Y. Wang, "Brassinosteroid regulates stomatal development by GSK3- mediated inhibition of a MAPK pathway," *Nature*, vol. 482, no. 7385, pp. 419–422, 2012.

[245] T. P. Brutnell, "BUNDLE SHEATH DEFECTIVE2, a Novel Protein Required for Post-Translational Regulation of the rbcL Gene of Maize," *THE PLANT CELL ONLINE*, vol. 11, no. 5, pp. 849–864, 1999.

[246] A. Gholizadeh, "Effects of drought on the activity of Phenylalanine ammonia lyase in the leaves and roots of maize inbreds," *Australian Journal of Basic and Applied Sciences*, vol. 5, no. 9, pp. 952–956, 2011.

[247] L. B. Liliane, C. Magioli, A. B. U. C. M. Perez, R. Margis, G. Sachetto-Martins, and M. Margis-Pinheiro, "AtchitIV gene expression is stimulated under abiotic stresses and is spatially and temporally regulated during embryo development," *Genetics and Molecular Biology*, vol. 27, no. 1, pp. 118–123, 2004.

[248] Y. Huang, X. Chen, Y. Liu, C. Roth, C. Copeland, H. E. McFarlane, S. Huang, V. Lipka, M. Wiermer, and X. Li, "Mitochondrial AtPAM16 is required for plant survival and the negative regulation of plant immunity.," *Nature communications*, vol. 4, p. 2558, 2013.

[249] C. Wasternack, "Jasmonates: An update on biosynthesis, signal transduction and action in plant stress response, growth and development," *Annals of Botany*, vol. 100, no. 4, pp. 681–697, 2007.

[250] G. A. Zhou, R. Z. Chang, and L. J. Qiu, "Overexpression of soybean ubiquitin-conjugating enzyme gene GmUBC2 confers enhanced drought and salt tolerance through modulating abiotic stress-responsive gene expression in Arabidopsis," *Plant Molecular Biology*, vol. 72, no. 4, pp. 357–367, 2010.

[251] P. Scherle, T. Behrens, and L. M. Staudt, "Ly-GDI, a GDP-dissociation inhibitor of the RhoA GTP-binding protein, is expressed preferentially in lymphocytes.," *Proceedings of the National Academy of Sciences*, vol. 90, no. 16, pp. 7568–7572, 1993.

[252] L. Chen and R. Bush, "LHT 1, A Lysine and Histidine Specific Amino Acid Transporter in Arabidopsis," *Plant Physiology*, no. 1 997, 1995.

[253] Y. Hu, W.-C. Li, Y.-Q. Xu, G.-J. Li, Y. Liao, and F.-L. Fu, "Differential expression of candidate genes for lignin biosynthesis under drought stress in maize leaves.," *Journal of applied genetics*, vol. 50, no. 3, pp. 213–223, 2009.

[254] R. Srinivasan and D. J. Oliver, "H-Protein of the Glycine Decarboxylase Multienzyme Complex : Complementary DNA Encoding the Protein from Arabidopsis thaliana.," *Plant physiology*, vol. 98, pp. 1518–1519, 1992.

[255] S. Ledger, C. Strayer, F. Ashton, S. A. Kay, and J. Putterill, "Analysis of the function

of two circadian-regulated CONSTANS-LIKE genes," *Plant Journal*, vol. 26, no. 1, pp. 15–22, 2001.

[256] A. L. Eamens, K. Wook Kim, and P. M. Waterhouse, "DRB2, DRB3 and DRB5 function in a non-canonical microRNA pathway in Arabidopsis thaliana.," *Plant signaling & behavior*, vol. 7, no. 10, pp. 1224–9, 2012.

[257] S. E. Ades, I. L. Grigorova, and C. A. Gross, "Regulation of the alternative sigma factor sigma(E) during initiation, adaptation, and shutoff of the extracytoplasmic heat shock response in Escherichia coli.," *Journal of bacteriology*, vol. 185, no. 8, pp. 2512–9, 2003.

[258] M. Shimizu, H. Kato, T. Ogawa, A. Kurachi, Y. Nakagawa, and H. Kobayashi, "Sigma factor phosphorylation in the photosynthetic control of photosystem stoichiometry," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 23, pp. 10760–10764, 2010.

[259] C. Periappuram, L. Steinhauer, D. L. Barton, D. C. Taylor, B. Chatson, and J. Zou, "The plastidic phosphoglucomutase from Arabidopsis. A reversible enzyme reaction with an important role in metabolic control," *Plant Physiology*, vol. 122, no. 4, pp. 1193–1199, 2000.

[260] V. K. Singh and M. Jain, "Genome-wide survey and comprehensive expression profiling of Aux/IAA gene family in chickpea and soybean," *Frontiers in Plant Science*, vol. 6, pp. 1–15, 2015.

[261] N. Wedel, J. Soll, and B. K. Paap, "CP12 provides a new mode of light regulation of Calvin cycle activity in higher plants.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 19, pp. 10479–10484, 1997.

[262] K. L. Lorick, J. P. Jensen, S. Fang, a. M. Ong, S. Hatakeyama, and a. M. Weissman, "RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquit-

ination.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 20, pp. 11364–11369, 1999.

[263] S. Lindemose, C. O'Shea, M. K. Jensen, and K. Skriver, "Structure, function and networks of transcription factors involved in abiotic stress responses," *International Journal of Molecular Sciences*, vol. 14, no. 3, pp. 5842–5878, 2013.

[264] L. Peng, H. Yamamoto, and T. Shikanai, "Structure and biogenesis of the chloroplast NAD(P)H dehydrogenase complex," *Biochimica et Biophysica Acta - Bioenergetics*, vol. 1807, no. 8, pp. 945–953, 2011.

[265] C. Liu, B. Mao, S. Ou, W. Wang, L. Liu, Y. Wu, C. Chu, and X. Wang, "OsbZIP71, a bZIP transcription factor, confers salinity and drought tolerance in rice," *Plant Molecular Biology*, vol. 84, no. 1-2, pp. 19–36, 2014.

[266] A. Marshall, R. B. Aalen, D. Audenaert, T. Beeckman, M. R. Broadley, M. A. Butenko, A. I. Caño-Delgado, S. de Vries, T. Dresselhaus, G. Felix, N. S. Graham, J. Foulkes, C. Granier, T. Greb, U. Grossniklaus, J. P. Hammond, R. Heidstra, C. Hodgman, M. Hothorn, D. Inzé, L. Østergaard, E. Russinova, R. Simon, A. Skirycz, Y. Stahl, C. Zipfel, and I. De Smet, "Tackling Drought Stress: RECEPTOR-LIKE KINASES Present New Approaches," *The Plant Cell*, vol. 24, no. 6, pp. 2262–2278, 2012.

[267] Z. Andrés, J. Pérez-Hormaeche, E. O. Leidi, K. Schlücking, L. Steinhorst, D. H. McLachlan, K. Schumacher, A. M. Hetherington, J. Kudla, B. Cubero, and J. M. Pardo, "Control of vacuolar dynamics and regulation of stomatal aperture by tonoplast potassium uptake.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 17, pp. E1806–14, 2014.

[268] P. Tripathi, R. C. Rabara, R. N. Reese, M. A. Miller, J. S. Rohila, S. Subramanian, Q. J. Shen, D. Morandi, H. Bücking, V. Shulaev, and P. J. Rushton, "A toolbox of

genes, proteins, metabolites and promoters for improving drought tolerance in soybean includes the metabolite coumestrol and stomatal development genes," *BMC Genomics*, vol. 17, no. 1, p. 102, 2016.

[269] H. J. Xia and G. Yang, "Inositol 1,4,5-trisphosphate 3-kinases: functions and regulations.," *Cell research*, vol. 15, no. 2, pp. 83–91, 2005.

[270] P. Nayamudeza, *Crop water use and the root systems ofbambara groundnut (Vigna subterranea (L.) Verde.) and groundnut (Arachis hypogaea (L.)) in response to irrigation and drought*. Msc, The University of Nottingham, 1989.

[271] D. T. Le, R. Nishiyama, Y. Watanabe, M. Tanaka, M. Seki, L. H. Ham, K. Yamaguchi-Shinozaki, K. Shinozaki, and L. S. P. Tran, "Differential Gene Expression in Soybean Leaf Tissues at Late Developmental Stages under Drought Stress Revealed by Genome-Wide Transcriptome Analysis," *PLoS ONE*, vol. 7, no. 11, 2012.

[272] D. Singh and A. Laxmi, "Transcriptional regulation of drought response: a tortuous network of transcriptional factors," *Frontiers in Plant Science*, vol. 6, pp. 1–11, 2015.

[273] J. Grundy, C. Stoker, and I. A. Carré, "Circadian regulation of abiotic stress tolerance in plants," vol. 6, no. August, pp. 1–15, 2015.

[274] J. Liu, J. Shen, Y. Xu, X. Li, J. Xiao, and L. Xiong, "Ghd2, a CONSTANS -like gene, confers drought sensitivity through regulation of senescence in rice," *Journal of Experimental Botany*, vol. 67, no. 19, 2016.

[275] E. J. Oh, Y. Kwon, J. Kim, Hyeok, and H. Noh, "A dual role for MYB60 in stomatal regulation and root growth of Arabidopsis thaliana under drought stress," *Plant Molecular Biology*, vol. 77, no. 1, 2011.

[276] T. Y. Reddy, V. R. Reddy, and V. Anbumozhi, "Physiological responses of groundnut

(Arachis hypogaea L.) to drought stress and its amelioration: A review," *Acta Agronomica Hungarica*, vol. 51, no. 2, pp. 205–227, 2005.

[277] V. Subramaniam and M. Maheshwari, "Physiological response of groundnut to water stress," *indian journal of plant physiology*, vol. 2, pp. 130–135, 1990.

[278] C. Holbrook, C. K. Kvien, K. S. Rucker, D. Wilson, and J. Hook, "Preharvest Aflatoxin Contamination in Drought-Tolerant and Drought-Intolerant Peanut Genotypes," *Peanut Science*, vol. 27, no. 2, pp. 45–48, 2000.

[279] D. M. Kambiranda and H. K. N. Vasanthaiah, "Impact of Drought Stress on Peanut (Arachis hypogaea L .) Productivity and Food Safety," 2011.

[280] R. Venuprasad, H. R. Lafitte, and G. N. Atlin, "Response to direct selection for grain yield under drought stress in rice," *Crop Science*, vol. 47, no. 1, pp. 285–293, 2007.

[281] R. Lafitte, "Relationship between leaf relative water content during reproductive stage water deficit and grain formation in rice," *Field Crops Research*, vol. 76, no. 2-3, pp. 165–174, 2002.

[282] P. Pimratch, P. Jogloy, N. Vorasoot, B. Toomsan, P. Patanothai, and C. Holbrook, "Relationship between Biomass Production and Nitrogen Fixation under Drought-Stress Conditions in Peanut Genotypes with Different Levels of Drought Resistance," *Journal of Agronomy and Crop Science*, vol. 194, no. 1, p. 15.25, 2007.

[283] B. Efeoglu, Y. Ekmekci, and N. Cicek, "Physiological responses of 3 maize cultivars to drought stress and recovery," *South African Journal of Botany*, vol. 75, no. 1, pp. 34–42, 2009.

[284] J. Flexas, J. Bota, F. Loreto, G. Cornic, and T. D. Sharkey, "Diffusive and Metabolic Limitations to Photosynthesis under Drought and Salinity in C3 Plants," *Plant Biology*, vol. 6, no. 3, pp. 269–279, 2004.

[285] M. M. Chaves, J. P. Maroco, and J. S. Pereira, "Understanding plant responses to drought - From genes to the whole plant," *Functional Plant Biology*, vol. 30, no. 3, pp. 239–264, 2003.

[286] A. L. Abdulai, *Morphological and Physiological Responses of Sorghum (Sorghum bicolor L . Moench) to Different Patterns of Drought Master of Science*. PhD thesis, The University of Bonn, 2005.

[287] W. H. Sun, Y. Y. Wu, X. Y. Wen, S. J. Xiong, H. G. He, Y. Wang, and G. Q. Lu, "Different mechanisms of photosynthetic response to drought stress in tomato and violet orychophragmus," *Photosynthetica*, vol. 54, no. 2, pp. 226–233, 2016.

[288] M. M. Chaves, J. Flexas, and C. Pinheiro, "Photosynthesis under drought and salt stress: Regulation mechanisms from whole plant to cell," *Annals of Botany*, vol. 103, no. 4, pp. 551–560, 2009.

[289] W. Davies and J. Zhan, "ROOT SIGNALS AND THE REGULATION OF GROWTH AND DEVELOPMENT OF PLANTS IN DRYING SOIL," *ANNUAL REVIEW OF PLANT PHYSIOLOGY AND PLANT MOLECULAR BIOLOGY*, vol. 42, pp. 55–76, 1991.

[290] H. M. Iskandar, R. E. Casu, A. T. Fletcher, S. Schmidt, J. Xu, D. J. Maclean, J. M. Manners, and G. D. Bonnett, "Identification of drought-response genes and a study of their expression during sucrose accumulation and water deficit in sugarcane culms," *BMC Plant Biology*, vol. 11, no. 1, p. 12, 2011.

[291] N. KUMAR, A. S. NANDWAL, R. S. WALDIA, S. SINGH, S. DEVI, K. D. SHARMA, and A. KUMAR, "Drought Tolerance in Chickpea As Evaluated By Root Characteristics, Plant Water Status, Membrane Integrity and Chlorophyll Fluorescence Techniques," *Experimental Agriculture*, vol. 48, no. 03, pp. 378–387, 2012.

[292] G. Thiyagarajan, D. Rajakumar, R. Kumaraperumal, and M. Manikandan, "Physiological responses of groundnut (Arachis hypogaea L .) varieties to drought stress," *ARCC*, vol. 30, no. 3, pp. 192–198, 2009.

[293] A. Bhagsari and R. Brown, "Photosynthesis in Peanut (Arachis) Genotypes," *Peanut Science*, vol. 3, pp. 1–5, 1976.

[294] M. Sarani, M. Namrudi, S. M. Hashemi, and M. M. Raoofi, "The effect of drought stress on chlorophyll content, root growth, glucosinolate and proline in crop plants," *International Journal of Farming and Allied Sciences*, vol. 3, no. 9, pp. 994–997, 2014.

[295] D. Huang, W. Wu, S. R. Abrams, and A. J. Cutler, "The relationship of drought-related gene expression in Arabidopsis thaliana to hormonal and environmental factors," *Journal of Experimental Botany*, vol. 59, no. 11, pp. 2991–3007, 2008.

[296] L. Pan, X. Zhang, J. Wang, X. Ma, M. Zhou, L. Huang, G. Nie, P. Wang, Z. Yang, and J. Li, "Transcriptional profiles of drought-related genes in modulating metabolic processes and antioxidant defenses in lolium multiflorum," *Frontiers in Plant Science*, vol. 7, pp. 1–15, 2016.

[297] A. Fracasso, L. M. Trindade, and S. Amaducci, "Drought stress tolerance strategies revealed by RNA-Seq in two sorghum genotypes with contrasting WUE," *BMC Plant Biology*, vol. 16, p. 115, 2016.

[298] M. Anwar Hossain, B. Biswavas Misra, M. Seki, S. Rasheed, K. Bashir, A. Matsui, and M. Tanaka, "Transcriptomic Analysis of Soil-Grown Arabidopsis thaliana Roots and Shoots in Response to a Drought Stress," *Frontiers in Plant Science Front. Plant Sci*, vol. 7, no. 7, pp. 1803389–180, 2016.

[299] N. J. Atkinson, C. J. Lilley, and P. E. Urwin, "Identification of genes involved in the

response of Arabidopsis to simultaneous biotic and abiotic stresses," *Plant Physiology*, vol. 162, no. 4, pp. 2028–2041, 2013.

[300] K. Shinozaki, K. Yamaguchi-Shinozaki, and M. Seki, "Regulatory network of gene expression in the drought and cold stress responses," *Current Opinion in Plant Biology*, vol. 6, no. 5, pp. 410–417, 2003.

[301] K. Shinozaki and K. Yamaguchi-Shinozaki, "Gene networks involved in drought stress response and tolerance," *Journal of Experimental Botany*, vol. 58, no. 2, pp. 221–227, 2007.

[302] K. Yamaguchi-Shinozaki and K. Shinozaki, "Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor.," *Novartis Foundation Symposium*, vol. 236, pp. 176–186, 2001.

[303] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under depencency," *The Annals of Statistics*, vol. 29, no. 4, pp. 1165–1188, 2001.

[304] F. WICKELMAIER, "An introduction to MDS," 2003.

[305] D. Golldack, C. Li, H. Mohan, and N. Probst, "Tolerance to drought and salt stress in plants: Unraveling the signaling networks," *Frontiers in Plant Science*, vol. 5, p. 151, 2014.

[306] L. Zhang, C. Zhang, P. Wu, Y. Chen, M. Li, H. Jiang, and G. Wu, "Global analysis of gene expression profiles in physic nut (Jatropha curcas L.) seedlings exposed to salt stress," *PLoS ONE*, vol. 9, no. 5, pp. 1–14, 2014.

[307] S. Bak, F. Beisson, G. Bishop, B. Hamberger, R. Höfer, S. Paquette, and D. Werck-Reichhart, "Cytochromes P450," *The Arabidopsis Book*, vol. 9, 2011.

[308] T. Hadiarto and L. S. P. Tran, "Progress studies of drought-responsive genes in rice," *Plant Cell Reports*, vol. 30, no. 3, pp. 297–310, 2011.

[309] L. Rizhsky, H. Liang, J. Shuman, V. Shulaev, S. Davletova, R. Mittler, Rizhsky L., Liang H., Shuman J., Shulaev V., Davletova S., and M. R., "When Defense Pathways Collide. The Response of Arabidopsis to a Combination of Drought and Heat Stress," *Plant Physiology*, vol. 134, pp. 1683–1696, 2004.

[310] Y. Fang and L. Xiong, "General mechanisms of drought response and their application in drought resistance improvement in plants," *Cellular and Molecular Life Sciences*, vol. 72, pp. 673–689, 2015.

[311] K. Nakashima, K. Yamaguchi-Shinozaki, and K. Shinozaki, "The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat," *Frontiers in Plant Science*, vol. 5, p. 170, 2014.

[312] M. H. Cruz de Carvalho, "Drought stress and reactive oxygen species: Production, scavenging and signaling.," *Plant signaling & behavior*, vol. 3, no. 3, pp. 156–65, 2008.

[313] B. V. Vardhini and N. a. Anjum, "Brassinosteroids make plant life easier under abiotic stresses mainly by modulating major components of antioxidant defense system," *Frontiers in Environmental Science*, vol. 2, pp. 1–16, 2015.

[314] E. Remy and P. Duque, "Beyond cellular detoxification: a plethora of physiological roles for MDR transporter homologs in plants," *Frontiers in Physiology*, vol. 5, no. 201, pp. 1–10, 2014.

[315] T. Umezawa, M. Fujita, Y. Fujita, K. Yamaguchi-Shinozaki, and K. Shinozaki, "Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future," *Current Opinion in Biotechnology*, vol. 17, no. 2, pp. 113–122, 2006.

[316] E. A. C. MacRobbie, "Evidence for a role for protein tyrosine phosphatase in the control of ion release from the guard cell vacuole in stomatal closure.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 18, pp. 11963–8, 2002.

[317] N. Focks and C. Benning, "wrinkled1: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism," *Plant Physiology*, vol. 118, no. 1, pp. 91–101, 1998.

[318] Z. Wang, M. Gerstein, and M. Snyder, "RNA-Seq: a revolutionary tool for transcriptomics," *Nature Reviews Genetics*, vol. 10, no. 1, pp. 57–63, 2009.

[319] M. A. Dillies, A. Rau, J. Aubert, C. Hennequet-Antier, M. Jeanmougin, N. Servant, C. Keime, N. S. Marot, D. Castel, J. Estelle, G. Guernec, B. Jagla, L. Jouneau, D. Laloë, C. Le Gall, B. Schaëffer, S. Le Crom, M. Guedj, and F. Jaffrézic, "A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis," *Briefings in Bioinformatics*, vol. 14, no. 6, pp. 671–683, 2013.

[320] Y. Lin, K. Golovnina, Z.-X. Chen, H. N. Lee, Y. L. S. Negron, H. Sultana, B. Oliver, and S. T. Harbison, "Comparison of normalization and differential expression analyses using RNA-Seq data from 726 individual Drosophila melanogaster.," *BMC Genomics*, vol. 17, no. 1, p. 28, 2016.

[321] Y. Benjamini and Y. Hochberg, "Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 Published by :," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

[322] E. Kristiansson, T. Österlund, L. Gunnarsson, G. Arne, D. G. Joakim Larsson, O. Nerman, T. Barrett, D. Troup, S. Wilhite, P. Ledoux, C. Evangelista, I. Kim, M. Tomashevsky, K. Marshall, K. Phillippy, P. Sherman, R. Muertter, M. Holko, O. Ayanbule, A. Yefanov, A. Soboleva, H. Parkinson, U. Sarkans, N. Kolesnikov, N. Abeygunawardena, T. Burdett, M. Dylag, I. Emam, A. Farne, E. Hastings, E. Holloway, N. Kurbatova, M. Lukk, J. Malone, R. Mani, E. Pilicheva, G. Rustici, A. Sharma, E. Williams, T. Adamusiak, M. Brandizi, N. Sklyar, A. Brazma, J. Raser, E. O'Shea, Y. Taniguchi, P. Choi, G. Li, H. Chen, J. M. Babu, A. Emili, X. Xie, D. Allison, X. Cui, G. Page, M. Sabripour, E. Kristiansson, A. Sjögren, M. Rudemo, O. Nerman, M. Consortium, W. Kuo, F. Liu, J. Trimarchi, C. Punzo, M. Lombardi, J. Sarang, M. Whipple, M. Maysuria, K. Serikawa, S. Lee, D. McCrann, J. Kang, J. Shearstone, J. Burke, D. Park, X. Wang, T. Rector, P. Ricciardi-Castagnoli, S. Perrin, S. Choi, R. Bumgarner, J. Kim, G. III, M. Freeman, B. Seed, R. Jensen, G. Church, E. Hovig, C. Cepko, P. Park, L. Ohno-Machado, T. Jenssen, U. Ala, R. Piro, E. Grassi, C. Damasco, L. Silengo, M. Oti, P. Provero, F. D. Cunto, E. Segal, N. Friedman, N. Kaminski, A. Regev, D. Koller, A. Sweet-Cordero, S. Mukherjee, A. You, J. Roix, C. Ladd-Acosta, J. Mesirov, T. Golub, T. Jacks, J. Miller, S. Horvath, D. Geschwind, A. Rasche, H. Al-Hasani, R. Herwig, F. Marques, A. Campain, Y. Yang, B. Morris, I. Ginis, Y. Luo, T. Miura, S. Thies, R. Brandenberger, S. Gerecht-Nir, M. Amit, A. Hoke, M. Carpenter, J. Itskovitz-Eldor, M. Rao, F. Pan, C. Chiu, S. Pulapura, M. Mehan, J. Nunez-Iglesias, K. Zhang, K. Kamath, M. Waterman, C. Finch, X. Zhou, J. de Magalhaes, J. Curado, G. Church, L. Gunnarsson, E. Kristiansson, C. Rutgersson, J. Sturve, J. Fick, L. Förlin, D. Larsson, L. Gunnarsson, E. Kristiansson, L. Förlin, O. Nerman, D. Larsson, C. Ung, S. Lam, M. Hiaing, C. Winata, S. Korzh, S. Mathavan, Z. Gong, D. Kristensen, Y. Wolf, A. Mushegian, E. Koonin, S. Ohno, Z. Gu, S. Rifkin, K. White, W. Li, L. Huminiecki, K. Wolfe, M. Lynch, V. Katju,

A. S. R, M. Robinson-Rechavi, X. Chen, J. Zhang, R. Fisher, P. Hu, C. Greenwood, J. Beyene, A. Campain, Y. Yang, G. Tseng, D. Ghosh, E. Feingold, J. Stuart, E. Segal, D. Koller, S. Kim, H. Le, Z. Oltvai, Z. Bar-Joseph, P. Cahan, A. Ahmad, H. Burke, S. Fu, Y. Lai, L. Florea, N. Dharker, T. Kobrinski, P. Kale, T. McCaffrey, J. Newman, A. Weiner, Y. Lu, R. Rosenfeld, Z. Bar-Joseph, Y. Lu, S. Mahony, P. Benos, R. Rosenfeld, I. Simon, L. Breeden, Z. Bar-Joseph, Y. Lu, R. Rosenfeld, G. Nau, Z. Bar-Joseph, A. Ramasamy, A. Mondry, C. Holmes, D. Altman, E. Sayers, T. Barrett, D. Benson, E. Bolton, S. Bryant, K. Canese, V. Chetvernin, D. Church, M. DiCuccio, S. Federhen, M. Feolo, I. Fingerman, L. Geer, W. Helmberg, Y. Kapustin, D. D. Landsman, Z. Lu, T. Madden, T. Madej, D. Maglott, A. Miller, I. Mizrachi, J. Ostell, A. Panchenko, L. Phan, K. Pruitt, G. Schuler, E. Sequeira, S. Sherry, M. Shumway, K. Sirotkin, D. Slotta, A. Souvorov, G. Starchenko, T. Tatusova, L. Wagner, Y. Wang, W. Wilbur, E. Yaschenko, J. Ye, F. Chen, A. Mackey, C. Jr, D. Roos, A. Berglund, E. Sjölund, G. Östlund, E. Sonnhammer, L. Li, C. Jr, D. Roos, R. Grützmann, H. Boriss, O. Ammerpohl, J. Lüttges, H. Kalthoff, H. Schackert, G. Klöppel, H. Saeger, C. Pilarsky, K. Richter, M. Haslbeck, J. Buchner, M. Feder, G. Hoffman, J. Laramie, T. Chung, B. Brownstein, G. Cobb, B. Vallant, S. Anderssson, H. Brown-Borg, H. Ren, S. Kersten, S. Jonnalagadda, R. Srinivasan, J. Corton, J. Sorensen, M. Nielsen, M. Kruhoffer, J. Justesen, V. Loeschcke, W. Hu, G. Hu, B. Han, J. Kilian, D. Whitehead, J. Horak, D. Wanke, S. Weinl, O. Batistic, C. D'Angelo, E. Bornberg-Bauer, J. Kudla, K. Harter, D. Chen, M. Toone, J. Mata, R. Lyne, G. Burns, K. Kivinen, A. Brazama, N. Jones, J. Bahler, D. Berry, A. Gasch, C. Purdom, P. Hardiman, V. Bye, N. Eno, C. Tyler, J. Sumpter, D. Larsson, M. Adolfsson-Erici, J. Parkkonen, M. Pettersson, A. Berg, P. Olsson, L. Förlin, E. Routledge, D. Sheahan, C. Desbrow, G. Brighty, M. Waldock, J. Sumpter, S. Jobling, S. Coey, J. Whitmore, D. Kime, K. van Look, B. McAllister, N. Beresford, A. AC, G. Brighty, C. Tyler, J. Sumpter, J. Sump-

ter, S. Jobling, E. Thomas-Jones, K. Thorpe, N. Harrison, G. Thomas, C. Morris, T. Hutchinson, S. Woodhead, C. Tyler, O. Carnevali, F. Maradonna, M. de Wit, D. Keil, K. van der Ven, S. Vandamme, E. Witters, W. Coen, A. Arukwe, A. Goksøyr, A. Davis, B. King, S. Mockus, C. Murphy, C. Saraceni-Richards, M. Rosenstein, T. Wiegers, C. Mattingly, T. Williams, A. Diab, S. George, V. Sabine, J. Chipman, F. Geoghegan, I. Katsiadaki, T. Williams, J. Chipman, C. Martyniuka, E. Gerrie, J. Popesku, M. Ekker, V. Trudeau, S. Tilton, S. Givan, C. Pereira, G. Bailey, D. Williams, M. Sárvári, E. Hrabovszky, T. Kalló, O. Galamb, N. Solymosi, T. Likó, B. Molnár, K. Tihanyi, Z. Szombathelyi, Z. Liposits, J. Kwekel, L. Burgoon, J. Burt, J. Harkema, T. Zacharewski, L. Henríquez-Hernández, A. Flores-Morales, R. Santana-Farré, M. Axelson, P. Nilsson, G. Norstedt, L. Fernández-Pérez, R. Xu, X. Li, F. Chen, A. Mackey, J. Vermunt, D. Roos, E. Kristiansson, A. Sjögren, M. Rudemo, O. Nerman, L. Klebanov, C. Jordan, A. Yakovlev, A. Sjögren, E. Kristiansson, M. Rudemo, O. Nerman, E. Forbes, P. Calow, N. Isaac, S. Turvey, B. Collen, C. Waterman, J. Baillie, I. Good, D. Bhoj, K. Schiefermayr, B. Bolstad, R. Irizarry, M. øAstrand, T. Speed, Y. Yang, S. Dudoit, P. Luu, D. Lin, V. Peng, J. Ngai, T. Speed, G. Smythe, A. Alexa, J. Rahnenführer, and T. Lengauer, "A novel method for cross-species gene expression analysis," *BMC Bioinformatics*, vol. 14, no. 1, p. 70, 2013.

[323] K. Apel and H. Hirt, "REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction," *Annual Review of Plant Biology*, vol. 55, pp. 373–399, apr 2004.

[324] N. Smirnoff, "Plant resistance to environmental stress," *Current Opinion in Biotechnology*, vol. 9, no. 2, pp. 214–219, 1998.

[325] A. R. Bhardwaj, G. Joshi, B. Kukreja, V. Malik, P. Arora, R. Pandey, R. N. Shukla, K. G. Bankar, S. Katiyar-Agarwal, S. Goel, A. Jagannath, A. Kumar, and M. Agarwal,

"Global insights into high temperature and drought stress regulated genes by RNA-Seq in economically important oilseed crop Brassica juncea.," *BMC Plant Biology*, vol. 15, no. 1, p. 9, 2015.

[326] W. Chen, Q. Yao, G. B. Patil, G. Agarwal, R. K. Deshmukh, L. Lin, B. Wang, Y. Wang, S. J. Prince, L. Song, D. Xu, Y. C. An, B. Valliyodan, R. K. Varshney, and H. T. Nguyen, "Identification and Comparative Analysis of Differential Gene Expression in Soybean Leaf Tissue under Drought and Flooding Stress Revealed by RNA-Seq," *Frontiers in Plant Science*, vol. 7, p. 1044, 2016.

[327] V. Padmanabhan, D. M. Dias, and R. J. Newton, "Expression analysis of a gene family in loblolly pine (Pinus taeda L.) induced by water deficit stress," *Plant molecular biology*, vol. 35, no. 6, pp. 801–807, 1997.

[328] C. Canel, J. N. Bailey-Serres, and M. L. Roose, "Pummelo fruit transcript homologous to ripening-induced genes.," *Plant physiology*, vol. 108, no. 3, pp. 1323–4, 1995.

[329] Z. Peleg and E. Blumwald, "Hormone balance and abiotic stress tolerance in crop plants," *Current Opinion in Plant Biology*, vol. 14, no. 3, pp. 290–295, 2011.

[330] X. Yao, W. Xiong, T. Ye, and Y. Wu, "Overexpression of the aspartic protease ASPG1 gene confers drought avoidance in Arabidopsis," *Journal of Experimental Botany*, vol. 63, no. 7, pp. 2579–2593, 2012.

[331] J. Iyama and Y. Murata, "1. Studies on the photosynthesis in upland field crops 2. Relationships between the soil moisture and photosynthesis of some upland crops and rice plant.," in *Crop Science*, 1961.

[332] J. I. Schroeder, J. M. Kwak, and G. J. Allen, "Guard cell abscisic acid signalling and engineering drought hardiness in plants.," *Nature*, vol. 410, no. 6826, pp. 327–330, 2001.

[333] B. Hong, S. J. Uknes, and T.-h. D. Ho, "Cloning and characterization of a cDNA encoding a mRNA rapidly-induced by ABA in barley aleurone layers," *Plant Molecular Biology*, vol. 11, no. 4, pp. 495–506, 1988.

[334] C. Yanhui, Y. Xiaoyuan, H. Kun, L. Meihua, L. Jigang, G. Zhaofeng, L. Zhiqiang, Z. Yunfei, W. Xiaoxiao, Q. Xiaoming, S. Yunping, Z. Li, D. Xiaohui, L. Jingchu, D. Xing-Wang, C. Zhangliang, G. Hongya, and Q. Li-Jia, "The MYB Transcription Factor Superfamily of Arabidopsis: Expression Analysis and Phylogenetic Comparison with the Rice MYB Family," *Plant Molecular Biology*, vol. 60, no. 1, pp. 107–124, 2006.

[335] L. P. Tran, K. Nakashima, Y. Sakuma, S. D. Simpson, Y. Fujita, K. Maruyama, M. Fujita, M. Seki, K. Shinozaki, and K. Yamaguchi-shinozaki, "Isolation and Functional Analysis of Arabidopsis Stress-Inducible NAC Transcription Factors That Bind to a Drought-Responsive cis -Element in the early responsive to dehydration stress 1 Promoter," *The Plant Cell*, vol. 16, no. 9, pp. 2481–2498, 2004.

[336] J. You, W. Zong, H. Hu, X. Li, J. Xiao, and L. Xiong, "A STRESS-RESPONSIVE NAC1-Regulated Protein Phosphatase Gene Rice Protein Phosphatase18 Modulates Drought and Oxidative Stress Tolerance through Abscisic Acid-Independent Reactive Oxygen Species Scavenging in Rice.," *Plant physiology*, vol. 166, no. 4, pp. 2100–14, 2014.

[337] R. Zhong, E. A. Richardson, and Z.-h. Ye, "The MYB46 Transcription Factor Is a Direct Target of SND1 and Regulates Secondary Wall Biosynthesis in Arabidopsis The MYB46 Transcription Factor Is a Direct Target of SND1 and Regulates Secondary Wall Biosynthesis in Arabidopsis," vol. 19, no. 9, pp. 2776–2792, 2007.

[338] T. Kamiya, M. Borghi, P. Wang, J. M. C. Danku, L. Kalmbach, P. S. Hosmani, S. Naseer, T. Fujiwara, N. Geldner, and D. E. Salt, "The MYB36 transcription factor

orchestrates Casparian strip formation.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 33, pp. 10533–8, 2015.

[339] G.-H. He, J.-Y. Xu, Y.-X. Wang, J.-M. Liu, P.-S. Li, M. Chen, Y.-Z. Ma, and Z.-S. Xu, "Drought-responsive WRKY transcription factor genes TaWRKY1 and TaWRKY33 from wheat confer drought and/or heat resistance in Arabidopsis," *BMC Plant Biology*, vol. 16, no. 1, p. 116, 2016.

[340] S. B. Lee and M. C. Suh, "Recent advances in cuticular wax biosynthesis and its regulation in arabidopsis," *Molecular Plant*, vol. 6, no. 2, pp. 246–249, 2013.

[341] G. C. Briggs, C. F. Mouchel, and C. S. Hardtke, "Characterization of the plant-specific BREVIS RADIX gene family reveals limited genetic redundancy despite high sequence conservation.," *Plant Physiology*, vol. 140, no. 4, pp. 1306–16, 2006.

[342] C. F. Mouchel, G. C. Briggs, and C. S. Hardtke, "Natural genetic variation in Arabidopsis identifies BREVIS RADIX, a novel regulator of cell proliferation and elongation in the root.," *Genes & Development*, vol. 18, no. 6, pp. 700–714, 2004.

[343] H. Abe, T. Urao, T. Ito, M. Seki, and K. Shinozaki, "Transcriptional Activators in Abscisic Acid Signaling," *Society*, vol. 15, pp. 63–78, 2003.

[344] A. Uzman, "Molecular biology of the cell (4th ed.): Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P.," *Biochemistry and Molecular Biology Education*, vol. 31, no. 4, pp. 212–214, 2003.

[345] W. K. Ho, H. H. Chai, P. Kendabie, N. S. Ahmad, J. Jani, F. Massawe, A. Kilian, and S. Mayes, "Integrating genetic maps in bambara groundnut [Vigna subterranea (L) Verdc.] and their syntenic relationships among closely related legumes," *BMC Genomics*, vol. 18, no. 1, p. 192, 2017.

[346] S. Cunhua, D. Wei, C. Xiangling, X. Xinna, Z. Yahong, S. Dong, and S. Jianjie, "The effects of drought stress on the activity of acid phosphatase and its protective enzymes in pigweed leaves," *African Journal of Biotechnology*, vol. 9, no. 6, pp. 825–833, 2010.

[347] T. J. Close, A. A. Kortt, and P. M. Chandler, "A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn," *Plant Molecular Biology*, vol. 13, no. 1, pp. 95–108, 1989.

[348] A. Ebert, "Potential of Underutilized Traditional Vegetables and Legume Crops to Contribute to Food and Nutritional Security, Income and More Sustainable Production Systems," *Sustainability*, vol. 6, no. 1, pp. 319–335, 2014.

[349] S. Padulosi, T. Hodgkin, J. Williams, and N. Haq, "Underutilised crops: trends, challenges and opportunities in the 21st Century," *Managing plant genetic diversity*, p. 323, 2002.

[350] Food and Agriculture Organization of the United Nations (FAO), "Biodiversity for Food and Agriculture â Contributing to food security and sustainable in a changing world,," 2011.

[351] Food and Agriculture Organization of the United Nations (FAO), "Harvesting nature's diversity," 1993.

[352] Food and Agriculture Organization of the United Nations (FAO), "How to feed the world in 2050," 2009.

[353] Food and Agriculture Organization of the United Nations (FAO), "The Commission on Genetic Resources for Food and Agriculture (CGRFA) Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture (PGRFA)," 2012.

[354] R. E. Black, L. H. Allen, Z. A. Bhutta, L. E. Caulfield, M. de Onis, M. Ezzati, C. Mathers, and J. Rivera, "Maternal and child undernutrition: global and regional

exposures and health consequences," *The Lancet*, vol. 371, no. 9608, pp. 243–260, 2008.

[355] Micronutrient Initiative, "Investing in the Future: A united call to action on vitamin and mineral deficiencies," 2009.

[356] B. EfeoÇ§lu, Y. Ekmekçi, and N. Çiçek, "Physiological responses of three maize cultivars to drought stress and recovery," *South African Journal of Botany*, vol. 75, no. 1, pp. 34–42, 2009.

[357] J. B. Agarwal PK, "Transcription factors in plants and ABA dependent and\rindependent abiotic stress signaling," *Biologia Plantarum*, vol. 54, no. 54, pp. : 201–212., 2010.

[358] L. Rizhsky, H. Liang, and R. Mittler, "The combined effect of drought stress and heat shock on gene expression in tobacco.," *Plant Physiology*, vol. 130, no. 3, pp. 1143–1151, 2002.

[359] H. Liu, M. A. R. F. Sultan, X. L. Liu, J. Zhang, F. Yu, and H. X. Zhao, "Physiological and comparative proteomic analysis reveals different drought responses in roots and leaves of drought-tolerant wild wheat (Triticum boeoticum)," *PLoS ONE*, vol. 10, no. 4, pp. 1–29, 2015.

[360] N. Tuteja, "Abscisic Acid and abiotic stress signaling.," *Plant signaling & behavior*, vol. 2, no. 3, pp. 135–8, 2007.

[361] M. M. Chaves and M. M. Oliveira, "Mechanisms underlying plant resilience to water deficits: Prospects for water-saving agriculture," *Journal of Experimental Botany*, vol. 55, no. 407, pp. 2365–2384, 2004.

[362] V. D. Fernando and D. F. Schroeder, "Role of ABA in Arabidopsis Salt, Drought, and

Desiccation Tolerance," in *Abiotic and Biotic Stress in Plants - Recent Advances and Future Perspectives* (A. Shankar, ed.), InTech, 2016.

[363] S. P. Hazen, M. S. Pathan, A. Sanchez, I. Baxter, M. Dunn, B. Estes, H.-S. Chang, T. Zhu, J. A. Kreps, and H. T. Nguyen, "Expression profiling of rice segregating for drought tolerance QTLs using a rice genome array," *Functional & Integrative Genomics*, vol. 5, no. 2, pp. 104–116, 2005.

[364] H. Kaur, P. Verma, B. P. Petla, V. Rao, S. C. Saxena, and M. Majee, "Ectopic expression of the ABA-inducible dehydration-responsive chickpea l-myo-inositol 1-phosphate synthase 2 (CaMIPS2) in Arabidopsis enhances tolerance to salinity and dehydration stress," *Planta*, vol. 237, no. 1, pp. 321–335, 2013.

[365] E. Baldoni, A. Genga, and E. Cominelli, "Plant MYB transcription factors: Their role in drought response mechanisms," *International Journal of Molecular Sciences*, vol. 16, no. 7, pp. 15811–15851, 2015.

[366] M. Fujita, Y. Fujita, K. Maruyama, M. Seki, K. Hiratsu, M. Ohme-Takagi, L. S. P. Tran, K. Yamaguchi-Shinozaki, and K. Shinozaki, "A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway," *Plant Journal*, vol. 39, no. 6, pp. 863–876, 2004.

[367] C. A. Dezar, G. M. Gago, D. H. González, and R. L. Chan, "Hahb-4, a sunflower homeobox-leucine zipper gene, is a developmental regulator and confers drought tolerance to Arabidopsis thaliana plants," *Transgenic Research*, vol. 14, no. 4, pp. 429–440, 2005.

[368] K. Shinozaki and K. Yamaguchi-Shinozaki, "Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways," *Curr Opin Plant Biol*, vol. 3, no. 3, pp. 217–223, 2000.

[369] X.-Y. Huang, D.-Y. Chao, J.-P. Gao, M.-Z. Zhu, M. Shi, and H.-X. Lin, "A previously unknown zinc finger protein, DST, regulates drought and salt tolerance in rice via stomatal aperture control.," *Genes & Development*, vol. 23, no. 15, pp. 1805–1817, 2009.

[370] E. Cominelli, M. Galbiati, A. Vavasseur, L. Conti, T. Sala, M. Vuylsteke, N. Leonhardt, S. L. Dellaporta, and C. Tonelli, "A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance," *Current Biology*, vol. 15, no. 13, pp. 1196–1200, 2005.

[371] C. V. Ha, D. T. Le, R. Nishiyama, Y. Watanabe, S. Sulieman, U. T. Tran, K. Mochida, N. V. Dong, K. Yamaguchi-Shinozaki, K. Shinozaki, and L.-S. P. Tran, "The auxin response factor transcription factor family in soybean: genome-wide identification and expression analyses during development and water stress.," *DNA research : an international journal for rapid publication of reports on genes and genomes*, vol. 20, no. 5, pp. 511–24, 2013.

[372] H. Jung, D. K. Lee, Y. D. Choi, and J. K. Kim, "OsIAA6, a member of the rice Aux/IAA gene family, is involved in drought tolerance and tiller outgrowth," *Plant Science*, vol. 236, pp. 304–312, 2015.

[373] Y. M. Kim, Y. J. Han, O. J. Hwang, S. S. Lee, A. Y. Shin, S. Y. Kim, and J. I. Kim, "Overexpression of arabidopsis translationally controlled tumor protein gene AtTCTP enhances drought tolerance with rapid ABA-induced stomatal closure," *Molecules and Cells*, vol. 33, no. 6, pp. 617–626, 2012.

[374] H. Min, C. Chen, S. Wei, X. Shang, M. Sun, R. Xia, X. Liu, D. Hao, H. Chen, and Q. Xie, "Identification of Drought Tolerant Mechanisms in Maize Seedlings Based on Transcriptome Analysis of Recombination Inbred Lines," *Frontiers in Plant Science*, vol. 7, p. 1080, 2016.

[375] B. Valliyodan and H. T. Nguyen, "Understanding regulatory networks and engineering for enhanced drought tolerance in plants," *Current Opinion in Plant Biology*, vol. 9, no. 2, pp. 189–195, 2006.

[376] K. Ifuku, Y. Yamamoto, T.-A. Ono, S. Ishihara, and F. Sato, "PsbP protein, but not PsbQ protein, is essential for the regulation and stabilization of photosystem II in higher plants.," *Plant Physiology*, vol. 139, no. 3, pp. 1175–84, 2005.

[377] N. Tayeh, G. Aubert, M.-L. Pilet-Nayel, I. Lejeune-Hénaut, T. D. Warkentin, and J. Burstin, "Genomic Tools in Pea Breeding Programs: Status and Perspectives," *Frontiers in Plant Science*, vol. 6, pp. 1–13, 2015.

[378] P. J. Hiremath, A. Farmer, S. B. Cannon, J. Woodward, H. Kudapa, R. Tuteja, A. Kumar, A. BhanuPrakash, B. Mulaosmanovic, N. Gujaria, L. Krishnamurthy, P. M. Gaur, P. B. KaviKishor, T. Shah, R. Srinivasan, M. Lohse, Y. Xiao, C. D. Town, D. R. Cook, G. D. May, and R. K. Varshney, "Large-scale transcriptome analysis in chickpea (Cicer arietinum L.), an orphan legume crop of the semi-arid tropics of Asia and Africa," *Plant Biotechnology Journal*, vol. 9, no. 8, pp. 922–931, 2011.

[379] A. C. Nica and E. T. Dermitzakis, "Expression quantitative trait loci: present and future," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 368, no. 1620, 2013.

[380] N. R. Street, O. Skogström, A. Sjödin, J. Tucker, M. Rodríguez-Acosta, P. Nilsson, S. Jansson, and G. Taylor, "The genetics and genomics of the drought response in Populus," *The Plant Journal*, vol. 48, no. 3, pp. 321–341, 2006.

[381] A. M. Pérez-de Castro, S. Vilanova, J. Cañizares, L. Pascual, J. M. Blanca, M. J. Díez, J. Prohens, and B. Picó, "Application of genomic tools in plant breeding.," *Current genomics*, vol. 13, no. 3, pp. 179–95, 2012.

[382] A. Korte and A. Farlow, "The advantages and limitations of trait analysis with GWAS: a review," *Plant Methods*, vol. 9, no. 1, p. 29, 2013.

[383] P. M. Visscher, M. A. Brown, M. I. McCarthy, and J. Yang, "Five Years of GWAS Discovery," *The American Journal of Human Genetics*, vol. 90, pp. 7–24, jan 2012.