Transcriptomic analysis of a *Sorghum bicolor* landrace identifies a role for beta-alanine betaine biosynthesis in drought tolerance

Natrisha Devnarain[1,2*], Bridget G. Crampton[2,3], Nicky Olivier[2,3], Chris van der Westhuyzen[1], John V. W. Becker[4], Martha M. O'Kennedy[1].

[1]*Biosciences, Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa*

[2]*Department of Plant and Soil Sciences, University of Pretoria, Pretoria, South Africa*

[3]*Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa*

[4]*African Centre for Gene Technologies (ACGT), Pretoria, South Africa*

*Corresponding author,
N. Devnarain
Department of Plant and Soil Sciences, University of Pretoria, Pretoria, South Africa
e-mail: natrisha.devnarain@gmail.com

Highlights

- Transcriptomic analysis of a South African sorghum landrace.
- Gene Ontology enrichment of 26 abiotic stress response genes.
- Role for β-alanine betaine identified in sorghum drought tolerance.

Abbreviations: MS, mild stress; MSC, mild stress control; RW, re-watered; RWC, re-watered control; SS, severe stress; SSC, severe stress control.

Abstract

Sorghum is indigenous to Africa and a remarkably drought tolerant cereal crop. In this study, the genetic response mechanisms involved in sorghum's tolerance to progressive water deficit and re-watering were investigated in a South African landrace (designated LR6), using cDNA microarrays comprising 35 899 transcript probes. Significant differential expression of 902 transcripts, including 128 transcripts with currently unknown functions, was altered in response to progressive water stress and re-watering. The modulated sorghum genes had homology to proteins involved in growth, regulation, and protection. Gene Ontology (GO) analysis identified significant enrichment of 26 genes involved in the 'response to abiotic stimulus' GO category during severe stress. The expression of two genes associated with beta(β)-alanine betaine biosynthesis was validated with quantitative RT-PCR. Importantly, the detection of β-alanine betaine in sorghum leaf extracts using NMR spectroscopy, and the significant increase in relative abundance during severe stress supports the microarray and qRT-PCR findings, thereby highlighting a role for β-alanine betaine biosynthesis in drought tolerance of sorghum. In future, crop improvement initiatives that consider metabolic engineering of the β-alanine betaine biosynthesis pathway should be explored.

Keywords: Beta-alanine betaine, Drought tolerance, Landrace, Sorghum, Transcriptomics

1. Introduction

Water is an essential resource for agriculture and its availability is dependent on climate variability. Water stress is one of the most important crop abiotic stresses and continues to impact agricultural productivity worldwide (Boyer 1982; Kholova et al. 2009). Drought is recognised as a complex phenomenon during which the accumulated effects of water scarcity over time may develop into a severe stress condition that can be devastating to crop yields (Wilhite 1993; Mishra and Singh 2010). Periodic El Niño episodes have influenced climate conditions by intensifying drought in certain regions and South Africa experienced its most severe drought period since 1904 due the 2015/2016 El Niño (Van Vuuren 2015; De Jager 2016).

The necessary resources for irrigation systems are largely unavailable in impoverished regions of the world and the selection of resilient cultivars through conventional crop breeding programmes is a slow process. In addition, screening for stress tolerance is secondary compared with crop yield and grain quality in breeding programmes (Cativelli et al. 2008; Dolferus et al. 2011). Therefore landraces and wild ancestors, which are often adapted to their indigenous climates, and exhibit better abiotic and biotic stress tolerance than commercial germplasm, are excluded due to yield traits (Newton et al. 2010; Dolferus et al. 2011). Camacho Villa et al. (2005) has defined a landrace as ‘a dynamic population(s) of a cultivated plant that has historical origin, distinct identity and lacks formal crop improvement, as well as often being genetically diverse, locally adapted and associated with traditional farming systems’. The important genetic diversity that can be contributed to agricultural germplasm by landraces has recently been realised and landraces are now being incorporated into conventional breeding programmes, as well as investigated for their genetic resources (Pinheiro De Carvalho et al. 2013).

Plant tolerance to abiotic stress is greatly influenced by inherent mechanisms of complex, multi-gene and multi-pathway responses. Therefore understanding the molecular interactions involved, as well as identification of signalling and gene networks that enable drought tolerance continue to be a challenge

for crop researchers (Agarwal et al. 2006; Dugas et al. 2011). Genomics technologies have been useful in deciphering specific abiotic stress responses in plants through platforms such as transcriptomics (Bohnert et al. 2006; Sreenivasulu et al. 2007). Intensive research has contributed to a better understanding of abiotic stress response in plants through *Arabidopsis thaliana*, and important agronomic crops such as maize, rice and wheat, to understand the interplay between gene expression and drought tolerance (Rabbani et al. 2003; Vincour and Altman 2005; Shinozaki and Yamaguchi-Shinozaki 2007; Dugas et al. 2011). Much knowledge can be gained by investigating other crop species which exhibit remarkable resilience in response to water deficit, such as sorghum [*Sorghum bicolor* (L.) Moench].

Sorghum is a grain crop indigenous to sub-Saharan Africa, and well adapted to water-limited and hot climates. It is related to important cereal crops belonging to the Poaceae family, such as rice, maize, wheat, barley, oats and millets; and is a close relative to sugarcane (Doggett 1988; Subudhi and Nguyen 2000). Worldwide, sorghum is the fifth most important grain crop with 63.9 million tonnes produced during 2016 (FAO 2018; http://faostat.fao.org). Investigation of sorghum landraces in particular, has great potential to elucidate key mechanisms of drought tolerance. Further, the sequencing of the sorghum genome by Paterson et al. (2009) has expedited molecular characterisation of sorghum and it has been proposed as a model C_4 cereal crop for genomic research due to its relatively small genome size (~730 Mb) (Paterson et al. 2009; Dugas et al. 2011).

Transcriptomic studies using microarrays have been used to investigate abiotic stress responses in some sorghum genotypes and revealed important findings (Buchanan et al. 2005; Johnson et al. 2014; Johnson et al. 2015). Buchanan et al. (2005) described the transcriptomic response to dehydration, salinity and ABA in cultivar BTx623. These researchers identified several osmotic stress inducible sorghum genes for the first time. Johnson et al. (2014) demonstrated evidence for both cross-talk and specificity to combined drought and heat stress with some aspects of the combined stress response unique compared with the individual stress in cultivar R16. A role for proline biosynthesis was identified in the stay-green trait in cultivar B35 (Johnson et al. 2015).

In this study, cDNA microarrays were used to characterise the molecular basis for drought tolerance in a South African sorghum landrace (designated LR6). Further, qRT-PCR and NMR spectroscopy were used to validate a role for β-alanine betaine biosynthesis in sorghum drought tolerance. Landrace LR6 was selected for transcriptomic analysis from a previous study (refer to Devnarain et al. 2016) during which the physiological responses to progressive water stress and re-watering was assessed at the reproductive growth stage of development.

2. Materials and Methods

2.1. Plant material and drought stress treatments

Sorghum plant material (landrace designated LR6) was obtained from a previous water stress investigation (refer to Devnarain et al. 2016 for further detail). Briefly, three treatments were imposed on the sorghum plants: mild stress (MS, four days of water deficit), severe stress (SS, six days of water deficit), and a re-watered (RW) treatment during which plants were harvested five hours after re-watering following seven days of water deficit. All treatments were complemented with well-watered controls. There were nine biological replicates for each treatment and control per genotype. During harvest, the second, fourth and sixth leaves of each plant were excised, combined and flash frozen in liquid nitrogen, and stored at -80^{o}C for transcriptomic and metabolite evaluations.

2.2. Transcriptomic analysis

2.2.1. Microarray design

Sorghum cDNA microarrays (4X44K format; Agilent Technologies Ltd) were designed by Genotypic Technologies Pvt LTD, India with sequence data from the Joint Genome Institute (JGI: http://genome.jgi.doe.gov/). Each array was comprised of probes for 35 899 transcripts in the sense orientation.

2.2.2. Sample preparation and hybridization

Total RNA was extracted from ground, frozen leaf material using the QIAzol Lysis Reagent and Qiagen RNeasy kit (Qiagen, Germany). Powdered leaf material (~50 mg) was combined with 1 ml QIAzol reagent in a 2 ml micro-centrifuge tube (QSP, USA), vortexed and incubated for 5 min at ambient temperature (24-28^{o}C). Chloroform (0.2 ml; Merck (Pty) Ltd., RSA) was added; tubes were vigorously shaken for 15 sec and incubated for 3 min at ambient temperature. Samples were centrifuged at 12 000 *g* for 10 min at 4^{o}C (Biofuge 13, Heraeus Instruments, Germany). The upper aqueous phase was combined with an equal volume of 70% ethanol (Merck (Pty) Ltd., RSA), gently mixed, then transferred onto RNeasy columns and centrifuged at 8000 *g* for 15 sec (Mikro 200, Hettich, Germany). Thereafter, total RNA was purified as per manufacturer's instructions. An on-column DNase I (Qiagen, Germany) digestion was included using 27 Kunitz units for 15 min. RNA was eluted with nuclease-free water (Bioline, USA). The $A_{260/280}$ and $A_{260/230}$ ratios were determined with 1 µl of sample using a NanoDrop ND-1000 (Thermo Scientific Inc., USA) spectrophotometer in order to assess the purity and concentration of RNA. An RNA 6000 Nano LabChip was used to evaluate the integrity of RNA with an Agilent 2100 Bioanalyzer (Agilent Technologies, USA).

There were three replicates for each treatment and control, with each representing equal amounts (5 µg each) of three pooled RNA samples. First strand complementary DNA (cDNA) was synthesised using 9 µg of pooled RNA with 1.5 µl oligo (dT) (2 µg/µl) and 6 µl Random primer stock (0.5 µg/µl). After incubation for 10 min at 70^{o}C (GeneAmp PCR System 2400, Perkin Elmer, USA), the following components were added: 2.25 µl Superscript® III reverse transcriptase (400 U), 15 µl 5X SST buffer, 7.5 µl 10X aa-dUTP mix, 7.5 µl 100mM DTT and 1.5 µl Ribolock RNase Inhibitor [Suppliers: Invitrogen, USA; New England BIOLABS Inc. (NEB), USA; Thermo Scientific Inc., USA]. Samples were incubated at 25^{o}C for 5 min, followed by 42^{o}C overnight (19 hours). To each reaction, 15 µl 0.5 M EDTA and 15 µl 1 N NaOH was added, and incubated for 10 min at 65^{o}C. Reactions were cleaned with the Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel, Germany) as per manufacturer's instructions. The cDNA samples were eluted with nuclease-free water (Bioline, USA) warmed to 37^{o}C and sample concentrations were determined on a Nanodrop ND-1000.

A direct comparison analysis was employed (stressed treatment vs. non-stressed control) for identifying differentially regulated gene expression during treatments. Each cDNA treatment and its corresponding control sample were dye labelled with either Cy3 or Cy5 reactive dye (Amersham, GE Healthcare, UK). The appropriate dye (2.5 µl) was added to 2.5 µl cDNA sample and 5 µl 0.2 M Na_2CO_3 buffer (pH 9.0), then incubated in the dark for 2 hours. A dye swap of one of the three pooled replicates per treatment was included. The Qiagen PCR Purification kit (Qiagen, Germany) was used to clean reactions and the concentration of dye coupled cDNA was determined on a Nanodrop ND-1000. Equal quantities (up to 100 pmol) of dye coupled cDNA were combined, and to this 11 µl of 10X Gene Expression Blocking Agent (Agilent Technologies Ltd) and 2.2 µl 25X Fragmentation buffer (Agilent Technologies Ltd) were added. Samples were incubated at 60^{o}C for 30 minutes. Agilent 2X Hi-RPM hybridisation buffer (55 µl, Agilent Technologies Ltd) was added to samples before 100 µl of the mix was carefully pipetted onto the centre of a gasket slide, over which the microarray slide was placed, and secured in the chamber with a clamp assembly. Slides were rotated at 10 rpm and maintained at 65^{o}C for 17 hours. Following careful separation from the cover gasket slide, the microarray was washed with Agilent Wash buffers 1 and 2 as per manufacturer's instructions. Microarray slides were spin-dried for 1 min and immediately scanned on an Axon Genepix Pro scanner (Molecular Devices LLC, USA).

2.2.3. Bioinformatic analyses

GenePix Pro 6.1 software (Molecular Devices LLC, USA) was used to align a template to hybridised spots, measure intensities (background and foreground) of Cy3 and Cy5 dyes, and flag unsuitable features (SNR < 3, % Sat > 20%) in order to generate a spreadsheet of spot intensities for further analysis (.gpr files). Data from individual genotypes were log transformed and analysed using the Bioconductor LIMMA (LInear Models for MicroArray data) package in the R statistical environment. Following Robustspline standardisation, Aquantile normalisation was conducted to ensure that the average intensities had the same empirical distribution across arrays. After correction for multiple testing (using FDR) the table with significantly differentially expressed transcripts was created, and fold change (FC) cut-off was set at 1.5.

Significant differentially expressed ($P < 0.05$) transcript sequences (60 bp) were subjected to BLASTN analysis (http://blast.ncbi.nlm.nih.gov/) to find available genomic information specific to sorghum. The BLAST sequence ID was used to ascertain the 'Sb' gene ID. The MOROKOSHI Sorghum transcriptome database was then used to obtain the functional annotation of sorghum gene IDs (http://sorghum.riken.jp/morokoshi/; Makita et al. 2015). Differentially expressed sorghum sequences with a gene ID and functional annotation were compiled into a spreadsheet. The SorghumCyc database (http://pathway.gramene.org/gramene/sorghumcyc.shtml) was used to obtain pathway information.

The enrichment of Gene Ontology (GO) terms for differentially expressed transcripts with FC ≥ 2 was evaluated with Plant GO Slim in terms of 'biological process' category in AgriGO (http://bioinfo.cau.edu.cn/agriGO/) by the Hypergeometric statistical test method ($P < 0.05$) and Hochberg (FDR) multi-test adjustment method.

2.3. Quantitative Real Time PCR

The cDNA samples for quantitative RT-PCR were synthesised from 1 µg of pooled RNA (from three replicates) using the ImProm-II™ Reverse Transcription System as per manufacturer's instructions (Promega Corporation, USA). Quantitative RT-PCR was performed on a QuantStudio™ 12K Flex Real-Time PCR System (Applied Biosystems™, USA) using SYBR chemistry for selected endogenous controls and genes of interest identified from the microarray analysis. Primers with similar GC and Tm characteristics were designed to selected genes using Primer 3 on Vector NTi® Express Software (Thermo Fisher Scientific Inc, USA) to allow amplification of ~100-bp products. Primers were synthesised by Integrated DNA Technologies (IDT, USA). Thermal cycling conditions were 2 min at 50°C, 10 min at 95°C, followed by 45 cycles of: 95°C for 15s and 57°C for 60s. Samples were quantified in triplicate by the standard curve method. Gene expression quantities were extrapolated from individual standard curves using the Thermo Fisher Cloud Software (Thermo Fisher Scientific, USA). Normalised gene expression data were analysed using STATISTICA version 8.0 (Statsoft® Inc., USA).

The oligo sequences (sequence direction 5′ to 3′) for Sb01g005340.1 Fwd-CATTACTCCAGGGACGCAGT, Rev- CTTGACAGACTGCCATCCAA and Sb10g022840.1 Fwd-AGCGAGGTTCCTGTTTGAGA, Rev- ATGTATCTCATCCGCCGAGT were selected for qRT-PCR. The microarray expression of Sb01g005340.1 was found to respond to progressive water stress and re-watering, and Sb10g022840.1 was significantly expressed in the microarray data. Sb01g001500.1 Fwd-TTCCTCGCTTGGTAAGGATG, Rev-CCTGATGCACAATCAACACC and Sb07g021090.1 Fwd- CGTGGACAGGAACTGTGCT, Rev- GACATCGTTGCCACTTTCAC were selected as endogenous reference genes from the microarray data as their expression were found to be consistent between treatments. In addition, their stability was assessed using qbase^{+} software (https://www.biogazelle.com/qbaseplus) and these genes were found to be suitable as reference genes. Further, PCR products were sequenced in the forward and reverse directions to confirm targets [ABI 3500XL Genetic Analyzer, POP7™ (ThermoScientific) - Inqaba Biotec^{TM}, RSA].

2.4. Nuclear magnetic resonance (NMR) spectroscopy

The protocol for extraction of metabolites from leaf tissue was adapted from Sun et al. (2016). Frozen ground leaf tissue (200 mg) from each plant was transferred to a 2 ml micro-centrifuge tube (QSP, USA) and 1 ml of pre-cooled water:methanol (1:1) mixture [Merck Millipore, Merck (Pty) Ltd., RSA; Romil Ltd., UK] was added, vortexed for 30 sec and sonicated in an ice bath for 1 min. Samples were centrifuged at 11 500 *g* for 20 min at 4^{o}C (Biofuge 13, Heraeus Instruments, Germany). For each sample, this procedure was performed three times and the aqueous fractions were combined. Thereafter, combined aqueous fractions from three biological replicates were pooled (as per the microarray analysis). Thus, there were three pooled biological replicates for each stress treatment and well-watered control.

The methanol was removed under vacuum to reduce percentage methanol to approximately 5%. The supernatants were frozen at -80^{o}C and lyophilised in a FreeZone Triad Freeze Dry System (Labconco®, USA) for 24 hours. Thereafter, 800 µl 100% deuterium oxide (D_2O; Sigma-Aldrich® Co. LLC, USA) and 160 µl sodium phosphate buffer (pH 7) were added to dried aqueous fractions and

gently mixed. The sodium phosphate buffer comprised of 10% D_2O and 0.02 mM 3-trimethylsilyl [2,2,3,3-d_4] propionate [TSP; Merck (Pty) Ltd., RSA] (Sun et al. 2016). The mixture was left undisturbed for 30 min to allow particulates to settle at the bottom of tubes. Thereafter, 0.65 ml of the supernatant was transferred to Norell® Standard Series™ 5 mm NMR tubes (rated at 600 MHz; Norell® Inc., USA). TSP was used as an internal standard.

All spectra were run on a Varian VNMRS PremiumShield system (Varian Inc., USA) operating at a ^1H frequency of 599.99 MHz. The system was equipped with a room temperature triple resonance (HCN) probe regulated at 30°C. All proton spectra were acquired and processed using VNMRJ 4.2A software over a spectral width of 14 ppm, using a 25 s relaxation delay, 256 scans, 90° pulse angle and 32 000 complex points under standard quantitative NMR conditions. No decoupling of ^{13}C was used. The pure compound of β-alanine betaine was synthesised as described by Mohamed Ahmed et al. (2014) and a sample was prepared for NMR analysis using the same buffered solvent mixture used for the sorghum samples. Free induction decays (FIDs) were manually processed and directly compared to the standard's spectrum for component identification, thereafter representative signals were manually integrated against TSP.

2.4.1. Data analysis

Data were analysed using GraphPad Prism® 5 (GraphPad Software Inc., USA) and were initially tested for normality using the D'Agostino & Pearson omnibus normality test, before a one-way ANOVA analysis was conducted. Microsoft© (2010) Excel was used to generate the graph, calculate means and Standard Error (SE) values.

2.5. Supporting data

The data supporting the results of this article are available in the GEO repository, (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE92487).

3. Results

3.1. Gene expression responses to two water stress treatments and re-watering

Transcriptomic analysis was conducted in order to elucidate the full gene expression profile in response to progressive water stress and re-watering in sorghum landrace LR6 using cDNA microarrays containing 35 899 transcript probes. Mild and severe water stress treatments, as well as a re-watered treatment were imposed on sorghum plants during the reproductive growth stage of development (refer to Devnarain et al. 2016 for further detail). A re-watered treatment was included following progressive water stress to identify the prompt transcriptome changes that may occur as the plant responded to the re-introduction of water.

Each array represented gene expression from a direct comparison of treatment vs. control. During mild stress, 165 transcripts were significantly up-regulated and 203 transcripts down-regulated. In response to severe stress, 322 transcripts were over-expressed and 92 transcripts down-regulated. Following the re-watered treatment, 95 and 45 transcripts were up-regulated and down-regulated, respectively.

Highly differentially expressed transcripts covered a wide spectrum of biological functions and the Top 50 significantly expressed transcripts based on absolute fold change during mild stress, severe stress and re-watering are presented in Tables 1-3.

3.1.1. Gene expression during mild water deficit

During mild stress, the greatest over-expression (FC = 34.59) was recorded for Sb07g019050.1, while Sb08g017260.1 recorded the greatest down-regulation (FC = -23.90). Within the Top 50 differentially expressed transcripts during mild stress, 16 transcripts were significantly over-expressed and 34 were significantly down-regulated (Table 1). Among the most highly elevated transcripts during mild stress are those encoding mitochondrial Transcription tERmination Factor (mTERF) family (FC = 34.59) and anion-transporting ATPase family (FC = 27.74) proteins, as well as a putative Late Embryogenic Abundant (LEA) hydroxyproline-rich glycoprotein family (FC = 17.75) protein. Of the top 50

Table 1: Top 50 differentially regulated genes in response to mild stress compared with controls in LR6. Expression values represent absolute fold change (FC) with P < 0.05, FDR = 5% and n = 3.

Gene ID	FC	Reg	Functional annotation
Sb07g019050.1	34.59	Up	mitochondrial transcription termination factor family protein/mTERF family protein
Sb06g018950.1	27.74	Up	anion-transporting ATPase family protein
Sb08g017260.1	23.90	Down	ATAPRL5 (APR-like 5)
Sb04g009840.1	17.75	Up	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family*
Sb01g036790.1	15.03	Down	ATECP63 (EMBRYONIC CELL PROTEIN 63)
Sb01g019970.1	14.44	Down	unknown protein
Sb02g039160.1	14.00	Up	inter-alpha-trypsin inhibitor heavy chain-related
Sb04g032690.1	13.95	Down	vesicle-associated membrane family protein/VAMP family protein
Sb04g029670.1	13.75	Down	Eukaryotic aspartyl protease family protein*
Sb05g019850.1	13.48	Up	TT4 (TRANSPARENT TESTA 4); naringenin-chalcone synthase
Sb09g002350.1	12.78	Down	oxidoreductase, 2OG-Fe(II) oxygenase family protein
Sb09g000210.1	12.52	Down	haloacid dehalogenase-like hydrolase family protein
Sb03g041640.1	12.05	Down	SET domain-containing protein
Sb06g015170.1	11.94	Down	unknown protein
Sb10g023665.1	11.72	Down	MYB120 (MYB DOMAIN PROTEIN 120); DNA binding/transcription factor
Sb04g031240.1	11.70	Down	zinc finger (C3HC4-type RING finger) family protein
Sb09g029900.1	11.53	Up	AAA-type ATPase family protein
Sb06g002300.1	11.10	Down	ATPase, coupled to transmembrane movement of substances
Sb09g000630.1	10.49	Down	receptor-like protein kinase, putative
Sb08g019162.1	10.28	Down	unknown protein
Sb04g002080.1	10.27	Down	LPA2 (low psii accumulation2)
Sb02g038080.1	9.87	Down	pentatricopeptide (PPR) repeat-containing protein
Sb01g005960.1	9.63	Down	(S)-2-hydroxy-acid oxidase, peroxisomal, putative/glycolate oxidase, putative/short chain alpha-hydroxy acid oxidase, putative
Sb01g002020.1	9.47	Down	NAPRT1 (NICOTINATE PHOSPHORIBOSYLTRANSFERASE 1); nicotinate phosphoribosyltransferase
Sb07g024330.1	9.34	Down	CYP71 (CYCLOPHILIN71); chromatin binding/histone binding/peptidyl-prolyl cis-trans isomerase
Sb04g036940.1	9.26	Up	catalytic/iron ion binding/oxidoreductase
Sb01g048160.1	9.17	Down	BSK1 (BR-SIGNALING KINASE 1); ATP binding/binding/kinase/protein kinase/protein tyrosine kinase
Sb04g023340.1	8.84	Down	transcription factor
Sb03g045990.1	8.63	Up	GLP5 (GERMIN-LIKE PROTEIN 5); manganese ion binding/nutrient reservoir
Sb04g029580.1	8.59	Up	SEC14 cytosolic factor, putative/phosphoglyceride transfer protein, putative
Sb01g038485.1	8.51	Down	PHD finger transcription factor, putative*
Sb03g009930.1	8.45	Down	zinc finger (CCCH-type) family protein
Sb08g018340.1	8.41	Down	metallopeptidase M24 family protein
Sb06g015150.1	8.38	Down	Protein phosphatase 2C family protein*
Sb03g011420.1	8.33	Up	F-box family protein
Sb06g003020.1	8.30	Down	ALE1 (ABNORMAL LEAF-SHAPE 1); serine-type endopeptidase
Sb10g026810.1	8.18	Down	MEE44 (maternal effect embryo arrest 44); nucleotidyltransferase
Sb05g026370.1	8.10	Up	Tetratricopeptide repeat (TPR)-like superfamily protein*
Sb10g024860.1	7.95	Down	MPL1 (MYZUS PERSICAE-INDUCED LIPASE 1); catalytic
Sb02g010830.1	7.83	Down	PIP2B (PLASMA MEMBRANE INTRINSIC PROTEIN 2); water channel
Sb05g002540.1	7.61	Up	IMS1 (2-ISOPROPYLMALATE SYNTHASE 1); 2-isopropylmalate synthase
Sb10g024320.1	7.57	Up	unknown protein
Sb02g033540.1	7.46	Up	small G protein family protein/RhoGAP family protein
Sb10g025970.1	7.25	Up	RDR6 (RNA-DEPENDENT RNA POLYMERASE 6); RNA-directed RNA polymerase/nucleic acid binding
Sb09g003860.1	7.19	Down	DRB2 (DSRNA-BINDING PROTEIN 2); double-stranded RNA binding
Sb01g043090.1	7.13	Down	SNAP33 (SOLUBLE N-ETHYLMALEIMIDE-SENSITIVE FACTOR ADAPTOR PROTEIN 33); SNAP receptor/protein binding
Sb01g007830.1	7.13	Up	ATKCO1; calcium-activated potassium channel/ion channel/outward rectifier potassium channel
Sb02g018570.1	7.10	Down	emsy N terminus domain-containing protein/ENT domain-containing protein
Sb06g027490.1	7.02	Down	calcium-binding EF hand family protein
Sb03g013687.1	6.91	Down	IMPA-1 (IMPORTIN ALPHA ISOFORM 1); binding/protein transporter

*Description from MOROKOSHI Sorghum transcriptome database – Functional annotation not available

differentially expressed transcripts during mild stress, four encode proteins of unknown function (Table 1).

3.1.2. Gene expression in response to severe water deficit

During severe stress, the highest over-expression (FC = 25.38) was Sb03g026070.1 and the greatest suppression (FC = -30.48) was recorded for Sb02g029700.1. Within the Top 50 differentially expressed transcripts during severe stress, 41 were significantly over-expressed and nine were significantly down-regulated (Table 2). Among the most highly elevated transcripts during severe stress are those that encode proteins with putative homology to ABscisic acid-Insensitive 2 (*ABI2*; FC = 25.38) and mannosyltransferase (FC = 20.60). Of the top 50 differentially expressed transcripts during severe stress, 12 encode proteins of unknown function (Table 2).

3.1.3. Gene expression during the re-watered treatment

Following re-watering, the highest over-expression (FC = 15.86) was found with Sb04g000660.1 and the greatest down-regulation (FC = -23.79) was recorded for Sb08g017260.1. Within the Top 50 differentially expressed transcripts during re-watering, 36 genes were significantly over-expressed and 14 genes were significantly down-regulated (Table 3). Highly suppressed transcripts in response to re-watering encode a putative Ran BP2/NZF zinc finger-like superfamily protein (FC = -23.79) and low-molecular-weight cysteine-rich 73 (LCR73, FC = -13.26). Among the most highly elevated transcripts during this treatment are those which encode a protein kinase family protein (FC = 15.86) and zinc ion binding protein (FC = 14.74). Of the top 50 differentially expressed transcripts following re-watering, 10 encode proteins of unknown function (Table 3).

Table 2: Top 50 differentially expressed genes in response to severe stress compared with controls in LR6. Expression values represent absolute fold change (FC) with $P < 0.05$, FDR = 5% and n = 3.

Gene ID	FC	Reg	Functional annotation
Sb02g029700.1	30.48	Down	unknown protein
Sb03g026070.1	25.38	Up	homology to ABI2*
Sb09g018420.1	20.71	Down	RAB18 (RESPONSIVE TO ABA 18)
Sb09g025250.1	20.60	Up	mannosyltransferase
Sb01g011620.1	18.97	Up	acid phosphatase/oxidoreductase/transition metal ion binding
Sb01g016420.1	18.38	Up	4-coumarate--CoA ligase, putative/4-coumaroyl-CoA synthase, putative
Sb02g003280.1	17.67	Down	unknown protein
Sb02g035610.1	16.49	Down	oxygen-evolving enhancer protein 3, chloroplast, putative (PSBQ1) (PSBQ)
Sb04g025650.1	15.43	Up	AtLa1 (Arabidopsis thaliana La protein 1); RNA binding
Sb02g031620.1	14.01	Up	Protein of unknown function (DUF3550/UPF0682)*
Sb07g021990.1	13.92	Up	unknown protein
Sb09g001780.1	13.63	Up	disease resistance protein (CC-NBS class), putative
Sb02g000430.1	13.20	Up	ATGLR3.4; intracellular ligand-gated ion channel
Sb01g014930.1	12.82	Up	ARA12; serine-type endopeptidase
Sb09g021870.1	12.67	Up	unknown protein
Sb06g028930.1	12.41	Up	disease resistance protein (NBS-LRR class), putative
Sb01g035810.1	12.00	Up	unknown protein
Sb03g012385.1	11.87	Up	3\'-5\' exonuclease/nucleic acid binding
Sb06g026090.1	11.64	Up	leucine-rich repeat family protein
Sb03g013630.1	11.15	Up	UNE5 (UNFERTILIZED EMBRYO SAC 5); protein disulfide isomerase
Sb03g009050.1	10.52	Up	unknown protein
Sb02g000690.1	10.29	Up	disease resistance-responsive family protein
Sb07g015090.1	9.72	Down	calmodulin-binding protein
Sb10g022680.1	9.16	Up	coclaurine N-methyltransferase, putative
Sb09g029130.1	8.66	Up	CTP synthase, putative/UTP--ammonia ligase, putative
Sb03g008160.1	8.59	Up	PR5K; kinase/transmembrane receptor protein serine/threonine kinase
Sb09g003220.1	8.56	Down	signal recognition particle receptor alpha subunit family protein
Sb01g044190.1	8.53	Up	GCN5-related N-acetyltransferase (GNAT) family protein
Sb02g023660.1	8.33	Up	glycosyl hydrolase family 81 protein
Sb08g002780.1	8.27	Up	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
Sb03g003810.1	8.26	Up	unknown protein
Sb05g022210.1	8.02	Up	glycosyl transferase family 29 protein/sialyltransferase family protein
Sb03g028800.1	8.02	Up	protein kinase family protein
Sb01g017260.1	7.90	Up	RNA binding
Sb04g007980.1	7.55	Up	unknown protein
Sb07g005450.1	7.49	Down	pentatricopeptide (PPR) repeat-containing protein
Sb01g033530.1	7.48	Up	basic helix-loop-helix (bHLH) family protein
Sb02g034730.1	7.37	Up	unknown protein
Sb01g042990.1	7.33	Up	unknown protein
Sb08g003250.1	7.26	Up	ACT domain-containing protein
Sb01g019940.1	7.01	Up	unknown protein
Sb08g022613.1	6.74	Up	Eukaryotic aspartyl protease family protein*
Sb03g042170.1	6.70	Up	peptide chain release factor, putative
Sb05g007070.1	6.56	Down	LHCB5 (LIGHT HARVESTING COMPLEX OF PHOTOSYSTEM II 5); chlorophyll binding
Sb10g002450.1	6.31	Up	cornichon family protein
Sb06g003180.1	6.23	Up	pathogenesis-related protein, putative
Sb04g036080.1	6.10	Up	unknown protein
Sb02g002810.1	6.03	Up	F-box family protein
Sb06g016720.1	5.92	Up	SDG26 (SET DOMAIN GROUP 26); histone-lysine N-methyltransferase
Sb05g020910.1	5.71	Down	galactosyl transferase GMA12/MNN10 family protein

*Description from MOROKOSHI Sorghum transcriptome database – Functional annotation not available

Table 3: Top 50 significantly regulated genes in response to re-watering compared with controls in LR6. Values represent absolute fold change (FC) with $P < 0.05$, FDR = 5% and n = 3.

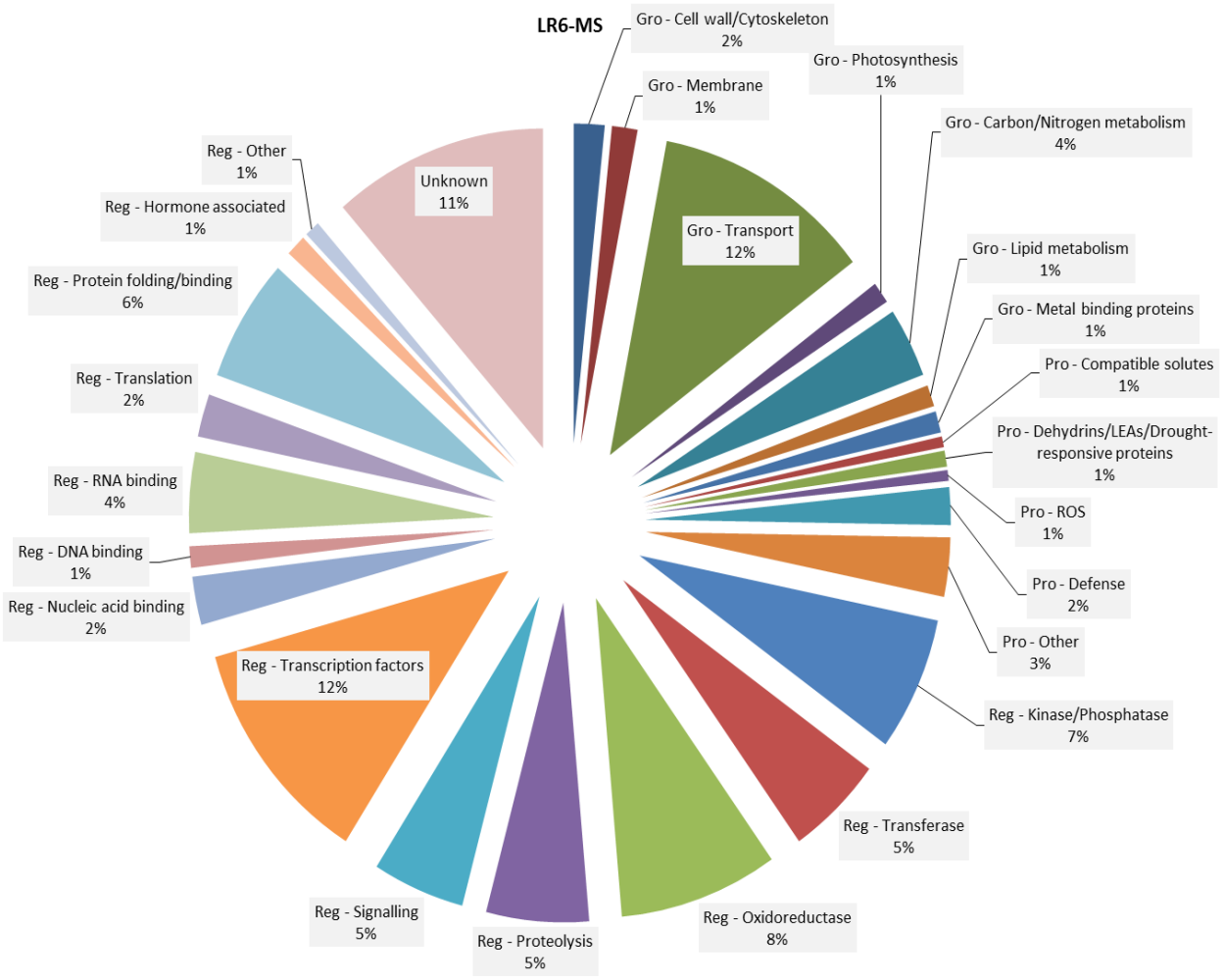
Gene ID	FC	Reg	Functional annotation
Sb01g010560.1	23.79	Down	Ran BP2/NZF zinc finger-like superfamily protein*
Sb04g000660.1	15.86	Up	protein kinase family protein
Sb01g012830.1	14.74	Up	zinc ion binding
Sb02g036850.1	13.26	Down	LCR73 (Low-molecular-weight cysteine-rich 73)
Sb01g021280.1	10.97	Down	emsy N terminus domain-containing protein/ENT domain-containing protein
Sb01g028650.1	10.61	Up	chloroplast chaperonin 10, putative
Sb04g015850.1	10.15	Up	FLA3 (FASCICLIN-LIKE ARABINOGALACTAN PROTEIN 3 PRECURSOR)
Sb03g002590.1	10.05	Up	glycosyltransferase family 14 protein/core-2/I-branching enzyme family protein
Sb01g002440.1	9.93	Up	DNA-binding enhancer protein-related
Sb02g040470.1	9.87	Up	pectinacetylesterase, putative
Sb04g019560.1	9.17	Up	SEC (secret agent); transferase, transferring glycosyl groups
Sb04g006710.1	9.07	Down	unknown protein
Sb03g038450.1	8.22	Down	VHS domain-containing protein/GAT domain-containing protein
Sb04g029160.1	7.66	Up	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein*
Sb05g021020.1	7.48	Down	GRF7 (GENERAL REGULATORY FACTOR 7); protein binding/protein phosphorylated amino acid binding
Sb01g041110.1	7.47	Down	unknown protein
Sb10g029430.1	7.45	Up	unknown protein
Sb01g006280.1	7.35	Down	GTP-binding protein-related*
Sb01g010910.1	7.24	Down	protein-protein interaction regulator family protein*
Sb08g021110.1	7.19	Up	GC-rich sequence DNA-binding factor-like protein with Tuftelin interacting domain*
Sb06g029600.1	7.04	Up	BAN (BANYULS); oxidoreductase
Sb03g009100.1	7.04	Up	unknown protein
Sb01g018000.1	6.94	Down	formin homology 2 domain-containing protein/FH2 domain-containing protein
Sb06g016470.1	6.93	Up	unknown protein
Sb04g029970.1	6.92	Up	ribose-phosphate pyrophosphokinase 4/phosphoribosyl diphosphate synthetase 4
Sb04g001190.1	6.53	Down	FEY (FOREVER YOUNG); binding/catalytic/oxidoreductase
Sb03g036970.1	6.52	Up	unknown protein
Sb04g024370.1	6.46	Up	FLU (FLUORESCENT IN BLUE LIGHT); binding
Sb03g037040.1	6.40	Up	3-hydroxybutyryl-CoA dehydrogenase, putative
Sb10g026850.1	6.37	Up	EYE (EMBRYO YELLOW)
Sb01g039960.1	6.33	Up	unknown protein
Sb07g022650.1	6.31	Down	CYP76C2; electron carrier/heme binding /iron ion binding/monooxygenase/oxygen binding
Sb02g039530.1	5.98	Up	clathrin adaptor complexes medium subunit family protein
Sb01g010250.1	5.94	Up	OMR1 (L-O-METHYLTHREONINE RESISTANT 1); L-threonine ammonia-lyase
Sb09g021810.1	5.83	Up	ATFD3 (ferredoxin 3); 2 iron, 2 sulfur cluster binding / electron carrier/ iron-sulfur cluster binding
Sb09g020580.1	5.71	Up	pentatricopeptide (PPR) repeat-containing protein
Sb07g026600.1	5.48	Up	ARA4; GTP binding / GTPase/ protein binding
Sb02g028968.1	5.38	Up	DNA binding / protein dimerization
Sb06g000730.1	5.38	Up	unknown protein
Sb06g029810.1	5.36	Up	S-locus protein kinase, putative
Sb03g046140.1	5.23	Up	NADPH:quinone oxidoreductase*
Sb01g045320.1	5.15	Up	GP ALPHA 1 (G PROTEIN ALPHA SUBUNIT 1); GTP binding / GTPase/ channel regulator/ signal transducer
Sb05g026050.1	5.13	Up	ARABIDOPSIS TRITHORAX-RELATED PROTEIN 6*
Sb03g028530.1	4.98	Down	WRKY DNA-binding protein 71
Sb10g028730.1	4.93	Up	RPM1 (RESISTANCE TO P. SYRINGAE PV MACULICOLA 1); nucleotide binding / protein binding
Sb02g005030.1	4.93	Up	unknown protein
Sb02g036690.1	4.90	Up	leucine-rich repeat protein kinase, putative
Sb01g041120.1	4.87	Up	Clp amino terminal domain-containing protein
Sb08g019510.1	4.77	Up	unknown protein
Sb05g021742.1	4.71	Down	ATNADP-ME2 (NADP-malic enzyme 2); malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+)/ malic enzyme/ oxidoreductase, acting on NADH or NADPH, NAD or NADP as acceptor

*Description from MOROKOSHI Sorghum transcriptome database – Functional annotation not available

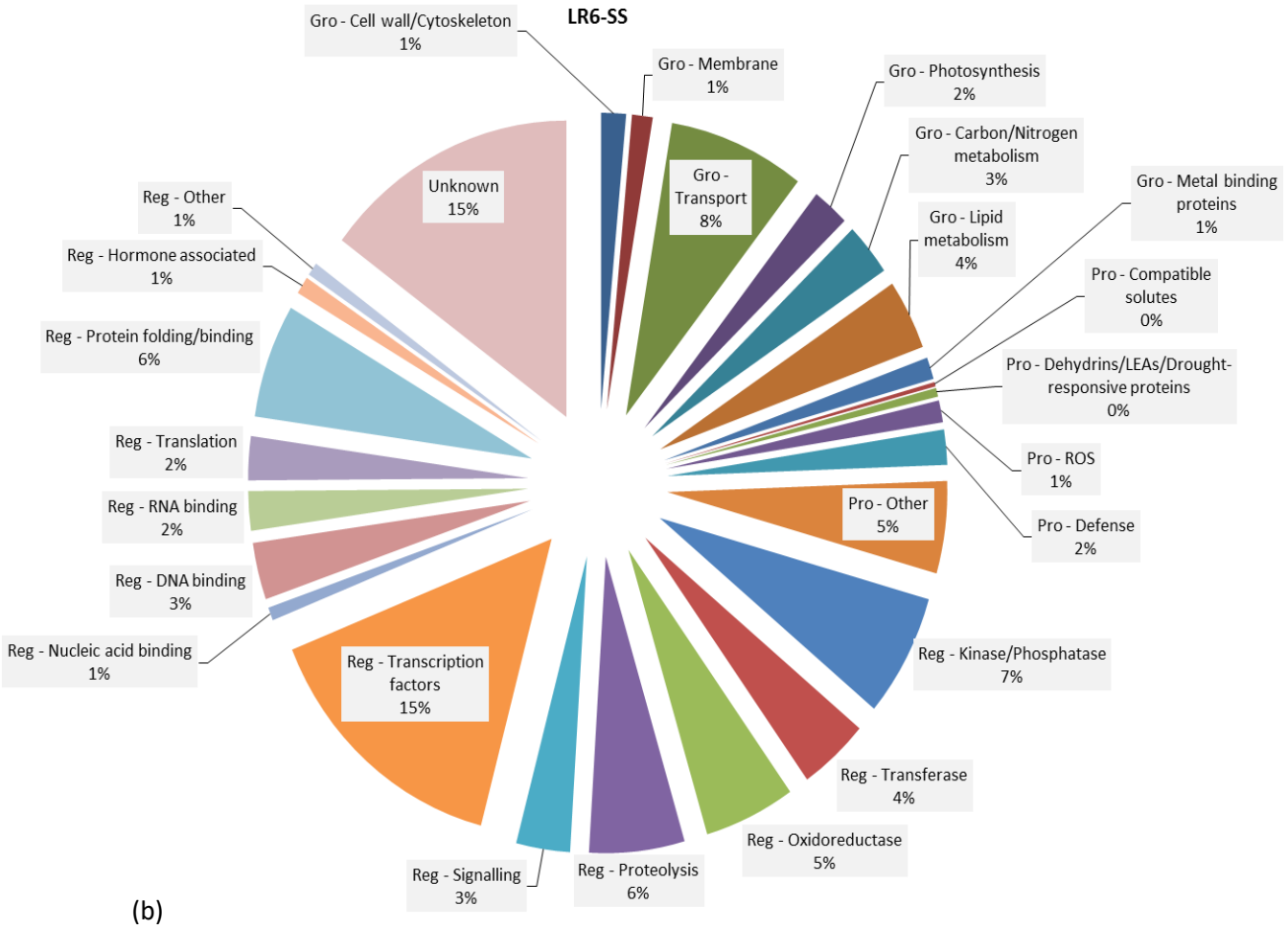
3.1.4. Significant differentially expressed genes categorised by functional annotation

The proportion of significant differentially expressed transcripts ($P < 0.05$, $FC \geq 1.5$) were categorised into functional groups during water stress treatments and re-watering, and visualised with the aid of pie charts. Functional annotations were classified into four categories: Growth (Gro), Protection (Pro), Regulation (Reg) and Unknown (Fig. 1).

The proportion of transcripts representing 'transcription factors' increased from mild stress (12%) to severe stress (15%), then decreased in response to re-watering [6%; Fig. 1 (a-c)]. The proportion of the differentially expressed gene complement associated with 'protein folding/binding' was 6% during mild stress and re-watering, and 4% during severe stress. Transcripts responsible for 'transport' represented 12% of the dataset during mild stress, 8% during severe stress and 4% following re-watering. The proportion of transcripts associated with 'kinase/phosphatase' activity represented 7% of the dataset during mild stress and re-watering, and 6% during severe stress. Transcripts associated with 'oxidoreductase' activity represented 8% of the expressed dataset during mild stress, 5% during severe stress and 9% following re-watering. A large proportion of transcripts that responded to progressive water stress and re-watering in LR6 encode proteins with 'unknown' function (MS = 11%, SS = 15% and RW = 19%).


LR6-MS
Gro - Cell wall/Cytoskeleton 2%
Gro - Membrane 1%
Gro - Photosynthesis 1%
Gro - Carbon/Nitrogen metabolism 4%
Gro - Transport 12%
Gro - Lipid metabolism 1%
Gro - Metal binding proteins 1%
Pro - Compatible solutes 1%
Pro - Dehydrins/LEAs/Drought-responsive proteins 1%
Pro - ROS 1%
Pro - Defense 2%
Pro - Other 3%
Reg - Kinase/Phosphatase 7%
Reg - Transferase 5%
Reg - Oxidoreductase 8%
Reg - Proteolysis 5%
Reg - Signalling 5%
Reg - Transcription factors 12%
Reg - Nucleic acid binding 2%
Reg - DNA binding 1%
Reg - RNA binding 4%
Reg - Translation 2%
Reg - Protein folding/binding 6%
Reg - Hormone associated 1%
Reg - Other 1%
Unknown 11%

(a)


LR6-SS
Gro - Cell wall/Cytoskeleton 1%
Gro - Membrane 1%
Gro - Photosynthesis 2%
Gro - Carbon/Nitrogen metabolism 3%
Gro - Metal binding proteins 1%
Gro - Lipid metabolism 4%
Pro - Compatible solutes 0%
Pro - Dehydrins/LEAs/Drought-responsive proteins 0%
Pro - ROS 1%
Pro - Defense 2%
Pro - Other 5%
Reg - Kinase/Phosphatase 7%
Reg - Transferase 4%
Reg - Oxidoreductase 5%
Reg - Proteolysis 6%
Reg - Signalling 3%
Reg - Transcription factors 15%
Reg - Nucleic acid binding 1%
Reg - DNA binding 3%
Reg - RNA binding 2%
Reg - Translation 2%
Reg - Protein folding/binding 6%
Reg - Hormone associated 1%
Reg - Other 1%
Unknown 15%
Gro - Transport 8%

(b)

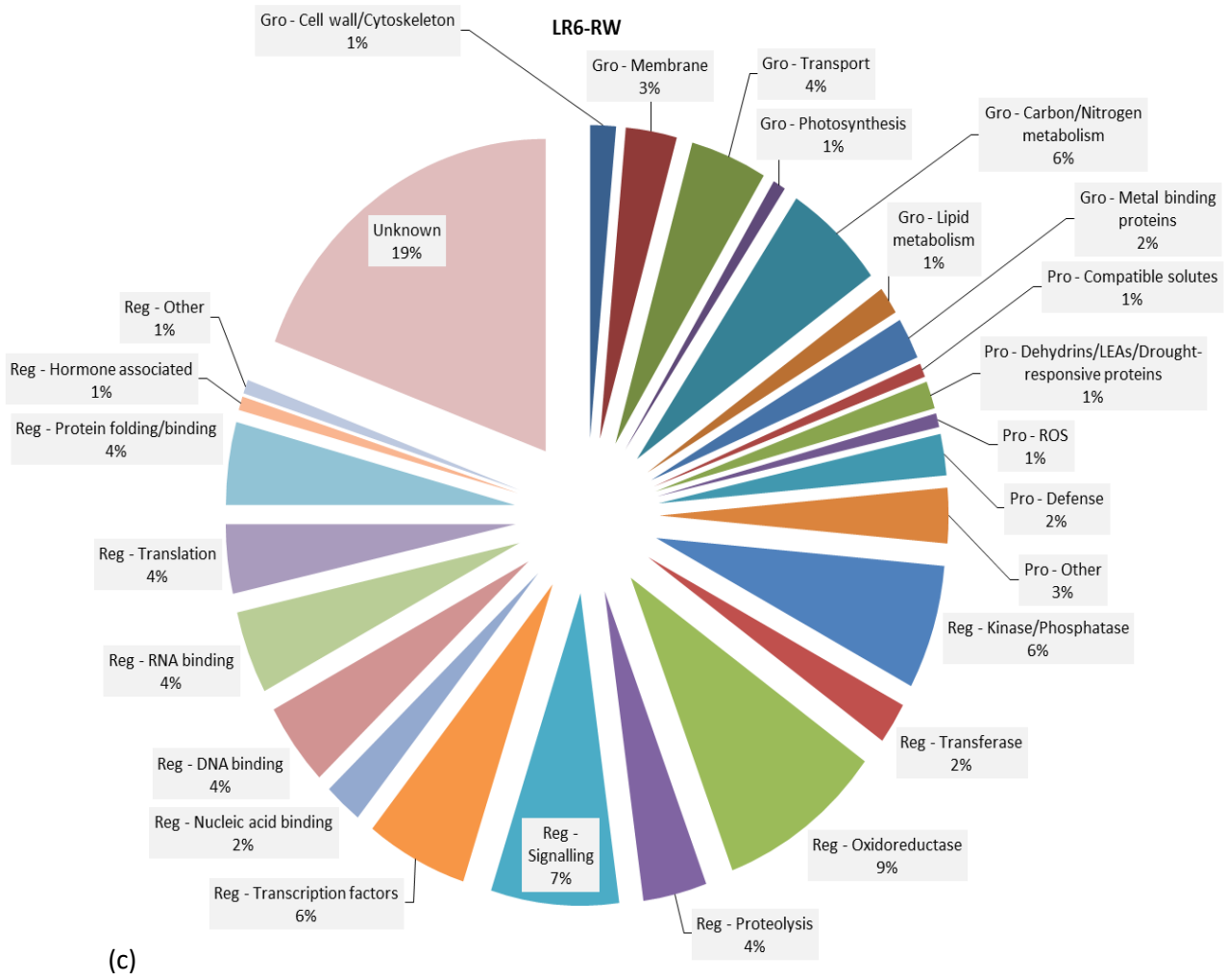


Figure 1(a-c): **Pie charts illustrating the proportions of significantly expressed transcripts in the total gene set of LR6 that responded to progressive water stress and re-watering. (a)** mild stress, **(b)** severe stress and **(c)** re-watered treatment. Transcripts were grouped to represent those with functions in growth (Gro), protection (Pro) and regulation (Reg) during water stress and re-watering.

As seen in Figure 1, transcripts representing 'transcription factors' contributed to a large proportion of the sorghum landrace's response during water deficit and re-watering. There were 112 instances where transcription factor transcripts were differentially expressed in LR6. In response to severe stress, 53 transcripts encoding transcription factor genes were over-expressed, which included zinc finger (11), basic helix-loop-helix (bHLH) family (4), WRKY (3), and myb family (2). During mild stress, the 18 transcription factors that were up-regulated included zinc finger proteins (5) and *ABF3* (*ABSCISIC ACID RESPONSIVE ELEMENTS-BINDING FACTOR 3*) (1).

3.2. Ontological analysis

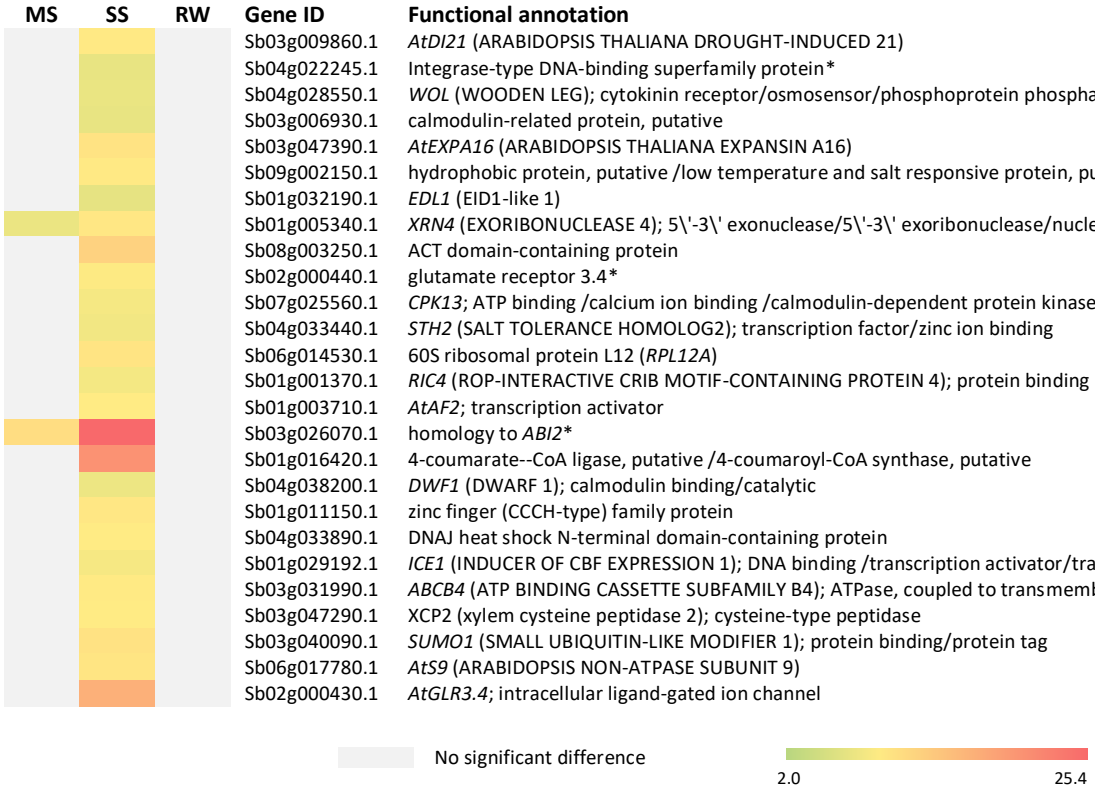
Gene Ontology SEA (singular enrichment analysis) was conducted using AgriGO (http://bioinfo.cau.edu.cn/agriGO/) in order to identify biological processes that were enriched within the differentially expressed gene sets ($P < 0.05$, $FC \geq 2$, FDR = 5%). During mild stress and re-watering, no significant enrichment of genes in comparison to well-watered controls was observed.

Ontological analysis of differentially expressed transcripts identified enriched biological processes in LR6 during severe stress. Three significant nodes were identified with enriched GO categories involved in 'response to stimulus' and 'response to abiotic stimulus' processes in one node. A second node highlighted enriched processes of 'regulation of biological quality', 'homeostatic process' and 'cellular homeostasis'. This node shared 'cellular homeostasis' with a third 'cellular process' node. Significantly enriched genes belonging to 'response to abiotic stimulus' (26), 'homeostatic process' (11), 'regulation of biological quality' (20), 'cellular homeostasis' (9) and 'response to stimulus' (52) were found in LR6 during severe stress (Table 4).

The expression of the 26 enriched genes during 'response to abiotic stimulus' were also considered during mild stress and re-watering in LR6 (Fig. 2). Up-regulated genes in response to severe stress in LR6 encode '*AtEXPA16* (*ARABIDOPSIS THALIANA EXPANSIN A16*)', 'glutamate receptor 3.4' polypeptide associated with transport activity, '*STH2* (*SALT TOLERANCE HOMOLOG2*); transcription factor/zinc ion binding', '*ATAF2*; transcription activator' and '*EDL1* (*EID1-like 1*)'.

Table 4: Gene Ontology (GO) analysis of significant differentially expressed genes in LR6. Data represents enriched biological processes in response to severe stress ($P < 0.05$, $FC \geq 2$, FDR = 5%).

GO term Biological process	Description	No. in selection (out of 153)	No. in genome (out of 26 245)	p-value	FDR
GO:0009628	response to abiotic stimulus	26	2423	0.0018	0.03
GO:0042592	homeostatic process	11	586	0.0008	0.03
GO:0065008	regulation of biological quality	20	1637	0.0015	0.03
GO:0019725	cellular homeostasis	9	433	0.0011	0.03
GO:0050896	response to stimulus	52	6230	0.0028	0.04



*description from MOROKOSHI Sorghum transcriptome database – Functional annotation not available

Figure 2: **Heat map representation of 26 significantly enriched genes in LR6 involved in the 'response to abiotic stimulus' GO category**.

Abbreviations: MS, mild stress; RW, re-watered; SS, severe stress

Two genes (Sb01g005340.1 and Sb03g026070.1) were found to respond to mild and severe stress in LR6. According to SorghumCyc database, Sb01g005340.1 encodes a zinc ion binding enzyme involved with β-alanine betaine biosynthesis. Sb03g026070.1 encodes a polypeptide (protein amino acid dephosphorylation) with no available associated pathway in SorghumCyc.

Other genes from the list of 26 genes up-regulated in LR6 during severe stress encode '*AtDI21* (*ARABIDOPSIS THALIANA DROUGHT-INDUCED 21*)', 'hydrophobic protein, putative/low temperature and salt responsive protein, putative', '*CPK13*; ATP binding/calcium ion binding/calmodulin-dependent protein kinase', 'zinc finger (CCCH-type) family protein', 'DNAJ heat shock N-terminal domain-containing protein', '*ICE1* (*INDUCER OF CBF EXPRESSION 1*); DNA

binding/transcription activator/transcription factor' and '*SUMO1* (*SMALL UBIQUITIN-LIKE MODIFIER 1*)', according to their functional annotations in the MOROKOSHI database. The significant over-expression of these genes associated with the 'response to abiotic stimulus' process likely contributes to tolerance to water deficit in LR6.

3.3. Expression of two β-alanine betaine biosynthesis genes in response to progressive water stress and re-watering

The expression of two genes (Sb01g005340.1 and Sb10g022840.1) associated with the β-alanine betaine biosynthesis pathway, as per SorghumCyc database, was analysed by qRT-PCR [Fig. 3 (a-b)]. The expression of Sb01g005340.1 was significantly elevated during severe stress (1.37 ± 0.07) compared with mild stress (1.02 ± 0.09) and after re-watering (0.93 ± 0.02). Similarly, the expression of Sb10g022840.1 significantly increased from mild stress (0.99 ± 0.03) to severe stress (2.02 ± 0.07) and decreased after re-watering (0.82 ± 0.06).

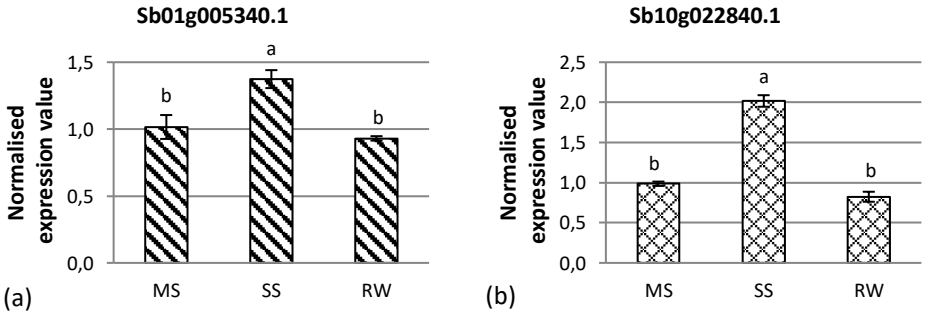


Figure 3 (a-b): Gene expression of two sorghum genes (Sb01g005340.1 and Sb10g022840.1) associated with β-alanine betaine biosynthesis, in response to progressive water stress and re-watering. Expression values were normalised to two endogenous reference genes (Sb01g001500.1 and Sb07g021090.1) and are represented as treatment relative to control with standard error bars. Dissimilar alphabet characters assigned using the Tukey HSD post-hoc test denote a statistical significance (data analysed by an ANOVA analysis, $P < 0.05$).

3.4. Detection of β-alanine betaine in sorghum leaf extracts during progressive water stress and re-watering

Beta(β)-alanine betaine was detected in LR6 leaf extracts using NMR spectroscopy (Fig. 4). The relative abundance of β-alanine betaine was recorded at 13.7 ± 2.13, 20.2 ± 0.82 and 18.3 ± 2.08 during mild stress, severe stress and after re-watering, respectively. The levels of β-alanine betaine in the controls ranged between 11.3 and 14.0. During severe stress, there was a significant difference (ANOVA analysis, $P < 0.05$) in the relative abundance of β-alanine betaine between the treatment (20.2 ± 0.82) and the corresponding control (11.30 ± 1.05).

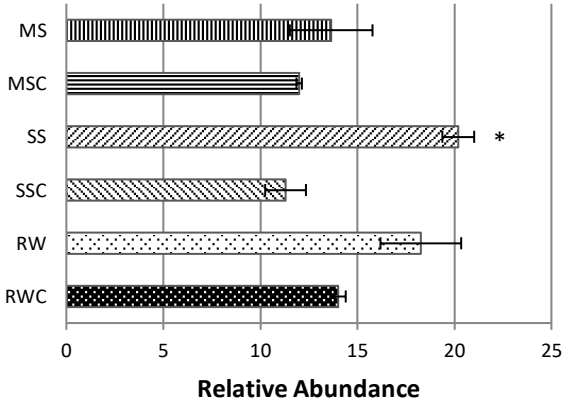


Figure 4: Relative abundance (against TSP standard) of β-alanine betaine in response to progressive water stress (MS and SS) and following re-watering (RW) detected by ^{1}H-NMR spectroscopy in LR6 leaf extracts. Data analysed with an ANOVA analysis and bar charts illustrate mean ± SE, n = 3.
* Significant difference in relative abundance compared with control.
Abbreviations: MS, mild stress; MSC, mild stress control; RW, re-watered; RWC, re-watered control; SS, severe stress; SSC, severe stress control.

4. Discussion

Transcriptome profiling is a useful approach to elucidate the underlying genetic mechanisms supporting abiotic stress tolerance in plants. Here, custom-designed cDNA microarrays comprising 35 899 transcript probes was used to examine the effect of progressive water deficit and re-watering on the transcriptome of a South African sorghum landrace (LR6). The inherent genetic response of sorghum to stress is multi-factorial and involves the interaction of multi-gene and multi-pathway

activation that are necessary to enable adequate tolerance during abiotic stress (Buchanan et al. 2005; Dugas et al. 2011; Johnson et al. 2014).

In response to mild water stress, a significant increase in the expression (FC = 34.59) of a mitochondrial Transcription tERmination Factor (mTERF) was noted in LR6 (Table 1). Plant mTERF genes comprise a large protein family and play a central role in regulating organelle gene expression (Roberti et al. 2009; Kleine 2012; Zhao et al. 2014). Plant mutation studies have found a role for altered mTERF in response to abiotic stresses, such as light sensitivity (Meskauskiene et al. 2009) and heat stress (Kim et al. 2012). Zhao et al. (2014) conducted a genome-wide analysis of mTERF family proteins in maize, revealing 31 putative mTERF genes. These authors found that mTERF genes played various important roles in light acclimation and stress-related responses in maize. According to a review by Quesada (2016), only eight plant mTERFs which encode chloroplast or mitochondrial proteins have been characterised. Metabolic processes within mitochondria and chloroplasts lead to excessive production of reactive oxygen species which require stringent control to avoid cellular damage (Atkin and Macherel 2009; Gill and Tuteja 2010). Over-expression of an mTERF during mild water stress and re-watering may be a potential protective mechanism to reduce the generation of oxidative products. While current bioinformatics and expression research reveal putative roles of mTERF during abiotic stress, further functional studies are necessary to reveal the mechanisms of action of this particular mTERF during abiotic stress.

An over-expression (FC = 17.75) of a putative Late Embryogenic Abundant (LEA) hydroxyproline-rich glycoprotein (HRGP) family protein was found in LR6 during mild water stress (Table 1). The expression of LEA proteins in response to osmotic stress is well documented in literature through functional studies (Babu et al. 2004; Duan and Cai 2012; Amara et al. 2013; Liu et al. 2016). LEA proteins are highly hydrophilic and fundamentally unstructured proteins with high glycine content (Garay-Arroyo et al. 2000). During osmotic stress, these proteins bind to reporter enzymes, thereby maintaining their structural conformation and preventing loss of activity (Olvera-Carrillo et al. 2011). Battaglia et al. (2008) provided a reference platform for hydrophilins (Groups 1-7) in response to

water deficit by reviewing their structural and functional characteristics. Plant HRGPs in particular, are ubiquitous and belong to a superfamily of cell wall proteins (Showalter 1993; Sommer-Knudsen et al. 1998). Within this superfamily, proteins may be categorised as hyperglycosylated arabinogalactan proteins, moderately glycosylated extensins or lightly glycosylated proline-rich proteins (Showalter et al. 2010). Over-expression of a putative LEA-HRGP family protein may have to some degree, contributed to protection against osmotic stress in LR6 during mild water deficit.

Severe water deficit resulted in an over-expression (FC = 25.38) of a protein with putative homology to *ABI2* in LR6 (Table 2). *ABI1* and *ABI2* genes encode homologous type-2C protein phosphatases (PP2C) that have overlapping functions in abscisic acid (ABA) signalling (Leung et al. 1997; Sheen 1998; Merlot et al. 2001). ABA modulates various plant developmental processes and stress responses, and ABA signalling involves both positive and negative regulating kinases and phosphatases (Bonetta and McCourt 1998). Merlot et al. (2001) demonstrated that ABI1 and ABI2 PP2Cs act in a negative feedback regulatory loop of the abscisic acid signalling pathway. A functional study of this particular *ABI2* gene is necessary to determine its precise contribution to ABA signalling and drought tolerance in LR6.

Further, GO analysis revealed 26 genes involved in 'response to abiotic stimulus' significantly enriched in LR6 during severe stress (Fig. 2). In addition to *ABI2*, highly elevated transcripts from this category (FC: 18.38 and 13.20, respectively) included a putative 4-coumarate--CoA ligase and *AtGLR3.4*; intracellular ligand-gated ion channel. 4-Coumarate--CoA ligase catalyses a step in the phenylpropanoid pathway leading to different pathways of secondary metabolism, resulting in the synthesis of various phenolic compounds such as flavonoids, including flavanols, anthocyanins and tannins (Douglas 1996; Weisshaar and Jenkins 1998; Li et al. 2015). Nakabayashi et al. (2014) demonstrated that increased accumulation of flavonoids in *Arabidopsis* enhanced tolerance to oxidative and drought stress. *AtGLR3.4* is a glutamate receptor (GLR) gene of *Arabidopsis* and may function as an amino acid sensor (Price et al. 2012). Meyerhoff et al. (2005) found that *AtGLR3.4* expression was stimulated by abiotic stress in a calcium-dependent and ABA-independent manner.

The simultaneous over-expression of these 26 significantly enriched genes in LR6 undoubtedly contributed to its tolerance during severe water deficit.

Re-watering resulted in suppression of a putative Ran BP2/NZF zinc finger-like superfamily protein in LR6 (FC = 23.79, Table 3). This protein family may be involved in ubiquitin interaction and associated with the nuclear transport protein, Ran (Alam et al. 2004; Higa et al. 2007). Functional studies demonstrating the roles of Ran BP2/NZF zinc finger-like superfamily proteins during plant abiotic stress are lacking and needed.

Significantly, a role for β-alanine betaine was identified in LR6 by microarray analysis and validated by qRT-PCR of two genes (Sb01g005340.1 and Sb10g022840.1) associated with the β-alanine betaine biosynthesis pathway [Fig. 3 (a-b)]. As per the SorghumCyc database, Sb01g005340.1 is associated with zinc ion binding during the biosynthesis of β-alanine betaine and Sb10g022840.1 is an isozyme from the pathway associated with nucleic acid binding. Moreover, β-alanine betaine was detected in leaf extracts of LR6 using NMR spectroscopy and a significant elevation in relative abundance compared with the control was found during severe stress (Fig. 4, ANOVA analysis, $P < 0.05$).

Beta(β)-alanine betaine is an osmo-protective (i.e.: compatible solute) compound that was first identified in the highly stress-tolerant Plumbaginaceae plant family (Hanson et al. 1991; Hanson et al. 1994). As a zwitterionic compatible solute, the accumulation of β-alanine betaine during osmotic stress protects against cellular damage by stabilising proteins and membranes (Yancy 1994; Bray et al. 2000). Dugas et al. (2011) examined the transcriptomic response of sorghum (BTx623) to osmotic stress and exogenous ABA, and found significant differentially expressed genes from the β-alanine betaine biosynthesis pathway in all sorghum tissues and samples that were examined. The biosynthesis of β-alanine betaine is independent of oxygen. In comparison to glycine betaine that has been extensively researched as an osmo-protectant for crop improvement, β-alanine betaine may be

more advantageous during hypoxic conditions and salinity stress (Hanson et al. 1991; Hanson et al. 1994; Rathinasabapathi 2000).

It was demonstrated in *Limonium latifolium* L. (family: Plumbaginaceae) that β-alanine betaine is synthesised through *S*-adenosyl-L-methionine dependent *N*-methylation of β-alanine, and the intermediates are *N*-methyl β-alanine and *N,N*-Dimethyl β-alanine (Rathinasabapathi et al. 2000; Rathinasabapathi et al. 2001; Raman and Rathinasabapathi 2003). Therefore, the availability of β-alanine is an essential component for β-alanine betaine synthesis.

The biosynthesis of β-alanine, which is a non-protein amino acid that is a precursor to pantothenic acid, has not been well studied in plants. Duhazé et al. (2003) and Fouad (2004) highlighted three possible routes for β-alanine biosynthesis: uracil degradation, polyamine oxidation and propionate catabolism. Duhazé et al. (2003) demonstrated that the uracil degradative pathway was important for β-alanine synthesis through dihydrouracil and N-carbamyl β-alanine with [5,6-^{3}H]-uracil radiotracer labelling experiments in two *Limonium sp.* β-alanine synthesis through oxidative degradation of the polyamines, spermine and spermidine in maize shoots was demonstrated by Terano and Suzuki (1978). Rathinasabapathi (2002) demonstrated that propionate metabolism lead to β-alanine in *L. latifolium*. In prokaryotes, the bacterial L-aspartate-α-decarboxylase enzyme catalyses the synthesis of β-alanine (Williamson and Brown 1979; Cronan 1980).

In terms of metabolic engineering initiatives for β-alanine betaine biosynthesis, two studies have considered engineering β-alanine accumulation for stress tolerance in tobacco (*Nicotiana tabacum* L.; Fouad and Rathinasabapathi 2006; Fouad and Altpeter 2009). These researchers demonstrated that expression of the *Escherichia coli* L-aspartate-α-decarboxylase (encoded by the *panD* gene) increased levels of β-alanine and pantothenate, was associated with improved thermotolerance and enhanced biomass production following heat stress. Therefore, the possibility of producing agro-economic crop cultivars resilient to abiotic stresses through metabolic engineering of the β-alanine betaine biosynthesis pathway should be further explored.

5. Conclusions

This study revealed an extensive gene expression profile underlying drought tolerance in a South African sorghum landrace (LR6). GO analysis highlighted 26 significantly enriched genes from the 'response to abiotic stimulus' category during severe stress. Importantly, the results from this study have highlighted a role for β-alanine betaine biosynthesis in sorghum drought tolerance. Further investigations of this compound in particular, may provide a basis for enhancing the productivity of sorghum and other closely related crop species during periods of water deficit.

Acknowledgements

The Council for Scientific and Industrial Research (CSIR) - Biosciences, Pretoria and the Gauteng Department of Agriculture and Rural Development (GDARD), South Africa have financially supported this research. Dr RK Chikwamba is acknowledged for her efforts in driving the Sorghum Drought Initiative Programme within the Plant Biotechnology Research Group (2009-2011) at CSIR Biosciences. Sorghum seeds for this study were obtained from the National Plant Genetic Resources Centre of the Department of Agriculture, Forestry and Fisheries (DAFF), South Africa. The National Research Foundation (NRF) is graciously acknowledged for the Professional Development Programme (PDP) PhD Scholarship received by the PhD candidate.

References

Agarwal, P.K., Agarwal. P., Reddy, M.K., Sopory, S.K., 2006. Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant Cell Report 25, 1263-1274.

Alam, S.L., Sun, J., Payne, M., Welch, B.D., Blake, B.K., Davis, D.R., Meyer, H.H., Emr, S.D., Sundquist, W.I., 2004. Ubiquitin interactions of NZF zinc fingers. EMBO Journal 23, 1411-1421.

Amara, I., Capellades, M., Dolors Ludevid, M., Pagès, M., Goday, A., 2013. Enhanced water stress tolerance of transgenic maize plants over-expressing LEA *Rab28* gene. Journal of Plant Physiology 170, 864-873.

Atkin, O.K., Macherel, D., 2009. The crucial role of plant mitochondria in orchestrating drought tolerance. Annals of Botany 103, 581-597.

Babu, R.C., Zhang, J., Blum, A., David Ho, T-H., Wu, R., Nguyen, H.T., 2004. *HVA1*, a LEA gene from barley confers dehydration tolerance in transgenic rice (*Oryza sativa* L.) via cell membrane protection. Plant Science 166, 855-862.

Battaglia, M., Olvera-Carrillo, Y., Garciarrubio, A., Campos, F., Covarrubias, A.A., 2008. The enigmatic LEA proteins and other hydrophilins. Plant Physiology 148, 6-24.

Bohnert, H.J., Gong, Q., Li, P., Ma, S., 2006. Unraveling abiotic stress tolerance mechanisms – getting genomics going. Current Opinions in Plant Biology 9, 180-188.

Bonetta, D., McCourt, P., 1998. Genetic analysis of ABA signal transduction pathways. Trends in Plant Science 3, 231-235.

Boyer, J.S., 1982. Plant productivity and environment. Science 218, 443-448.

Bray, E.A., Bailey-Serres, J., Weretilnyk, E., 2000. Responses to abiotic stresses, in: Buchanan, B.B., Gruissem, W., Jones, R. (Eds.), Biochemistry and Molecular Biology of Plants. American Society of Plant Physiologists, Rockville, MD, USA, pp. 1158-1203.

Buchanan, C.D., Lim, S., Salzman, R.A., Kagiampakis, I., Morishige, D.T., Weers, B.D., Klein, R.R., Pratt, L.H., Cordonnier-Pratt, M-M., Klein, P.E., Mullet, J.E., 2005. *Sorghum bicolor's* transcriptome response to dehydration, high salinity and ABA. Plant Molecular Biology 58, 699-720.

Camacho Villa, T.C., Maxted, N., Scholten, M., Ford-Lloyd, B., 2005. Defining and identifying crop landraces. Plant Genetic Resources 3, 373-384.

Cativelli, L., Rizza, F., Badeck, F-W., Mazzucotelli, E., Mastrangelo, A.M., Francia, E., Máre, C., Tondelli, A., Stanca, A.M., 2008. Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. Field Crops Research 105, 1-14.

Cronan, J.E., 1980. β-alanine synthesis in *Escherichia coli*. Journal of Bacteriology 141, 1291-1297.

De Jager, E., 2016. South Africa – Annual Total Rainfall. South African Weather Service, pp. 1-4.

Devnarain, N., Crampton, B.G., Chikwamba, R., Becker, J.V.W., O'Kennedy, M.M., 2016. Physiological responses of selected African sorghum landraces to progressive water stress and re-watering. South African Journal of Botany 103, 61-69.

Doggett, H., 1988. Sorghum. Longman Scientific & Technical, New York, USA, pp. 1-512.

Dolferus, R., Ji, X., Richards, R.A., 2011 Abiotic stress and control of grain number. Plant Science 181, 331-341.

Douglas, C.J., 1996. Phenylpropanoid metabolism and lignin biosynthesis: From weeds to trees. Trends in Plant Science 1, 171-178.

Duan, J., Cai, W., 2012. *OsLEA3-2*, an abiotic stress induced gene of rice plays a key role in salt and drought tolerance. PLoS ONE 7, e45117.

Dugas, D.V., Monaco, M.K., Olson, A., Klein, R.R., Kumar, S., Ware, D., Klein, P.E., 2011. Functional annotation of the transcriptome of *Sorghum bicolor* in response to osmotic stress and abscisic acid. BMC Genomics 12, 514.

Duhazé, C., Gagneul, D., Leport, L., Larher, F.R., Bouchereau, A., 2003. Uracil as one of the multiple sources of β-alanine in *Limonium latifolium*, a halotolerant β-alanine betaine accumulating Plumaginaceae. Plant Physiology and Biochemistry 41, 993-998.

Food and Agricultural Organization of the United Nations (FAO) (2018) FAOSTAT online database. http://faostat.fao.org/. Date accessed 01 June 2018

Fouad, W.M.M., 2004. Metabolic engineering for beta-alanine overproduction and stress tolerance in plants: Expression of *Escherichia coli* L-aspartate-alpha-decarboxylase in transgenic tobacco. PhD Dissertation, University of Florida, pp. 1-106.

Fouad, W.M., Rathinasabapathi, B., 2006. Expression of bacterial L-aspartate-α-decarboxylase in tobacco increases β-alanine and pantothenate levels and improves thermotolerance. Plant Molecular Biology 60, 495-505.

Fouad, W.M., Altpeter, F., 2009. Transplastomic expression of bacterial L-aspartate-α-decarboxylase enhances photosynthesis and biomass production in response to high temperature stress. Transgenic Research 18, 707-718.

Garay-Arroyo, A., Colmenero-Flores, J.M., Garciarrubio, A., Covarrubias, A.A., 2000. Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. Journal of Biological Chemistry 275, 5668-5674.

Gill, S.S., Tuteja, N., 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiology and Biochemistry 48, 909-930.

Hanson, A.D., Rathinasabapathi, B., Chamberlin, B., Gage, D.A., 1991. Comparative physiological evidence that β-alanine betaine and choline-O-sulfate act as compatible osmolytes in halophytic *Limonium* species. Plant Physiology 97, 1199-1205.

Hanson, A.D., Rathinasabapathi, B., Rivoal, J., Burnet, M., Dillon, M.O., Gage, D.A., 1994. Osmoprotective compounds in the Plumbaginaceae: A natural experiment in metabolic engineering of stress tolerance. Proceedings of the National Academy of Sciences (PNAS) 91, 306-310.

Higa, M.M., Alam, S.L., Sundquist, W.I., Ullman, K.S., 2007. Molecular characterization of the Ran-binding zinc finger domain of Nup153. The Journal Biological Chemistry 282, 17090-17100.

Johnson, S.M., Lim, F-L., Finkler, A., Fromm, H., Slabas, A.R., Knight, M.R., 2014. Transcriptomic analysis of *Sorghum bicolor* responding to combined heat and drought stress. BMC Genomics 15, 456.

Johnson, S.M., Cummins, I., Lim, F.L., Slabas, A.R., Knight, M.R., 2015. Transcriptomic analysis comparing stay-green and senescent *Sorghum bicolor* lines identifies a role for proline biosynthesis in the stay-green trait. Journal of Experimental Botany 66, 7061-7073.

Kholova, J., Hash, C.T., Kakkera, A., Kočová, M., Vadez, V., 2009. Constitutive water-conserving mechanisms are correlated with the terminal drought tolerance of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. Journal of Experimental Botany 61, 369-377.

Kim, M., Lee, U., Small, I., Colas Des Francs-Small, C., Vierling, E., 2012. Mutations in an *Arabidopsis* mitochondrial transcription termination factor-related protein enhance thermotolerance in the absence of the major molecular chaperone HSP101. The Plant Cell 24, 3349-3365.

Kleine, T., 2012. *Arabidopsis thaliana* mTERF proteins: Evolution and functional classification. Frontiers in Plant Science 3, 233.

Leung, J., Merlot, S., Giraudat, J., 1997. The Arabidopsis *ABSCISIC ACID-INSENSITIVE2* (*ABI2*) and *ABI1* genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. The Plant Cell 9, 759-771.

Li, Y., Kim, J.I., Pysh, L., Chapple, C., 2015. Four isoforms of *Arabidopsis* 4-coumarate:CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. Plant Physiology 169, 2409-2421.

Liu, Y., Liang, J., Sun, L., Yang, X., Li, D., 2016. Group 3 LEA protein, ZmLEA3, is involved in protection from low temperature stress. Frontiers in Plant Science 7, 1011.

Makita, Y., Shimada, S., Kawashima, M., Kondou-Kuriyama, T., Toyoda, T., Matsui, M., 2015. MOROKOSHI: transcriptome database in *Sorghum bicolor*. Plant and Cell Physiology 56, e6.

Merlot, S., Gosti, F., Guerrier, D., Vavasseur, A., Giraudat, J., 2001. The ABI1 and ABI2 protein phosphatases2C act in a negative feedback regulatory loop of the abscisic acid signalling pathway. The Plant Journal 25, 295-303.

Meskauskiene, R., Würsch, M., Laloi, C., Vidi, P-A., Coll, N.S., Kessler, F., Baruah, A., Kim, C., Apel, K., 2009. A mutation in the *Arabidopsis* mTERF-related plastid protein *SOLDAT10* activates retrograde signalling and suppresses $^1O_2$-induced cell death. The Plant Journal 60, 399-410.

Meyerhoff, O., Müller, K., Roelfsema, M.R.G., Latz, A., Lacombe, B., Hedrich, R., Dietrich, P., Becker, D., 2005. *AtGLR3.4*, a glutamate receptor channel-like gene is sensitive to touch and cold. Planta 222, 418-427.

Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. Journal of Hydrology 391, 202-216.

Mohamed Ahmed, I.A., Maimaiti, A., Mori, N., Yamanaka, N., Taniguchi, T., 2014. Simultaneous determination of β-alanine betaine and trimethylamine in bacterial culture and plant samples by capillary electrophoresis. Journal of Analytical Science and Technology 5, 38.

Nakabayashi, R., Yonekura-Sakakibara, K., Urano, K., Suzuki, M., Yamada, Y., Nishizawa, T., Matsuda, F., Kojima, M., Sakakibara, H., Shinozaki, K., Michael, A.J., Tohge, T., Yamazaki, M., Saito, K., 2014. Enhancement of oxidative and drought tolerance in Arabidopsis by overaccumulation of antioxidant flavonoids. The Plant Journal 77, 367-379.

Newton, A.C., Akar, T., Baresel, J.P., Bebeli, P.J., Bettencourt, E., Bladenopoulos, K.V., Czembor, J.H., Fasoula, D.A., Katsiotis, A., Koutis, K., Koutsika-Sotiriou, M., Kovacs, G., Larsson, H., Pinheiro De Carvalho, M.A.A., Rubiales, D., Russel, J., Dos Santos, T.M.M., Vaz Patto, M.C., 2010. Cereal landraces for sustainable agriculture. A review. Agronomy for Sustainable Development 30, 237-269.

Olvera-Carrillo, Y., Luis Reyes, J., Covarrubias, A.A., 2011. Late embryogenesis abundant proteins. Plant Signaling and Behavior 6, 586-589.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schutz, J., Spannagi, M., Tang, H., Wang, X., Wicker, T., Bharti, A.K., Chapman, J., Feltus, F.A., Gowik, U., Grigoriev, I.V., Lyons, E., Maher, C.A., Martis, M., Narechania, A., Otillar, R.P., Penning, B.W., Salamov, A.A., Wang, Y., Zhang, L., Carpita, N.C., Freeling, M., Gingle, A.R., Hash, C.T., Keller, B., Klein, P., Kresovich, S., McCann, M.C., Ming, R., Peterson, D.G., Rahman, M., Ware, D., Westhoff, P., Mayer, K.F.X., Messing, J., Rokhsar, D.S., 2009. The *Sorghum bicolor* genome and the diversification of grasses. Nature 457, 551-556.

Pinheiro De Carvalho, M.A.A., Bebeli, P.J., Bettencourt, E., Costa, G., Dias, S., Dos Santos, T.M.M., Slaski, J.J., 2013. Cereal landraces genetic resources in worldwide Genebanks. A review. Agronomy for Sustainable Development 33, 177-203.

Price, M.B., Jelesko, J., Okumoto, S., 2012. Glutamate receptor homologs in plants: Functions and evolutionary origins. Frontiers in Plant Science 3, 235.

Quesada, V., 2016. The roles of mitochondrial transcription termination factors (mTERFs) in plants. Physiologia Plantarum 157, 389-399.

Rabbani, M.A., Maruyama, K., Abe, H., Khan, M.A., Katsura, K., Ito, Y., Yoshiwara, K., Seki, M., Shinozaki, K., Yamaguchi-Shinozaki, K., 2003. Monitoring expression profiles in rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. Plant Physiology 133, 1755-1767.

Raman, S.B., Rathinasabapathi, B., 2003. β-alanine *N*-methyltransferase of *Limonium latifolium*. cDNA cloning and functional expression of a novel *N*-methyltransferase implicated in the synthesis of the osmoprotectant β-alanine betaine. Plant Physiology 132, 1642-1651.

Rathinasabapathi, B., 2000. Metabolic engineering for stress tolerance: Installing osmoprotectant synthesis pathways. Annals of Botany 86, 709-716.

Rathinasabapathi, B., 2002. Propionate, a source of β-alanine, is an inhibitor of β-alanine methylation in *Limonium latifolium*, Plumbaginaceae. Journal of Plant Physiology 159, 671-674.

Rathinasabapathi B, Sigua C, Ho J, Gage DA. 2000. Osmoprotectant β-alanine betaine synthesis in the Plumbaginaceae: *S*-adenosyl-L-methionine dependent *N*-methylation of β-alanine to its betaine is via *N*-methyl and *N,N*-dimethyl β-alanines. *Physiologia Plantarum* 109: 225-231.

Rathinasabapathi, B., Fouad, W.M., Sigua, C.A., 2001. β-alanine betaine synthesis in the Plumbaginaceae. Purification and characterization of a trifunctional, *S*-adenosyl-L-methionine-dependent *N*-methyltransferase from *Limonium latifolium* leaves. Plant Physiology 126, 1241-1249.

Roberti, M., Polosa, P.L., Bruni, F., Manzari, C., Deceglie, S., Gadaleta, M.N., Cantatore, P., 2009. The mTERF family proteins: Mitochondrial transcription regulators and beyond. Biochimica et Biophysica Acta 1787, 303-311.

Sheen, J., 1998. Mutational analysis of protein phosphatase 2C involved in abscisic acid signal transduction in higher plants. Proceedings of the National Academy of Science (PNAS) 95, 975-980.

Shinozaki, K., Yamaguchi-Shinozaki, K., 2007. Gene networks involved in drought stress response and tolerance. Journal of Experimental Botany 58, 221-227.

Showalter, A.M., 1993. Structure and function of plant cell wall proteins. The Plant Cell 5, 9-23.

Showalter, A.M., Keppler, B., Lichtenberg, J., Gu, D., Welch, L.R., 2010. A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. Plant Physiology 153, 485-513.

Sommer-Knudsen, J., Bacic, A., Clarke, A.E., 1998. Hydroxyproline-rich plant glycoproteins. Phytochemistry 47, 483-497.

Sreenivasulu, N., Sopory, S.K., Kavi Kishor, P.B., 2007. Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches. Gene 388, 1-13.

Subudhi, P.K., Nguyen, H.T., 2000. New horizons in biotechnology, in: Smith, C.W., Frederiksen, R.A., (Eds.) Sorghum: Origin, History, Technology, and Production. John Wiley & Sons, Inc., USA, pp. 349-398.

Sun, C.X., Li, M.Q., Gao, X.X., Liu, L.N., Wu, X.F., Zhou, J.H., 2016. Metabolic response of maize plants to multi-factorial abiotic stresses. Plant Biology 18, 120-129.

Terano, S., Suzuki, Y., 1978. Formation of β-alanine from spermine and spermidine in maize shoots. Phytochemistry 17, 148-149.

Van Vuuren, L., 2015. Drought Management – Strengthening our knowledge armoury. The Water Wheel November/December 2015, pp. 14-16.

Vincour, B., Altman, A., 2005. Recent advances in engineering plant tolerance to abiotic stress: Achievements and limitations. Current Opinion in Biotechnology 16, 123-132.

Weisshaar, B., Jenkins, G.I., 1998. Phenylpropanoid biosynthesis and its regulation. Current Opinion in Plant Biology 1, 251-257.

Wilhite, D.A., 1993. The Enigma of Drought, in: Wilhite, D.A. (Ed.) Drought Assessment, Management, and Planning: Theory and Case Studies. Kluwer Academic Publishers, USA, pp. 3-15.

Williamson, J.M., Brown, G.M., 1979. Purification and properties of L-aspartate-α-decarboxylase, an enzyme that catalyzes the formation of β-alanine in *Escherichia coli*. The Journal of Biological Chemistry 254, 8074-8082.

Yancy, P.H., 1994. Compatible and counteracting solutes, in: Strange, K. (Ed.) Cellular and Molecular Physiology of Cell Volume Regulation. CRC Press Inc., FL, USA, pp. 81-109.

Zhao, Y., Cai, M., Zhang, X., Li, Y., Zhang, J., Zhao, H., Kong, F., Zheng, Y., Qiu F., 2014. Genome-wide identification, evolution and expression analysis of mTERF gene family in maize. PLoS ONE 9, e94126.