

RESEARCH ARTICLE

Identification of Gene Modules Associated with Low Temperatures Response in Bambara Groundnut by Network-Based Analysis

Venkata Suresh Bonthala^{1,2,3*}, Katie Mayes⁴, Joanna Moreton^{5,6,7}, Martin Blythe⁷, Victoria Wright⁷, Sean Tobias May⁴, Festo Massawe², Sean Mayes^{3,4}, Jamie Twycross¹

1 School of Computer Sciences, Jubilee Campus, University of Nottingham, Nottingham, United Kingdom, **2** School of Biosciences, University of Nottingham Malaysia Campus, Kuala Lumpur, Malaysia, **3** Crops For the Future, Jalan Broga, 43500 Semenyih, Malaysia, **4** Plant and Crop Sciences, Biosciences, Sutton Bonington Campus, Loughborough, Leicestershire, United Kingdom, **5** Advanced Data Analysis Centre, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire, United Kingdom, **6** School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire, United Kingdom, **7** Deep Seq, School of Life Sciences, University of Nottingham, Medical School, Queen's Medical Centre, Nottingham, United Kingdom

* psxvb@nottingham.ac.uk; gandipalem@gmail.com



OPEN ACCESS

Citation: Bonthala VS, Mayes K, Moreton J, Blythe M, Wright V, May ST, et al. (2016) Identification of Gene Modules Associated with Low Temperatures Response in Bambara Groundnut by Network-Based Analysis. PLoS ONE 11(2): e0148771. doi:10.1371/journal.pone.0148771

Editor: Jinfa Zhang, New Mexico State University, UNITED STATES

Received: August 26, 2015

Accepted: January 22, 2016

Published: February 9, 2016

Copyright: © 2016 Bonthala et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data files are available from the NCBI GEO database (accession number(s) GSE72255, GSE75982).

Funding: This work was funded by the Crops for Future (CFF) and University of Nottingham.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Bambara groundnut (*Vigna subterranea* (L.) Verdc.) is an African legume and is a promising underutilized crop with good seed nutritional values. Low temperature stress in a number of African countries at night, such as Botswana, can effect the growth and development of bambara groundnut, leading to losses in potential crop yield. Therefore, in this study we developed a computational pipeline to identify and analyze the genes and gene modules associated with low temperature stress responses in bambara groundnut using the cross-species microarray technique (as bambara groundnut has no microarray chip) coupled with network-based analysis. Analyses of the bambara groundnut transcriptome using cross-species gene expression data resulted in the identification of 375 and 659 differentially expressed genes ($p < 0.01$) under the sub-optimal (23° C) and very sub-optimal (18° C) temperatures, respectively, of which 110 genes are commonly shared between the two stress conditions. The construction of a Highest Reciprocal Rank-based gene co-expression network, followed by its partition using a Heuristic Cluster Chiseling Algorithm resulted in 6 and 7 gene modules in sub-optimal and very sub-optimal temperature stresses being identified, respectively. Modules of sub-optimal temperature stress are principally enriched with carbohydrate and lipid metabolic processes, while most of the modules of very sub-optimal temperature stress are significantly enriched with responses to stimuli and various metabolic processes. Several transcription factors (from MYB, NAC, WRKY, WHIRLY & GATA classes) that may regulate the downstream genes involved in response to stimulus in order for the plant to withstand very sub-optimal temperature stress were highlighted. The identified gene modules could be useful in breeding for low-temperature stress tolerant bambara groundnut varieties.

Introduction

Bambara groundnut (*Vigna subterranea* (L.) Verdc, $2n = 2x = 22$) is a nutritionally rich, under-utilised, indigenous African legume crop and mainly grown for its protein rich seed. This crop continues to be the third most important food legume crop after groundnut and cowpea in semi-arid Africa [1, 2]. Due to its good seed nutritional values, particularly for the protein component in developing country diets, bambara groundnut has been receiving increased interest and could undergo a shift from subsistence crop to a cash crop i.e., bambara groundnut has been canned at a commercial level [3]. The remarkable feature of this crop is that it can produce yield in soils which are too poor for cultivation of other crops like groundnut (*Arachis hypogaea*) or where drought stress is too extreme [4].

Low temperature (LT) can be a major abiotic stress, particularly in non-equatorial regions of Africa and elevated planes. LT stress can limit the growth and development of the crop, leading to the loss of potential yield, with the limited reports for bambara groundnut suggesting base temperatures (below which no growth and development occurs) of 12–13°C. Similar to other crops [5,6], the growth and productivity of bambara groundnut is also effected by LT stress in a number of ways, such as delays in the germination of seeds, a reduction in total dry matter accumulated (TDM), reduced shelling percentage and reduced pod and seed yields. Usually the optimum temperature (T_0) for bambara groundnut ranges from 20–28°C [4, 7]. Thus, the low temperatures restrict the times of the year when the farmers can plant seed, particularly where the temperature at night is below the T_0 . However, while landraces which are grown within such stressed environments have been indirectly selected by growth in the target environment, any cold tolerance present may not be optimised and without an understanding of the trait, it may be difficult to introduce new genetic variation from other environments into the target environment. Despite its importance as a promising food and economically valuable crop, bambara groundnut is mainly cultivated as landraces and there is a need to develop bambara groundnut varieties which are tolerant to low temperature stress, which could further improve the options for growth and productivity of this drought tolerant legume as part of sustainable low input agriculture.

The advancement of molecular technologies and high-throughput “omics” tools such as microarrays and deep-sequencing studies have become a useful strategy for the global analysis of plant gene expression under a particular biotic and/or abiotic stress condition. Using microarrays, the abiotic stress responses in *Arabidopsis thaliana* [8], *Oryza sativa* [9], *Glycine max* [10], *Solanum lycopersicum* [11] and in other plants have been widely analyzed and several stress responsive genes have been identified, although these genes have yet to be evaluated in field crop situations. Unfortunately, commercial microarray chips are not available for under-utilized and understudied minor crops like bambara groundnut and alternatives such as comprehensive RNA sequencing—while becoming more affordable—can still be costly, prohibiting routine large-scale experiments. One possible solution for this problem is a cross-species (Xspecies) microarray approach [12], involving the hybridisation between the cRNA of the species of interest and a closely related species for which a custom microarray chip is available. This approach has been successfully used to study the transcriptomes of different important crops [13–15]. However, these microarray analyses may provide less information on gene-gene functional relationships. To study the gene-gene interrelationships under a particular biological treatment, such as experiencing biotic and/or abiotic stress conditions, network-based analyses, such as gene co-expression network analyses, have been popularly used [16–20].

Gene co-expression networks are constructed from expression data generated by using either microarrays or deep-sequencing (RNA-Seq) methods to provide a global view of gene-gene interrelationships and can help in the identification of candidate genes as well as the better

understanding of how the inter-connected genes interplay to carry out specific biological functions under a specific stress condition. In recent years, co-expression networks have been constructed for model species, *Arabidopsis* [16], and for other important food crops such as *Oryza sativa* [17], *Citrus* [18], *Glycine max* [19], *Vitis vinifera* [20] and have been provided a global view of the transcriptional relationships under specific stress conditions. From these gene co-expression network analyses, a number of candidate genes and gene functional modules have been identified that are associated with a specific biological process in plants. In addition, several web-based databases for plant gene co-expression networks, including LegumeGRN [21], CoP [22], PLANEX [23], PLEXdb [24] and CressExpress [25], have been developed to enable the visualization of gene-gene interrelationships, gene-specific expression profiles across stress conditions and data mining of co-expression networks for plant breeders and biologists.

In the present study we report the use of cRNA from bambara groundnut grown under sub-optimal and very sub-optimal temperature stresses, hybridized with Affymetrix Soybean's GeneChipTM array to study low temperature responsive candidate genes expression in bambara groundnut. To further study gene-gene functional relationships and identify the gene functional modules associated with LT, we have constructed a gene co-expression network, followed by the partitioning of the co-expression networks to identify LT responsive functional modules to provide a better understanding of the underlying molecular mechanisms of the LT response in bambara groundnut.

Materials & Methods

Plant Materials

Three plants of the bambara groundnut genotype 'S19-3' were grown in controlled environment growth rooms at Sutton Bonington Campus, University of Nottingham, under a 12 hour photoperiod and at a constant temperature of 27°C. Plants were grown in soil columns containing a growing medium of 1 part John Innes 2 compost to 1 part sand, and were watered as required to main approx. field capacity. A single, fully-expanded leaflet was sampled from each of the three plants growing at 27°C, snap frozen in liquid nitrogen and stored at -80°C. Plants were given a further 3 days at 27°C to recover from sampling before being moved to a controlled environment room at 23°C. On the fifth day at 23°C a single fully-expanded leaflet was sampled and stored, as described above. After 3 further days plants were moved to 18°C, and then five days later were sampled again.

Genomic DNA and total RNA extraction

Frozen bambara groundnut leaflets were ground to a fine powder, under liquid nitrogen, in a mortar and pestle. Extraction of genomic DNA was carried using 100 mg ground leaf tissue, using a DNeasy Plant Mini Kit (Qiagen), according to manufacturer's directions. The quality and concentration of DNA was assessed by visualization on agarose gel and by spectrophotometry, using a nanodrop. For extraction of total RNA, 100 mg ground leaf tissue was processed, as per manufacturer's instructions, using an RNeasy Plant Mini Kit (Qiagen) and including the additional DNA digestion step. RNA samples were then run on the Agilent Bioanalyzer to determine quality and concentration.

Genomic DNA and cRNA hybridization

Genomic DNA from bambara groundnut leaf tissue was labelled with the Bioprime DNA labeling system kit and then hybridized with Soybean's GeneChipTM (Affymetrix) for 16 hours at 45°C using the standard hybridization protocol (Affymetrix) and analysed by scanning. Total

RNA (5 µg) from bambara groundnut of optimal, sub-optimal and very sub-optimal temperatures was reverse-transcribed to synthesize double stranded cDNA using the standard protocol and then the resulting samples were *in-vitro* transcribed to generate complementary RNAs (cRNAs) incorporating biotinylated nucleotides using T7 DNA polymerase. Purified cRNA (15 µg) were heat-treated and then hybridized to the Soybean GeneChip™ (Affymetrix) for 16 hours at 45°C [12]. A triplicate hybridization (biological, from different plants) was used for each temperature (18°, 23° and 27°C) and all the hybridization data have been submitted to NCBI's GEO database (<https://www.ncbi.nlm.nih.gov/geo>; Acc. No. GSE72255).

Probe selection and Normalization

Probe-pair information from the genomic DNA (gDNA) hybridization file (.cel) were extracted using a.cel file parser available at <https://www.affymetrix.arabidopsis.info/xspecies/>, which generates a customised chip definition file (CDF) compatible with various microarray data analysis programs, including R/BioConductor[12]. The RNA hybridisation files (.cel) were loaded into the R environment applying the customized CDF files generated in the previous step and then were pre-processed using the RMA (Robust Multichip Average) pre-normalization algorithm present in the “affy” package [26].

Identification of differentially expressed genes

Differentially expressed genes were identified based on the baseline data generated under 27°C (control) in both sub-optimal (23°C) and very sub-optimal temperature (18°C) stresses using a t-test and corrected by False Discovery Rate (FDR) approach for multiple hypothesis testing [27]. Average signal intensities of all biological replicates for each sample were used to calculate the fold-change of gene expression.

RNA-seq data based expression profiling of DE genes

Unpublished SOLiD4 reads were generated from the same RNA samples as the microarray data and have been submitted to NCBI's GEO database (<https://www.ncbi.nlm.nih.gov/geo>; Acc. No. GSE75982). Reads which were at least 50bp in length, were aligned to an unpublished transcriptome assembly of Bambara groundnut using LifeScope version 2.5.1 [28] and read counts were extracted from the alignment file. DESeq package version 1.22.0 [29] were used to identify the differentially expressed genes. Further, the expression patterns of DE genes identified from RNA-seq data were matched with the DE genes identified from microarray data via BLASTN [30] searching of bambara groundnut's unpublished transcriptome sequences against the Soybean GeneChip design sequences with default parameters. The matched differentially expressed probes were visualised, based on variance stabilization transformed data, using heatmaps by heatmap.2 function of R's gplots package version 2.1 (<http://cran.r-project.org/package=gplots>).

Construction of the co-expression network

Highest-reciprocal rank (HRR) based co-expression network methodology was employed to further investigate the interactions between the identified DE genes by constructing co-expression networks for each of the stress conditions. The HRR-based co-expression network were created using the HRRNetworkCreator tool [31] and then the co-expression networks were visualized using Cytoscape [32].

Module detection and Functional enrichment

HCCA [31] was used to detect the modules in the constructed co-expression networks using default parameters. The GO annotation of the *Glycine max* genome was downloaded from Soy-Base (<http://www.soybase.org>). Significantly enriched Gene Ontology terms for each of the detected module was carried out by AgriGO Tool [33].

Results

Both soybean and bambara groundnut are phylogenetically related legume crops (around 20 million years divergence since a common ancestor; [34]). Hence in this study we have used Soybean's GeneChip™ array (Affymetrix) to study the transcriptome of Bambara groundnut under low temperatures. Further, we have developed a computational pipeline (Fig 1) to identify the low temperature stress associated genes/gene modules using cross-species based expression datasets coupled with network-based analysis. The pipeline consists of the following steps: (1) probe selection using a gDNA hybridization approach, (2) pre-processing of raw expression data, (3) identification of DE genes, (4) construction of an HRR-based co-expression network, (5) detection of modules, and (6) detection of biologically relevant modules by GO analysis.

Probe Selection and identification of differentially expressed genes

After hybridizing the genomic DNA of bambara groundnut with the Soybean GeneChip™ array (Affymetrix), a probe-pair was retained if its perfect-match hybridization intensity value was greater than the user defined threshold values (ranging from 0 to 500) using a.cel file parser, which also generated a CDF file for each of the threshold values. These CDF files only contain the probe-pairs in which the perfect-match (PM) oligonucleotides has a gDNA hybridization intensity greater than the user defined threshold [12]. Table 1 shows the number of retained probe pairs and probe-sets for each threshold value for further analysis.

The raw expression intensity values (RNA.cel files) were filtered for low- or non-specific hybridization using the custom generated CDF files from the previous step and then normalized using the RMA algorithm present in “affy” package of BioConductor in an R environment [26]. To further reduce the noise in the normalized data, only the probe-sets present (P) in all the array slides (total 9 slides) were selected. Differentially expressed genes (DEGs) were calculated for each threshold value (0–500) using a t-test and then corrected by using False Discovery Rates (FDR) for multiple hypothesis testing [27]. To consider a probe-set as differentially expressed, it should contain <0.01 corrected p-value with a >2 fold-change (either up- or down-regulated). Based on this criteria, the CDF file of 100 threshold values was considered as a best threshold value (Table 1) as it returned highest number of DEGs (375 in optimal vs. sub-optimal temperature and 659 in optimal vs. very sub-optimal temperature). Out of 375 DEGs, 204 genes are up-regulated while 171 genes are down-regulated under the sub-optimal temperature (S1 Table). On the other hand, out of 659 DEGs, 403 genes are up-regulated while 257 genes are down-regulated in the very sub-optimal temperature (S2 Table). There were 110 DEGs which are common between sub-optimal and very sub-optimal temperatures suggesting that a similar genetic response is likely to underlie the response of bambara groundnut to these two different temperatures. Out of these 110 common DEGs, 76 genes are up-regulated while 34 genes are down-regulated (S3 Table).

RNA-seq data based expression profiling of DE genes

We have chosen top hundred DE genes (up- and down-regulated) to validate their expression patterns using unpublished RNA-seq data in three steps: (1) Aligned RNA-seq reads of three

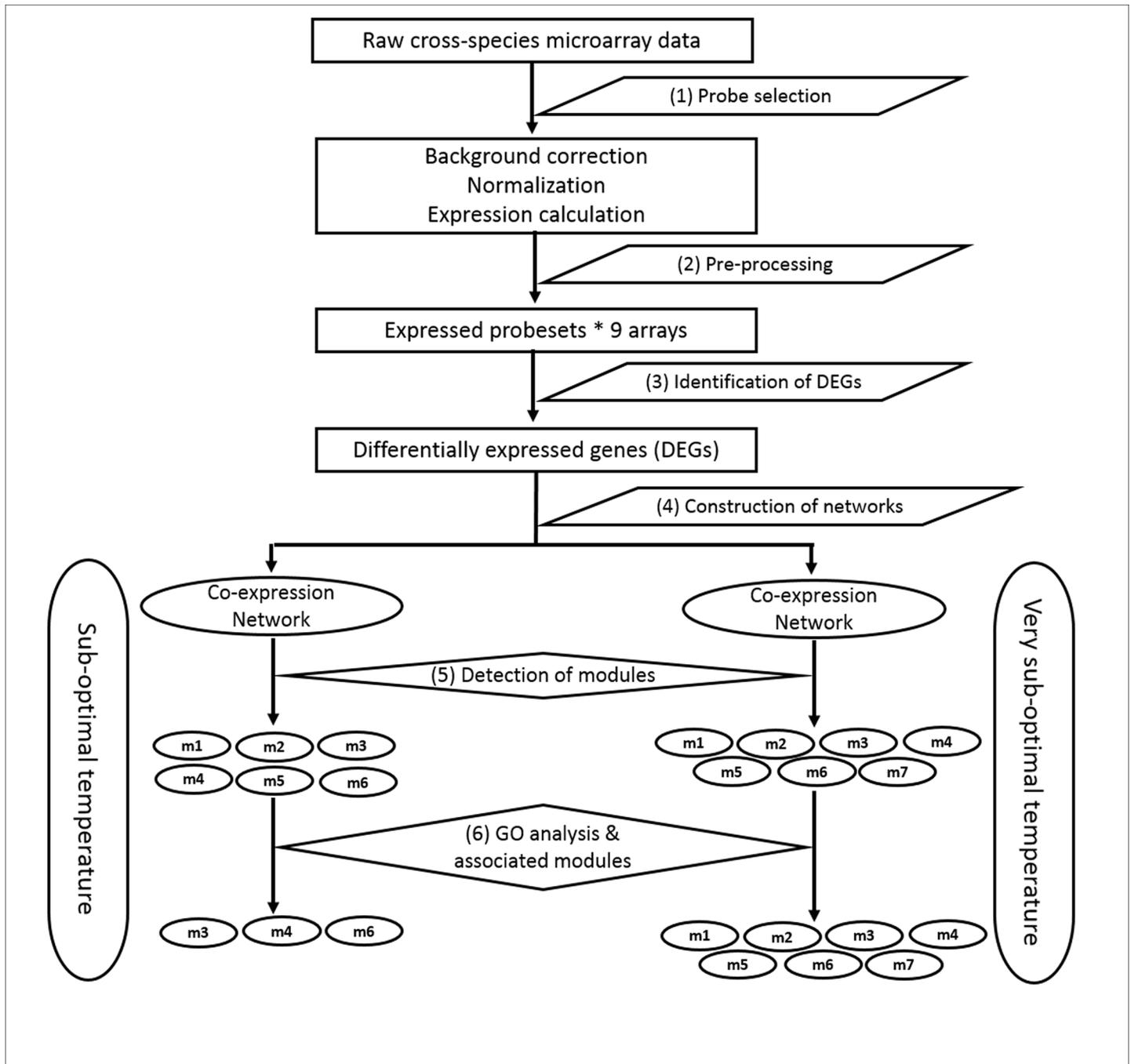


Fig 1. Flowchart of the pipeline for identification of gene modules associated with low temperature response in bambara groundnut.

doi:10.1371/journal.pone.0148771.g001

stress conditions individually to the unpublished transcriptome assembly of bambara groundnut using LifeScope version 2.5.1 [28] and read count for each of the isotig were extracted; (2) Then the count data were load into DESeq [29] and identified DE genes with adjusted p-value < 0.1; and (3) The top hundred DE genes (up- and down-regulated) of microarray data were matched against the DE genes obtained from RNA-seq data via BLASTN [30] searching of isotigs of bambara groundnut’s transcriptome against the Soybean GeneChip design

Table 1. Statistics of retained probes and probe-sets on the Affymetrix Soybean GeneChip array before further analysis. This table contains information on the number of probe-pairs and probe-sets retained for further analysis corresponding to each threshold value. The last two columns indicate the number of identified DEGs under sub-optimal and very sub-optimal temperature stresses compared to the optimal temperature, corresponding to each threshold value, based on masked probe-sets in the custom.cdf file.

Threshold values	Number of probes (complete Soybean chip)	Number of probesets (complete Soybean chip)	Number of DEGs (Sub-optimal temperature)	Number of DEGs (Very sub-optimal temperature)
0	670388	61072	256	513
50	602176	61058	342	612
100	326207	59978	375	659
200	130572	50822	335	592
300	65127	36716	254	372
400	36933	25017	190	238
500	23184	17163	127	192

doi:10.1371/journal.pone.0148771.t001

sequences with default parameters which resulted in 52 common DE genes (up- and down-regulated) between microarray-based top hundred DE genes and RNA-seq data based DE genes (adjusted p-value <0.1). Finally, the 52 common DE genes were used to generate the heatmap (S1 Fig and S4 Table) using heatmap.2 function. The heatmap indicates that the 52 DE genes identified from microarray data were consistent when using either cross-species microarray approaches or RNA-seq based analysis.

Construction of co-expression networks and detection of functional modules

We have employed the Highest Reciprocal Rank (HRR) based method to construct the co-expression networks. In this approach, the calculated Pearson Correlation Coefficient (r) values were normalized using the highest reciprocal rank i.e., the mutual co-expression relationship between two genes of interest. This method overcomes the problem of calling biological relationships between expressed genes at different cut-off r-values [31]. In general, the statistical significance of the co-expression relationships between two genes of interest may not reflect the biological relevance [35]. Therefore, we determined HRR value that optimized biological relevance and found that the $10 \leq \text{HRR} \leq 20$ produced biologically relevant networks in both the stress conditions i.e., HRR = 10 produced 28% and 34% of the nodes disconnected from networks of sub-optimal and very sub-optimal temperature stress conditions, respectively, while HRR = 20 decreased this proportion to 0.8% and 0.3% of nodes, respectively. Thus, combining both biological relevance and selecting for the maximum number of nodes connected in both the networks, we found that HRR = 20 resulted in biologically relevant co-expression networks with 4218 edges between 375 nodes (S5 Table) and 7178 edges between 659 nodes (S6 Table) in sub-optimal and very sub-optimal temperatures, respectively.

Responses to any stress are usually organized as relatively separable functional modules of highly interconnected genes in the co-expression network at a given cut-off. Genes present in the same functional module are co-expressed across diverse conditions, and thus, functional consistency among the genes present in the same modules is expected [17, 36, 37]. In this study we used the Heuristic Cluster Chiseling Algorithm (HCCA) [31], with the number of steps away from the seed node (n) = 3 and the average cluster size ranging from 40–200 nodes, to detect biologically related modules and this resulted in 6 (S7 Table) and 7 (S8 Table) modules in sub-optimal and very sub-optimal temperatures, respectively, ranging in size from 40–200 genes per cluster. Tables 2 & 3 shows the statistics of the detected modules in sub-optimal and very sub-optimal temperature co-expression networks, respectively.

Table 2. Summary of the detected modules in the sub-optimal temperature co-expression network. Module column indicates the module number in the network; The Total Genes column indicates the number of genes present in the module, with direction of gene expression change given in the next two columns.

Module	Total Genes	Up-regulated	Down-regulated
1	58	28	30
2	44	21	23
3	58	29	29
4	72	41	31
5	90	54	36
6	53	30	23

doi:10.1371/journal.pone.0148771.t002

Functional enrichment of detected modules

Gene Ontology (GO) annotation is one important steps in network analysis to understand the biological functions and then the identification of over-represented biological processes can reveal the functional features of each detected module. To GO annotate, followed by identifying the over-represented GO terms in the detected modules, we carried out a GO enrichment analysis using the AgriGO tool with the Soybean genome locus annotation of Phytozome as the reference set [33]. Hypergeometric distribution adjusted by Bonferroni correction for the testing of multiple hypotheses with an adjusted threshold of $p < 0.01$ and a minimum number of mapping entries of 10 were used to evaluate the statistical significance of the functional enrichment within the detected gene modules.

Modules associated with sub-optimal temperature. Functional enrichment analysis resulted in the significant enrichment of three modules (3rd, 4th & 5th) with GO terms with the genes present in these modules principally involved in the biological processes of carbohydrate and lipid metabolic processes (S9 and S10 Tables). Fig 2 displays mapping of significantly enriched GO terms on the sub-optimal temperature stress responsive gene co-expression network.

Module 3 consists of 58 gene nodes and is principally enriched with carbohydrate metabolic processes (GO:0005975, FDR = 0.0023). There are a total of 12 genes whose functions are associated with carbohydrate metabolic processes, which represent nearly 28% of the annotated genes in the module. Interestingly, 50% of genes in this module are up-regulated by sub-optimal temperature stress, while another 50% of genes are down-regulated. Module 4 consists of 72 gene nodes and is also principally enriched with carbohydrate metabolic processes (GO:0005975, FDR = 0.00015). There are a total of 15 genes whose functions are associated

Table 3. Summary of the detected modules in the very sub-optimal temperature co-expression network. Module column indicates the module number in the network; The Total Genes column indicates the number of genes present in the module, with direction of gene expression change given in the next two columns.

Module	Total Genes	Up-regulated	Down-regulated
1	155	93	62
2	185	103	82
3	87	53	34
4	59	37	22
5	86	56	30
6	42	30	12
7	42	27	15

doi:10.1371/journal.pone.0148771.t003

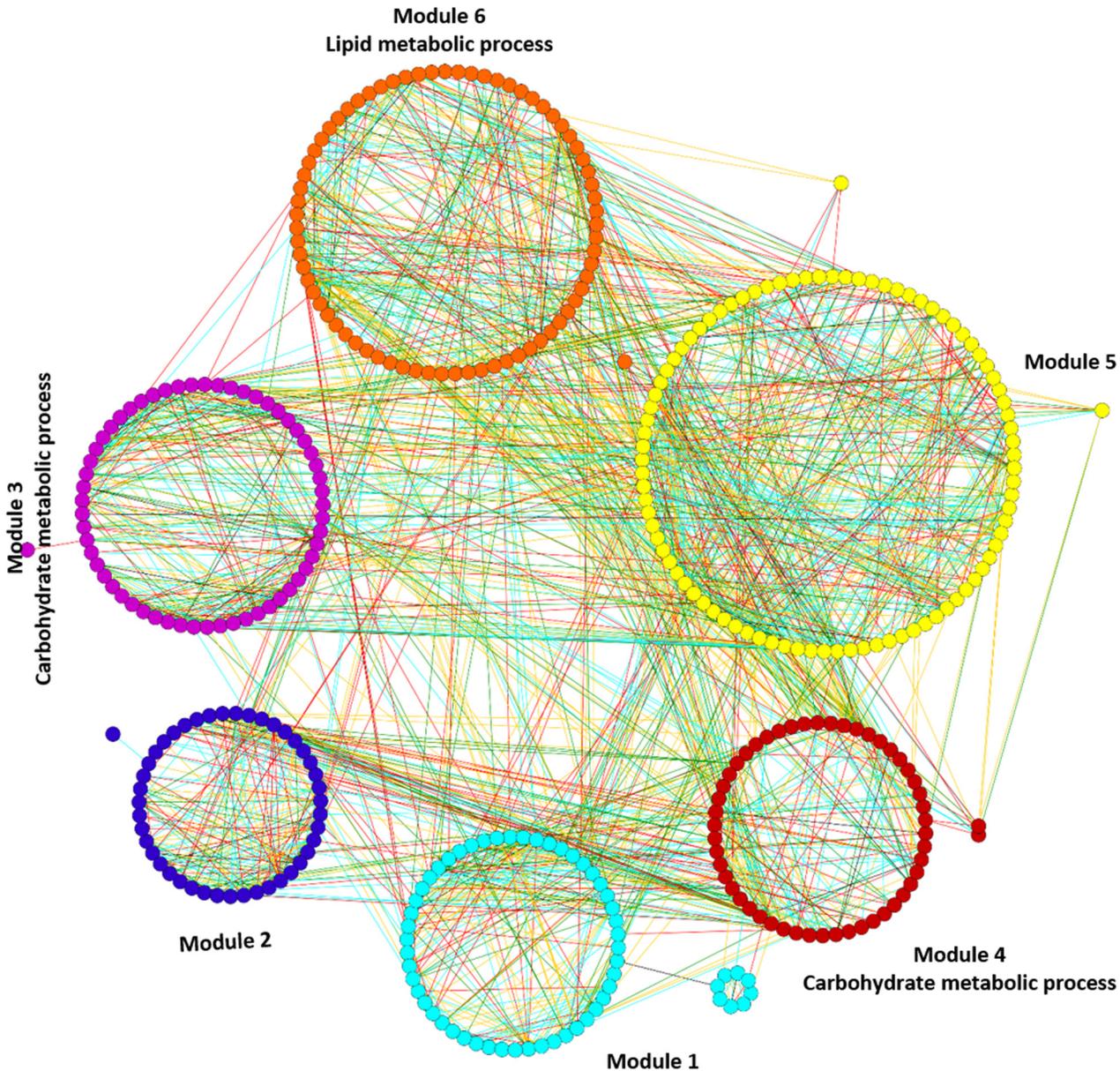


Fig 2. Mapping of enriched GO terms on the sub-optimal temperature stress responsive gene co-expression network. Modules 3, 4 and 6 were significantly enriched with GO terms. The different edge colors indicates different range of HRR values and as follows: Cyan: HRR1-HRR5; Red: HRR6-HRR10; Green: HRR11-HRR15 and Orange yellow: HRR16-HRR20.

doi:10.1371/journal.pone.0148771.g002

with carbohydrate metabolic processes, which represents nearly 24% of the annotated genes in the module. Fifty-seven percent of genes are up-regulated in this module and the rest of the genes are down-regulated by sub-optimal temperature stress. In contrast, module 6 consists of 53 gene nodes and is principally enriched with lipid metabolic process genes (GO:0006629, FDR = 0.004). There are a total of 10 genes whose functions are associated with lipid metabolic processes, which represents nearly 24% of the annotated genes in the module. Fifty-seven percent of genes are up-regulated and the rest of the genes are down-regulated by sub-optimal temperature stress. Overall, 50–57% of genes in all the enriched modules were up-regulated, while 43–50% of genes are down-regulated by sub-optimal temperature stress, which indicates

that the sub-optimal temperature stress has little effect on alteration in expression of genes associated with various metabolic enzymes of carbohydrate and lipid metabolic processes only.

Modules associated with very sub-optimal temperature. Functional enrichment analysis revealed that the very sub-optimal temperature stress responsive module genes were mainly involved in biological processes of various kinds of stress response, photosynthesis and cell and metabolic processes, including carbohydrate metabolism. In the very sub-optimal temperature all the detected modules were significantly enriched for GO terms (S11 and S12 Tables). Fig 3 displays mapping of enriched GO terms in the very sub-optimal temperature stress responsive gene co-expression network. Unlike the sub-optimal temperature stress, the very sub-optimal temperature stress regulated different kinds of stimulus responses along with cell and metabolic processes, which indicates that the plant responds to the decrease in temperature by altering more than one biological process in order to survive under unfavourable temperatures. All the detected modules in very sub-optimal temperature stress are positive regulators for low temperature stress in bambara groundnut.

Module 1 consists of 155 gene nodes and is enriched with stimulus responses (GO:0050896, FDR = 0.00039). There are a total of 67 genes whose functions are associated with responses to stimuli, which represent 54% of the annotated genes in the module. 60% of genes in the module are up-regulated by the very sub-optimal temperature stress, while rest of the genes are down-regulated. The presence of the up-regulated heat-shock protein (HSP) gene (Gma.10282.2.S1_at, Glyma18g01416.1) is indicative among the 67 genes enriched which show upregulation with the “stimulus response” in this module. In general, HSP genes are induced and secrete HSPs when there is sudden change in the genotypic expression of various genes triggered by any kind of stress. HSPs play an essential role as chaperones by assisting the correct folding of stress-accumulated misfolded proteins enhancing cell and plant survival during low temperature stress. It has also been reported that these HSP genes are induced in *Arabidopsis* and other plant species by low temperature stress [38] and thus this module may be preferentially associated with very sub-optimal temperature stress response in bambara groundnut.

Module 3 consists of 87 gene nodes and is enriched with abiotic stimulus response genes (GO:0009628, FDR = 0.0042). There are a total of 22 genes whose functions are associated with responses to abiotic stimuli, which represents 20% of the annotated genes in the module. Nearly 61% of genes in the module are up-regulated by the very sub-optimal temperature, while the rest of the genes are down-regulated. The up-regulated chaperone DnaJ gene is important among these 22 genes. Recent studies indicates that this gene contributes to maintenance of photosystem-II under chilling stress in tomatoes [39] and therefore, this gene may be participating in a similar response in bambara groundnut under very sub-optimal temperature.

Module 4 consists of 59 gene nodes and is enriched with responses to stress (GO:0006950, FDR = 0.0032). There are a total of 22 genes whose functions are associated with stress responses, which represents 47% of the annotated genes in the module. 63% of genes in the module are up-regulated, while the rest of the genes are down-regulated by the very sub-optimal temperature stress. This module contains an important up-regulated gene called WD40 repeat gene (Gma.13316.2.S1_a_at, Glyma18g52040.1) which plays crucial role in diverse protein-protein interactions by acting as a scaffolding molecules and thus assisting the proper molecular activity of proteins. In recent studies, these genes show higher expression during a long duration of cold stress in *Setaria italica*, indicating the up-regulation of these genes in bambara groundnut may play a role in correct scaffolding of low temperature stress-accumulated misfolded proteins [40].

Module 5 consists of 86 gene nodes and is also enriched with responses to stress (GO:0006950, FDR = 0.0019). There are a total of 29 genes whose functions are associated with stress responses, which represents 46% of the annotated genes in the module. 65% of genes in

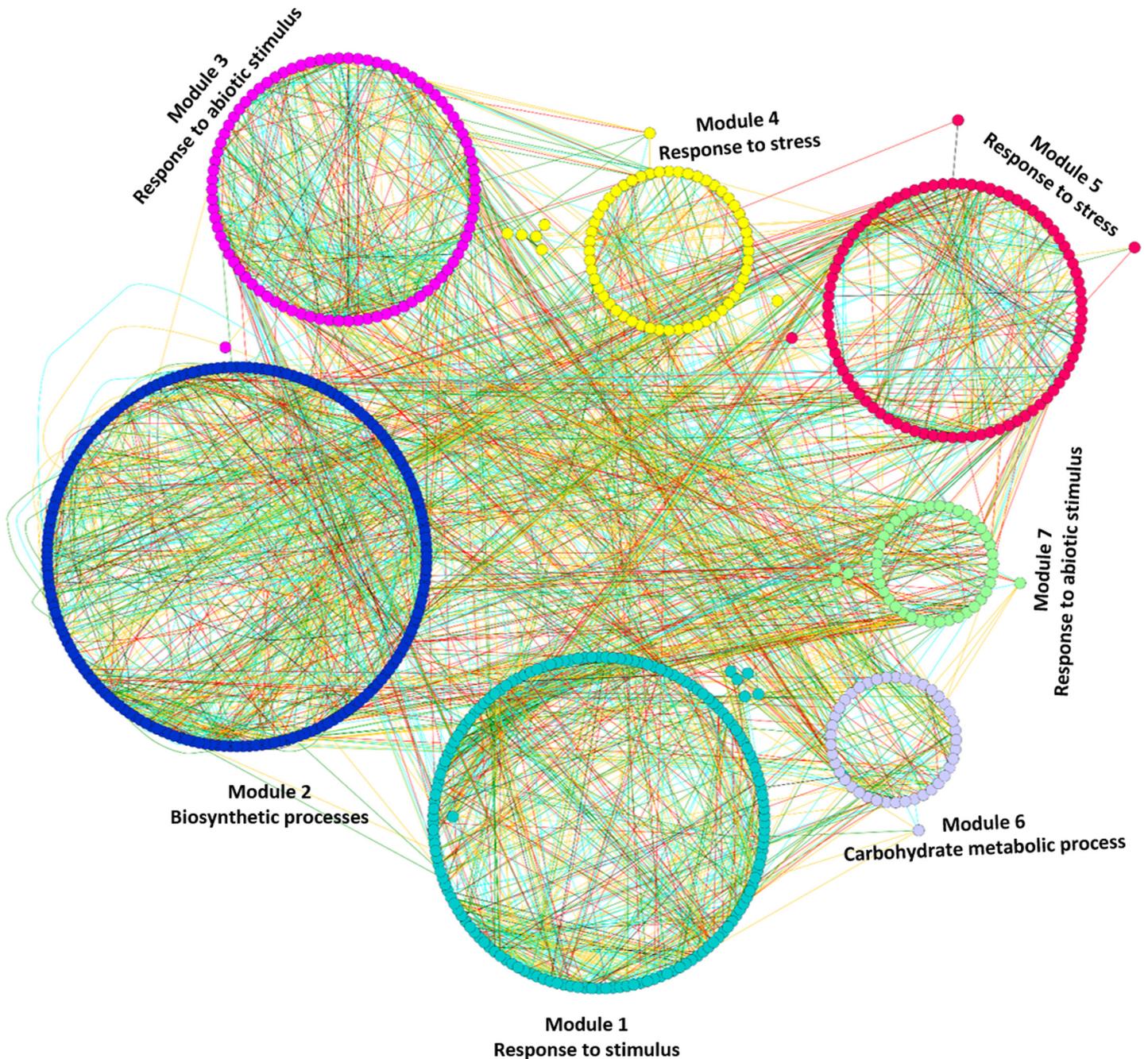


Fig 3. Mapping of enriched GO terms onto the very sub-optimal temperature stress responsive gene co-expression network. Modules with different colours indicates that these modules were significantly enriched with GO terms. The different edge colors indicates different range of HRR values and as follows: Cyan: HRR1-HRR5; Red: HRR6-HRR10; Green: HRR11-HRR15 and Orange yellow: HRR16-HRR20.

doi:10.1371/journal.pone.0148771.g003

the module are up-regulated, while the rest of the genes are down-regulated by the very sub-optimal temperature stress.

Module 7 consists of 42 gene nodes and is also enriched with abiotic stimulus response genes (GO:0009628, FDR = 0.0044). There are a total of 13 genes whose functions are associated with the abiotic stimulus responses, which represents 37% of the annotated genes in the

Table 4. Details of identified Transcription Factors (TFs). This table contains information for each identified TF with respect to the module number to which the TF belongs, the corresponding probe-set and gene ID, description of the TF, whether the identified TF is up- or down-regulated, the enriched GO term identified by the AgriGO tool and the number of genes connected to each TF.

Module	Probeset ID	Gene ID	Description	Regulation	Enriched GO Term	Intramodular connectivity
1	Gma.4281.1.S1_at	Glyma18g44560.1	WRKY70	Down	response to stimulus (GO:0050896)	11
1	Gma.593.2.S1_a_at	Glyma17g10250.2	MYB173	Down	response to stimulus (GO:0050896)	15
1	GmaAffx.36677.1.S1_at	Glyma04g33210.1	MYB16	Down	response to stimulus (GO:0050896)	14
1	Gma.4324.3.S1_a_at	Glyma20g31210.4	NAC17	Up	response to stimulus (GO:0050896)	5
2	GmaAffx.32030.1.A1_at	Glyma04g08990.1	GATA9	Down	response to stimulus (GO:0050896)	14
2	GmaAffx.20956.1.S1_at	Glyma10g05560.1	MYB118	Up	response to stimulus (GO:0050896)	13
3	Gma.13676.1.A1_at	Glyma03g41270.1	WHIRLY2	Up	response to stimulus (GO:0050896)	18
3	GmaAffx.50811.2.S1_at	Glyma15g08480.4	NAC73	Up	response to stimulus (GO:0050896)	18
7	GmaAffx.16512.1.S1_at	Glyma13g09010.1	MYB25	Up	response to stimulus (GO:0050896)	9

doi:10.1371/journal.pone.0148771.t004

module. Sixty-five percent% of genes in the module are up-regulated, while rest of the genes are down-regulated by the very sub-optimal temperature stress.

Transcription factors associated with the very sub-optimal temperature stress. Transcription factors (TFs) play a key role in the biological processes involved in plant growth and development, biosynthetic pathways and biotic and abiotic stress tolerances by regulating the downstream target genes. Recent studies have revealed their role in various biological processes in *Glycine max* [41, 42], *Brassica rapa* [43, 44], *Brassica oleracea* [45], *Vitis vinifera* [46] and *Setaria italica* [47–51]. In this study, we have identified several TFs only in the very sub-optimal temperature stress which may be regulate downstream target genes involved in stimulus responses contained in the GO term and the details of identified TFs are displayed in Table 4.

Module 1 contains four genes encoding TFs from three families and are MYB16 (Glyma04g33210.1), MYB173 (Glyma17g10250.2), NAC17 (Glyma20g31210.4) and WRKY70 (Glyma18g44560.1). Out of four TFs, three TFs were down-regulated, while one TF is up-regulated by the very sub-optimal temperature stress. MYB173 is down-regulated by the very sub-optimal temperature stress and had the highest number of node genes connected (S13 Table) in module 1 (Fig 4A). Previous studies revealed that the MYB family TFs were induced in various stress responses and in stress tolerance [43, 49, 52] which indicates that the down-regulated MYB family TFs in bambara groundnut needs to be induced in order to withstand the decreasing temperature. Module 2 contains two genes encoding TFs from two families and are GATA9 (Glyma04g08990.1) and MYB118 (Glyma10g05560.1). GATA9 is down-regulated, while MYB118 is up-regulated by the very sub-optimal temperature stress. GATA9 had the highest number of intramodular connections (S13 Table) in module 2 (Fig 4B). Previous studies have revealed that the GATA family TFs were differentially expressed in high temperatures and drought [53] which indicates that GATA9 is induced in response to the decreasing temperature.

Similarly, module 3 contains two genes encoding two TFs from different families. WHIRLY2 (Glyma03g41270.1) and NAC73 (Glyma15g08480.4). Interestingly, both TFs were up-regulated by the very sub-optimal temperature stress which had the joint highest intramodular connectivity (S13 Table) among identified TFs (Fig 5A and 5B). WHIRLY genes are known to be an important part of disease resistance mechanisms in *Arabidopsis* [54] and induced during salinity stress in *Hordeum vulgare* [52]. So far hundreds of NAC genes have been identified in various plants. Studies on NAC genes revealed that these genes regulate salt and drought tolerance [55, 56]. Interestingly, both WHIRLY2 and NAC73 genes were up-regulated by the very sub-optimal temperature stress which implies that these TFs may have an

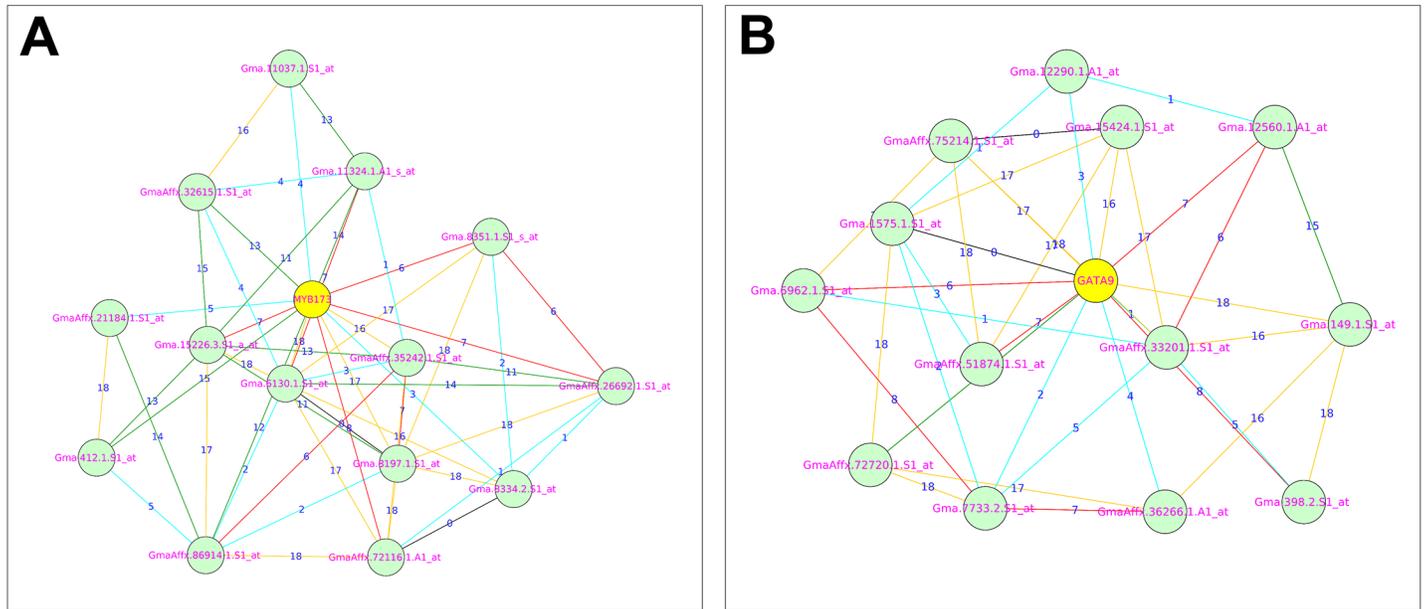


Fig 4. Co-expression relationships of MYB173 and GATA9 transcription factors with their related genes: [A] Shows the co-expression relationships of MYB173 with its related genes. [B] Shows the co-expression relationships of GATA9 with its related genes. The numbers on top of edges indicates the highest reciprocal rank between node genes. The different edge colors indicates different range of HRR values and as follows: Cyan: HRR1-HRR5; Red: HRR6-HRR10; Green: HRR11-HRR15 and Orange yellow: HRR16-HRR20.

doi:10.1371/journal.pone.0148771.g004

important role in raising low temperature tolerance by inducing genes associated with cold stimulus response in bambara groundnut. Module 3 is the only module that contains

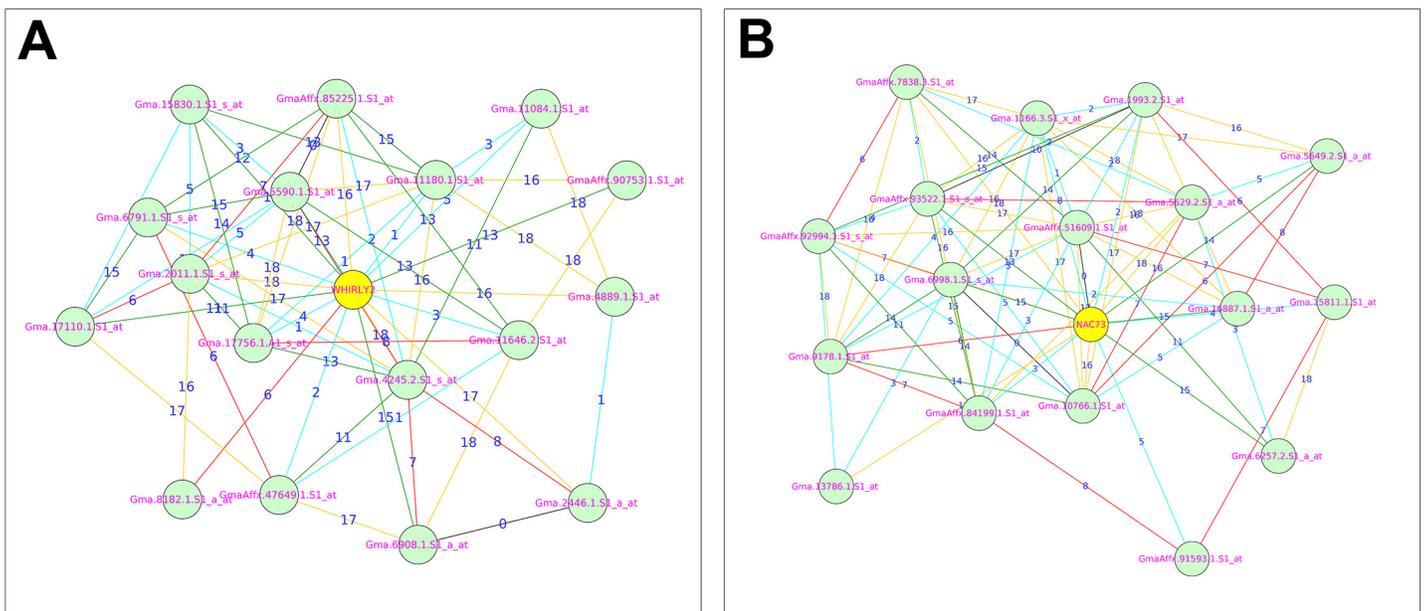


Fig 5. Co-expression relationships of WHIRLY2 and NAC73 genes with their related genes: [A] Shows the co-expression relationships of WHIRLY2 with its related genes. The details of related genes are present in [S12 Table](#). [B] Shows the co-expression relationships of NAC73 with its related genes. The numbers on top of the edges indicates the highest reciprocal rank between node genes. The different edge colors indicates different range of HRR values and as follows: Cyan: HRR1-HRR5; Red: HRR6-HRR10; Green: HRR11-HRR15 and Orange yellow: HRR16-HRR20.

doi:10.1371/journal.pone.0148771.g005

up-regulated TFs indicating that this module may be significantly associated with the very sub-optimal temperature stress. Further investigation is required to confirm whether these two TFs are involved in raising the very sub-optimal temperature stress tolerance in bambara groundnut. Finally, there is only one gene encoding a TF in module 7 (MYB25 Glyma13g09010.1) which is up-regulated by the very sub-optimal temperature stress, consistent with the results of previous studies [43, 52].

Discussion

Despite being nutritionally rich, bambara groundnut is still cultivated mainly as landraces. Bambara groundnut is well adapted to drought conditions. In contrast, low temperature stress ($<T_0$) causes loss of potential yield in bambara groundnut in a similar way to other crops by limiting growth and development [5, 6] and hence restricts farmers' option for growing the crop, particularly where the temperature is close to T_0 . Hence, the development of low-temperature stress tolerant Bambara groundnut varieties is important in order to provide more options for farmers.

So far most of the existing studies used either array based or sequence based methods to identify stress responsive candidate genes of crops under different kind of stresses [16–19]. But for bambara groundnut so far there is no chip is available and deep sequencing based methods are still expensive and hence we have used a heterologous microarray (Soybean Affymetrix GeneChip™) coupled with network-based analysis to identify the genes and gene modules associated with low temperatures (23^o and 18^o C) in bambara groundnut. In this study, we developed a computational approach (Fig 1) to identify the genes and gene modules significantly associated with the low temperature stress in bambara groundnut and which can be used to identify the genes and gene modules associated within any particular crop.

In this study, we identified 375 and 659 genes which are differentially expressed in the sub-optimal and the very sub-optimal temperature stresses respectively, 110 of which are commonly shared between stress levels. On average, 41% of annotated genes present in five functional modules of the very sub-optimal temperature stress were significantly enriched with 'response to stimuli', 'response to abiotic stimuli' and 'stress response', genes while none of the modules of the sub-optimal temperature stress treatment were significantly enriched for any stress-related biological process. Among all the annotated genes present in various modules of very sub-optimal temperature stress, there are several important genes which are up-regulated by very sub-optimal temperature stress condition including, HSP, Chaperone DnaJ and WD40 repeat genes. These genes mainly play crucial roles in correct folding of low-temperature stress-accumulated misfolded proteins in order to withstand low-temperatures. Functional annotation of the detected modules identified nine transcription factors (TF) belonging to five transcription factor families in the very sub-optimal temperature stress only (MYB, NAC, WRKY, WHIRLY and GATA). These TFs have already been shown to be significantly associated with various abiotic stresses in various crops [43, 52–56] and therefore, these TFs may regulate downstream target genes involved in various stress responses related to biological processes in bambara groundnut. The identified gene modules and genes could be useful in developing in low-temperature stress tolerant bambara groundnut varieties, via analysis of the existing genetic diversity available within bambara groundnut, through germplasm collections, such as those held at the International Institute for Tropical Agriculture in Nigeria.

Supporting Information

S1 Fig. Validation of 52 DE genes using heatmap.
(TIF)

S1 Table. List of identified DE genes in Sub_optimal temperature stress.
(XLSX)

S2 Table. List of identified DE genes in Very sub_optimal temperature stress.
(XLSX)

S3 Table. List of common DE genes between two stresses.
(XLSX)

S4 Table. Data used to generate heatmaps.
(XLSX)

S5 Table. HRR-based co-expression network of Sub_optimal temperature stress.
(XLSX)

S6 Table. HRR-based co-expression network of Very sub_optimal temperature stress.
(XLSX)

S7 Table. Detected modules in Sub_optimal temperature stress.
(XLSX)

S8 Table. Detected modules in Very sub_optimal temperature stress.
(XLSX)

S9 Table. GO annotation of Sub_optimal temperature stress.
(XLSX)

S10 Table. Gene-level GO annotation of Sub_optimal temperature stress.
(XLSX)

S11 Table. GO annotation of Very sub_optimal temperature stress.
(XLSX)

S12 Table. Gene-level GO annotation of Very sub_optimal temperature stress.
(XLSX)

S13 Table. List of connected nodes to each of the transcription factor.
(XLSX)

Acknowledgments

We would like to acknowledge Ray Wilson for his work on cDNA library preparation.

Author Contributions

Conceived and designed the experiments: VSB JT. Performed the experiments: VSB JT. Analyzed the data: VSB JT SM FM KM. Contributed reagents/materials/analysis tools: VSB JT SM FM KM JM MB VW STM. Wrote the paper: VSB JT SM FM KM.

References

1. Azam-Ali SN, Sesay A, Karikari SK, Massawe FJ, Aguilar-Manjarrez J, Bannayan M, et al. Assessing the potential of an underutilized crop—a case study using Bambara groundnut. *Experimental Agriculture* 2001, 37:433–472.
2. Ahmad NS, Basu SM, Redjeki ES, Murchie E, Massawe F, Azam-Ali SN, et al. Developing Genetic Mapping and Marker-Assisted Techniques in Bambara Groundnut (*Vigna subterranea* L.) Breeding. *Acta Horticulturae* 2013, 979, ISHS.

3. Makanda I, Tongoona P, Madamba R, Icishahayo D, Derera J. Evaluation of bambara groundnut varieties for off-season production in Zimbabwe. *African Crop Science Journal* 2009, 16, 3:175–183.
4. Ibraheem AS, Debbie S, Azam-Ali S. Temperature and Drought Stress Effects on Growth and Development of Bambara Groundnut (*Vigna Subterranea* L.). *Experimental Agriculture* 2014, 50(1):72–89.
5. Ha TN, Jorg L, Peter S, Orlene GP. Low temperature stress in maize (*Zea mays* L.) induces genes involved in photosynthesis and signal transduction as studied by suppression subtractive hybridization. *Plant Physiology and Biochemistry* 2009, 47:116–122. doi: [10.1016/j.plaphy.2008.10.010](https://doi.org/10.1016/j.plaphy.2008.10.010) PMID: [19042136](https://pubmed.ncbi.nlm.nih.gov/19042136/)
6. Shizen O, Tomoaki M, Shigehisa S. Low temperature stress at different flower developmental stages affects pollen development, pollination, and pod set in soybean. *Environmental and Experimental Botany* 2010, 69:56–62.
7. Linnemann AR, Azam-Ali SN. Bambara groundnut (*Vigna subterranean* (L.)Verdc.). In *Underutilized Crops, Pulses and Vegetables* (Ed Williams J. T.) 1993, London, UK, Champan and Hall.
8. Atkinson NJ, Lilley CJ, Urwin PE. Identification of Genes Involved in the Response of Arabidopsis to Simultaneous Biotic and Abiotic Stresses. *Plant Physiology* 2013, 162(4):2028–2041. doi: [10.1104/pp.113.222372](https://doi.org/10.1104/pp.113.222372) PMID: [23800991](https://pubmed.ncbi.nlm.nih.gov/23800991/)
9. Shinji K, Chris B, Michael D, Hong W, Susan B, Kiyoshi K, et al. Gene Expression Profiles during the Initial Phase of Salt Stress in Rice. *The Plant Cell* 2001, 13, 4:889–905. PMID: [11283343](https://pubmed.ncbi.nlm.nih.gov/11283343/)
10. Gregory AW, Roayaei JA, Quiñones OA, Schneider K. A microarray analysis for differential gene expression in the soybean genome using Bioconductor and R. *Brief Bioinform.* 2007, 8(6):415–31. PMID: [17906332](https://pubmed.ncbi.nlm.nih.gov/17906332/)
11. Rob A, Paxton P, Zhanjun F, Ryan M, Paul D, Gregory BM, et al. Transcriptome and Selected Metabolite Analyses Reveal Multiple Points of Ethylene Control during Tomato Fruit Development. *The Plant Cell* 2005, 17, 11:2954–2965. PMID: [16243903](https://pubmed.ncbi.nlm.nih.gov/16243903/)
12. Hammond JP, Broadley MR, Craigon DJ, Higgins J, Emmerson ZF, Townsend HJ, et al. Using genomic DNA-based probe-selection to improve the sensitivity of high-density oligonucleotide arrays when applied to heterologous species. *Plant Methods* 2005, 1, 10. PMID: [16280083](https://pubmed.ncbi.nlm.nih.gov/16280083/)
13. Bagnaresi P, Moschella A, Beretta O, Vitulli F, Ranalli P, Perata P. Heterologous microarray experiments allow the identification of the early events associated with potato tuber cold sweetening. *BMC Genomics* 2008, 9:176. doi: [10.1186/1471-2164-9-176](https://doi.org/10.1186/1471-2164-9-176) PMID: [18416834](https://pubmed.ncbi.nlm.nih.gov/18416834/)
14. Fabi JP, Seymour GB, Graham NS, Martin RB, Sean TM, Franco ML, et al. Analysis of ripening-related gene expression in papaya using an Arabidopsis-based microarray. *BMC Plant Biology* 2012, 12:242. doi: [10.1186/1471-2229-12-242](https://doi.org/10.1186/1471-2229-12-242) PMID: [23256600](https://pubmed.ncbi.nlm.nih.gov/23256600/)
15. Davey MW, Graham NS, Vanholme B, Swennen R, May ST, Keulemans J. Heterologous oligonucleotide microarrays for transcriptomics in a non-model species; a proof-of-concept study of drought stress in *Musa*. *BMC Genomics* 2009, 10:436. doi: [10.1186/1471-2164-10-436](https://doi.org/10.1186/1471-2164-10-436) PMID: [19758430](https://pubmed.ncbi.nlm.nih.gov/19758430/)
16. Obayashi T, Hayashi S, Saeki M, Ohta H, Kinoshita K. ATTED-II provides coexpressed gene networks for Arabidopsis. *Nucleic Acids Research* 2009, 37:D987–D991. doi: [10.1093/nar/gkn807](https://doi.org/10.1093/nar/gkn807) PMID: [18953027](https://pubmed.ncbi.nlm.nih.gov/18953027/)
17. Zhang L, Yu S, Zuo K, Luo L, Tang K. Identification of Gene Modules Associated with Drought Response in Rice by Network-Based Analysis. *PLoS ONE* 2012, 7(5):e33748. doi: [10.1371/journal.pone.0033748](https://doi.org/10.1371/journal.pone.0033748) PMID: [22662107](https://pubmed.ncbi.nlm.nih.gov/22662107/)
18. Wong DC, Sweetman C, Ford CM. Annotation of gene function in citrus using gene expression information and co-expression networks. *BMC Plant Biology* 2014, 14:186. doi: [10.1186/1471-2229-14-186](https://doi.org/10.1186/1471-2229-14-186) PMID: [25023870](https://pubmed.ncbi.nlm.nih.gov/25023870/)
19. Zhu M, Deng X, Joshi T, Xu D, Stacey G, Cheng J. Reconstructing differentially co-expressed gene modules and regulatory networks of soybean cells. *BMC Genomics* 2012, 13:437. doi: [10.1186/1471-2164-13-437](https://doi.org/10.1186/1471-2164-13-437) PMID: [22938179](https://pubmed.ncbi.nlm.nih.gov/22938179/)
20. Wong DC, Sweetman C, Drew DP, Ford CM. VTCdb: a gene co-expression database for the crop species *Vitis vinifera* (grapevine). *BMC Genomics* 2013, 14:882. doi: [10.1186/1471-2164-14-882](https://doi.org/10.1186/1471-2164-14-882) PMID: [24341535](https://pubmed.ncbi.nlm.nih.gov/24341535/)
21. Wang M, Verdier J, Benedito VA, Tang Y, Murray JD, Ge Y. LegumeGRN: A Gene Regulatory Network Prediction Server for Functional and Comparative Studies. *PLoS ONE* 2013, 8(7): e67434. doi: [10.1371/journal.pone.0067434](https://doi.org/10.1371/journal.pone.0067434) PMID: [23844010](https://pubmed.ncbi.nlm.nih.gov/23844010/)
22. Yoshiyuki O, Hideyuki S, Nozomu S, Daisuke S. CoP: a database for characterizing co-expressed gene modules with biological information in plants. *Bioinformatics* 2010, 26 (9): 1267–1268. doi: [10.1093/bioinformatics/btq121](https://doi.org/10.1093/bioinformatics/btq121) PMID: [20305269](https://pubmed.ncbi.nlm.nih.gov/20305269/)
23. Won CY, Yongbin Y, Kitae S, Cheo SJ, Byung-Moo L. PLANEX: the plant co-expression database. *BMC Plant Biology* 2013, 13:83. doi: [10.1186/1471-2229-13-83](https://doi.org/10.1186/1471-2229-13-83) PMID: [23688397](https://pubmed.ncbi.nlm.nih.gov/23688397/)

24. Dash S, Van Hemert J, Hong L, Wise RP, Dickerson JA. PLEXdb: gene expression resources for plants and plant pathogens. *Nucleic Acids Research* 2012, 40 (D1): D1194–D1201.
25. Srinivasasainagendra V, Page GP, Mehta T, Coulibaly I, Loraine AE. CressExpress: A Tool for Large-Scale Mining of Expression Data from Arabidopsis. *Plant Physiology* 2008, 147 (3):1004–1016. doi: [10.1104/pp.107.115535](https://doi.org/10.1104/pp.107.115535) PMID: [18467456](https://pubmed.ncbi.nlm.nih.gov/18467456/)
26. Gautier L, Cope L, Bolstad BM, Irizarry RA. "affy—analysis of Affymetrix GeneChip data at the probe level." *Bioinformatics* 2004, 20(3):307–315. PMID: [14960456](https://pubmed.ncbi.nlm.nih.gov/14960456/)
27. Benjamini Y, Hochberg Y. "Controlling the false discovery rate: a practical and powerful approach to multiple testing". *Journal of the Royal Statistical Society, Series B* 1995, 57(1):289–300.
28. LifeScope: <http://www.appliedbiosystems.com/lifescopy>
29. Simon A, Wolfgang H. Differential expression analysis for sequence count data. *Genome Biology* 2010, 11: R106. doi: [10.1186/gb-2010-11-10-r106](https://doi.org/10.1186/gb-2010-11-10-r106) PMID: [20979621](https://pubmed.ncbi.nlm.nih.gov/20979621/)
30. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J. Mol. Biol.* 1990, 215:403–410. PMID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
31. Mutwil M, Usadel B, Schütte M, Loraine A, Ebenhöf O, Persson S. Assembly of an Interactive Correlation Network for the Arabidopsis Genome Using a Novel Heuristic Clustering Algorithm. *Plant Physiology* 2010, 152(1):29–43. doi: [10.1104/pp.109.145318](https://doi.org/10.1104/pp.109.145318) PMID: [19889879](https://pubmed.ncbi.nlm.nih.gov/19889879/)
32. Michael S, Keiichiro O, Johannes R, Peng-Liang W, Trey I. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 2011, 27(3):431–432. doi: [10.1093/bioinformatics/btq675](https://doi.org/10.1093/bioinformatics/btq675) PMID: [21149340](https://pubmed.ncbi.nlm.nih.gov/21149340/)
33. Zhou D, Xin Z, Yi L, Zhenhai Z, Zhen S. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research* 2010, 38: W64–W70. doi: [10.1093/nar/gkq310](https://doi.org/10.1093/nar/gkq310) PMID: [20435677](https://pubmed.ncbi.nlm.nih.gov/20435677/)
34. Yang JK, Sue KK, Moon YK, Puji L, Kil HK, Bo-Keun H, et al. Genome sequence of mungbean and insights into evolution within Vigna species. *Nature Communications*. 2014, 5, 5443. doi: [10.1038/ncomms6443](https://doi.org/10.1038/ncomms6443) PMID: [25384727](https://pubmed.ncbi.nlm.nih.gov/25384727/)
35. Usadel B, Obayashi T, Mutwil M, Giorgi FM, Bassel GW, Tanimoto M, et al. Co-expression tools for plant biology: Opportunities for hypothesis generation and caveats. *Plant Cell Environ.* 2009, 32: 1633–1651. doi: [10.1111/j.1365-3040.2009.02040.x](https://doi.org/10.1111/j.1365-3040.2009.02040.x) PMID: [19712066](https://pubmed.ncbi.nlm.nih.gov/19712066/)
36. Linyong M, John LVH, Sudhansu D, Julie AD. Arabidopsis gene co-expression network and its functional modules. *BMC Bioinformatics* 2009, 10:346. doi: [10.1186/1471-2105-10-346](https://doi.org/10.1186/1471-2105-10-346) PMID: [19845953](https://pubmed.ncbi.nlm.nih.gov/19845953/)
37. Bin C, Cheng-Hui L, Jian H. Systematic identification of cell-wall related genes in populus based on analysis of functional modules in Co-expression network. *PLoS ONE* 2014, 9(4):e95176. doi: [10.1371/journal.pone.0095176](https://doi.org/10.1371/journal.pone.0095176) PMID: [24736620](https://pubmed.ncbi.nlm.nih.gov/24736620/)
38. Al-Whaibi Mohamed H. Plant heat-shock proteins: A mini review. *Journal of King Saud University—Science* 2011, 23:2, pp:139–150.
39. Kong F, Deng Y, Zhou B, Wang G, Wang Y, Meng Q. A chloroplast-targeted DnaJ protein contributes to maintenance of photosystem II under chilling stress. *Journal of Experimental Botany* 2014, 65 (1):143–158. doi: [10.1093/jxb/ert357](https://doi.org/10.1093/jxb/ert357) PMID: [24227338](https://pubmed.ncbi.nlm.nih.gov/24227338/)
40. Mishra AK, Muthamilarasan M, Khan Y, Parida SK, Prasad M. Genome-Wide Investigation and Expression Analyses of WD40 Protein Family in the Model Plant Foxtail Millet (*Setaria italica* L.). *PLoS ONE* 2014, 9(1):e86852. doi: [10.1371/journal.pone.0086852](https://doi.org/10.1371/journal.pone.0086852) PMID: [24466268](https://pubmed.ncbi.nlm.nih.gov/24466268/)
41. Guangjun Y, Hongliang X, Shuyang X, Yajuan Q, Yaxuan L, Yueming Y, et al. The large soybean (*Glycine max*) WRKY TF family expanded by segmental duplication events and subsequent divergent selection among subgroups. *BMC Plant Biology* 2013, 13:148. doi: [10.1186/1471-2229-13-148](https://doi.org/10.1186/1471-2229-13-148) PMID: [24088323](https://pubmed.ncbi.nlm.nih.gov/24088323/)
42. Valliyodan B, Van Toai TT, Alves JD, Patricia de Fátima PG, Jeong DL, Felix BF, et al. Expression of Root-Related Transcription Factors Associated with Flooding Tolerance of Soybean (*Glycine max*). *International Journal of Molecular Sciences* 2014, 15(10):17622–17643. doi: [10.3390/ijms151017622](https://doi.org/10.3390/ijms151017622) PMID: [25268626](https://pubmed.ncbi.nlm.nih.gov/25268626/)
43. Zhen W, Jun T, Rong H, Peng W, Xi-Lin H, Xiao-Ming S, et al. Genome-wide analysis of the R2R3-MYB transcription factor genes in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) reveals their stress and hormone responsive patterns. *BMC Genomics* 2015, 16:17. doi: [10.1186/s12864-015-1216-y](https://doi.org/10.1186/s12864-015-1216-y) PMID: [25613160](https://pubmed.ncbi.nlm.nih.gov/25613160/)
44. Kayum MA, Jung HJ, Park J, Ahmed NU, Saha G, Yang TJ, et al. Identification and expression analysis of WRKY family genes under biotic and abiotic stresses in *Brassica rapa*. *Mol Genet Genomics* 2015, 290(1):79–95. doi: [10.1007/s00438-014-0898-1](https://doi.org/10.1007/s00438-014-0898-1) PMID: [25149146](https://pubmed.ncbi.nlm.nih.gov/25149146/)
45. Bin Z, Zongli H, Yanjie Z, Yali L, Shuang Z, Guoping C. A putative functional MYB transcription factor induced by low temperature regulates anthocyanin biosynthesis in purple kale (*Brassica Oleracea* var.

- acephala f. tricolor). *Plant Cell Reports* 2012, 31, 2: 281–289. doi: [10.1007/s00299-011-1162-3](https://doi.org/10.1007/s00299-011-1162-3) PMID: [21987119](https://pubmed.ncbi.nlm.nih.gov/21987119/)
46. Wang L, Zhu W, Fang L, Sun X, Su L, Liang, Wang N, et al. Genome-wide identification of WRKY family genes and their response to cold stress in *Vitis vinifera*. *BMC Plant Biology* 2014, 14:103. doi: [10.1186/1471-2229-14-103](https://doi.org/10.1186/1471-2229-14-103) PMID: [24755338](https://pubmed.ncbi.nlm.nih.gov/24755338/)
 47. Puranik S, Sahu PP, Mandal SN, Bonthala VS, Parida SK, Prasad M. Comprehensive Genome-Wide Survey, Genomic Constitution and Expression Profiling of the NAC Transcription Factor Family in Foxtail Millet (*Setaria italica* L.). *PLoS ONE* 2013, 8(5): e64594. doi: [10.1371/journal.pone.0064594](https://doi.org/10.1371/journal.pone.0064594) PMID: [23691254](https://pubmed.ncbi.nlm.nih.gov/23691254/)
 48. Muthamilarasan M, Bonthala VS, Awdhesh KM, Rohit K, Yusuf K, Riti R, et al. C2H2 type of zinc finger transcription factors in foxtail millet define response to abiotic stresses. *Functional & Integrative Genomics* 2014, 14, 3:531–543.
 49. Muthamilarasan M, Khandelwal R, Yadav CB, Bonthala VS, Khan Y, Prasad M. Identification and Molecular Characterization of MYB Transcription Factor Superfamily in C4 Model Plant Foxtail Millet (*Setaria italica* L.). *PLoS ONE* 2014, 9(10):e109920. doi: [10.1371/journal.pone.0109920](https://doi.org/10.1371/journal.pone.0109920) PMID: [25279462](https://pubmed.ncbi.nlm.nih.gov/25279462/)
 50. Muthamilarasan M, Bonthala VS, Khandelwal R, Jaishankar J, Shweta S, Nawaz K, et al. Global analysis of WRKY transcription factor superfamily in *Setaria* identifies potential candidates involved in abiotic stress signaling. *Front. Plant Sci.* 2015, 6:910. doi: [10.3389/fpls.2015.00910](https://doi.org/10.3389/fpls.2015.00910) PMID: [26635818](https://pubmed.ncbi.nlm.nih.gov/26635818/)
 51. Lata C, Mishra AK, Muthamilarasan M, Bonthala VS, Khan Y, Prasad M. Genome-Wide Investigation and Expression Profiling of AP2/ERF Transcription Factor Superfamily in Foxtail Millet (*Setaria italica* L.). *PLoS ONE* 2014, 9(11): e113092. doi: [10.1371/journal.pone.0113092](https://doi.org/10.1371/journal.pone.0113092) PMID: [25409524](https://pubmed.ncbi.nlm.nih.gov/25409524/)
 52. Runhong G, Ke D, Guimei G, Zhizhao D, Zhiwei C, Liang L, et al. Comparative Transcriptional Profiling of Two Contrasting Barley Genotypes under Salinity Stress during the Seedling Stage. *International Journal of Genomics* 2013, Article ID 972852, 19.
 53. Ankur RB, Gopal J, Bharti K, Vidhi M, Priyanka A, Ritu P, et al. Global insights into high temperature and drought stress regulated genes by RNA-Seq in economically important oilseed crop *Brassica juncea*. *BMC Plant Biology* 2015, 15:9. doi: [10.1186/s12870-014-0405-1](https://doi.org/10.1186/s12870-014-0405-1) PMID: [25604693](https://pubmed.ncbi.nlm.nih.gov/25604693/)
 54. Desveaux D, Subramaniam R, Després C, Mess JN, Lévesque C, Fobert PR, et al. “A “Whirly” transcription factor is required for salicylic acid-dependent disease resistance in *Arabidopsis*,” *Developmental Cell* 2004, 6, 2:229–240. PMID: [14960277](https://pubmed.ncbi.nlm.nih.gov/14960277/)
 55. Hisako O, Kouji S, Koji D, Toshifumi N, Yasuhiro O, Kazuo M, et al. Comprehensive Analysis of NAC Family Genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Research* 2003, 10(6):239–247. PMID: [15029955](https://pubmed.ncbi.nlm.nih.gov/15029955/)
 56. Mingku Z, Guoping C, Jianling Z, Yanjie Z, Qiaoli X, Zhiping Z, et al. The abiotic stress-responsive NAC-type transcription factor SINAC4 regulates salt and drought tolerance and stress-related genes in tomato (*Solanum lycopersicum*). *Plant Cell Reports* 2014, 33, 11:1851–1863. doi: [10.1007/s00299-014-1662-z](https://doi.org/10.1007/s00299-014-1662-z) PMID: [25063324](https://pubmed.ncbi.nlm.nih.gov/25063324/)