Informative Established Sequence Traits in Tepary Bean (*Phaseolus Acutifolius* A. Gray) for Drought Stress

Satya S S Narina

Virginia State University-ARS, Petersburg, VA, USA

USDA-ARS, Stoneville, MS, USA

Brian E Scheffler

USDA-ARS, Stoneville, MS, USA

Linda A Ballard

USDA-ARS, Stoneville, MS, USA

Sheron E Simpson

USDA-ARS, Stoneville, MS, USA

Ramesh Buyyarapu

Dow-Agro Sciences, Indianapolis, Indiana, USA

Rao K Kottapalli

USDA-ARS, Texas Tech, Texas, USA

Brian L Sayre

Department of Biology, Virginia State University, Petersburg, VA, USA

Harbans L Bhardwaj

Virginia State University-ARS, Petersburg, VA, USA

Abstract

Drought stress, a major abiotic stress adversely effects crop growth and yield. Tepary bean was identified as one of the drought adapted legumes for arid and semi-arid regions. Drought responsive functional genes in tepary bean were identified through current investigation. The complementary DNAs (cDNAs) prepared from leaf messenger RNA (mRNA) of three

genotypes (#10, #15 and #20) were sequenced using 454, and 54,165 Expressed Sequence Tags (ESTs) were generated and functionally annotated.

Results: A total of 18523, 16190 and 19452 unique ESTs respectively from cDNA libraries from genotype #10, genotype #15 and genotype #20 had significant hits to NCBI's non redundant protein database (e value < 1e-05). All the ESTs were annotated using Gene Ontology (GO) terms and drought related genes were identified. A total of 2922 microsatellite markers were designed and 648 were selected for testing and 538 were successfully amplified and were polymorphic.

Conclusion: The new EST resource provides more gene sequence information compared to the previously developed 700 ESTs for drought stress. The polymorphic microsatellite markers identified will be used to characterize the tepary bean germplasm for various traits of importance and select superior lines.

1. Background

Tepary bean (*Phaseolus acutifolius*, Fabaceae), native to the southwestern United States & Mexico, was a major item of trade with the agrarian tribes of the pueblos of the southwest, such as the Papago, Zuni, Hopi, Pima, and others. It is much more drought-resistant than common beans and is grown in desert and semi-desert conditions from Arizona through Mexico to Costa Rica. Tepary bean seed has excellent nutritional value (Bhardwaj and Hamama, 2004, Blair et al., 2010, Narina et al., 2014, 2015) including anticancer and antidiabetic properties.

The research on tepary bean is gaining momentum since 2002 in Virginia (Bhardwaj et al., 2002 and Hamama and Bhardwaj, 2002) due to increase in human consumption (Scheerens et al., 1983), its economical and nutritional values (Narina et al., 2014, & 2015). Drought studies are in progress in locations like Nebraska and Virginia to screen drought tolerant varieties (Narina et al., 2012) for food, feed and forage (Bhardwaj, 2013) and utilize these traits by introgression through breeding practices, to develop improved varieties with drought tolerance in common bean as well as in tepary bean (Urea, 2009).

There are several mechanisms like drought avoidance, escape and tolerance observed in tepary bean (Mohamed et al., 2002 and 2005), and in general used by plants for survival in adverse environments such as drought, high temperature and salinity. Wild common bean can be used as a source for exploitation of variation for drought tolerance (Cortes et al., 2013). One of the inevitable consequences of drought stress is enhanced reactive oxygen species (ROS) production in the different cellular compartments, namely in the chloroplasts, the peroxisomes and the mitochondria. ROS signaling is linked to abscissic acid (ABA), Ca2+ fluxes and sugar sensing and is likely to be involved both upstream and downstream of the ABA-dependent signaling pathways under drought stress (Maria, 2008). The rise of PvPIP2; 1 gene expression and PIP1 protein abundance in the leaves of *P. vulgaris* plants subjected to drought was correlated with a decline in the transpiration rate (Ricardo, 2006).

Herbaceous native shrub species have evolved mechanisms to withstand water deficit and drought stress (Blum, 2011) through genes involved in anatomical, cellular and metabolic components. Drought tolerant variety (Pinto Villa) was observed with well-developed root vascular tissues and a number of upregulated functional genes for abiotic stress in common bean (Montero-Tavera, et al., 2008). Plants adapt to drought stress through response mechanisms like osmotic adjustment, reduction in DNA content, changes in metabolic pathways with the accumulation of compounds like proline, appearance of new proteins, synthesis of the hormone, and ABA(Caplan et al., 1989).

Tepary bean is an important crop in North America especially for farmers in Arizona and Virginia. Identification of transcriptome regions associated with drought tolerance would enable breeders to develop improved cultivars through marker-assisted selection. EST analysis is an efficient approach for gene discovery and identification of polymorphic markers (Gupta and Varshney, 2000, Varshney et al., 2013) such as microsatellites and single nucleotide polymorphisms (SNPs). The lack of ESTs available for tepary bean in public data bases (only 751 available at NCBI database on August 1, 2010 for tepary bean) created interest to conduct research and develop crop specific genomic libraries for identification of highly abundantly expressed genes related to various economically and nutritionally important traits.

Therefore the objective of the current study is to construct and sequence the cDNA library for three tepary bean genotypes showing differential response to drought, identify the genes involved in drought stress tolerance by analyzing the ESTs developed and to generate polymorphic EST - Simple Sequence Repeat (EST-SSR) markers for future germplasm characterization, genetic linkage mapping and QTL analysis in tepary bean. Six tepary bean lines were selected out of 200 Plant Introduction (PI) lines screened for drought stress at three locations by Bhardwaj during 2008-2009 (unpublished). Three genotypes with the collection numbers 10, 15 and 20 were selected for the current gene expression studies.

2. Methods

2.1 Tissue Source and RNA isolation

The drought susceptible (#15) and resistant (#10 and 20) genotypes of tepary bean were selected for current study (Table1). These three lines were selected based on the field studies at three locations giving treatments to water stress. The line 15 gave poor yields and 10 and 20 gave highest yield under severe drought condition. The seeds from these plants were collected by PI, Dr. Harbans L Bhardwaj at VSU-ARS, Petersburg, VA. These seeds were planted in pots in a greenhouse to collect fresh leaf tissue for RNA isolation. The plants at 4 - 6 trifoliate leaf stage were given water stress for seven complete days. The fully expanded young leaves were collected on seventh day and total RNA was isolated using Qiazen's DNeasy plant Mini kit (Qiazen, Inc) on the same day. The RNA quality was tested using UV-visible spectrophotometry and electrophoresing the samples on 1.1% formaldehyde agarose gel to confirm the quality of RNA.

2.2 The cDNA preparation and 454 sequencing

The cDNA was prepared and sequenced using the Roche 454 Life Sciences Genome Sequencer FLX titanium by following the manufacturer's instructions (454 Life Science, Roche). Libraries were prepared according to the 454 protocol: nebulization, purification, and ligation of adaptors. The libraries were prepared with ~1 μg (cDNA fragments) using the "Titanium General Library Preparation Kit". Sequencing on the Genome Sequencer FLX was performed using the "GS FLX Sequencing Kit Titanium Reagents XLR70 (Research and Testing laboratory, Lubback, Texas). NGEN was utilized by RT lab for assembly annotation and functional assignments were performed using Kraken (www.krakenblast.com) and the SwissProt/Tremble GOA curated database and cross references.

2.3 Gene Ontology (GO) annotation

After high quality sequence reads and aligned ESTs were obtained from Research and Testing laboratory, Lubbock, Texas, the bioinformatic analysis was performed using Blast2Go tool (Conesa, 2005). The Blast2Go tool was also used to assign Go ids, enzyme commission numbers and mapping to Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The E value cut off was set to 1e-5 and P-value to $</= 0.05$. The distribution of genes in each ontology categories was examined and the percentages of unique sequences in each of the assigned Go terms were computed (Jonas, 2007) and assigned to three main categories of GO, namely, biological process, molecular function, and cellular component.

2.4 Bioinformatic mining of microsatellites

The ESTs with more than 50 base pairs were searched for microsatellites using SSR finder software (www.SSRIT.com). In this work, SSRs were considered for primer design that fitted the following criteria: a minimum pattern length of 12 bp, excluding polyA and polyT repeat, at least 7 repeat units in case of di-nucleotide and at least 5 repeat units for tri-, tetra-, penta- and hexa-nucleotide SSRs. The default settings used for Primer3 input were optimum temperature of 63°C and an optimum primer size of 24 bases. All the ESTs were also analyzed for Blastx and Blastn similarities to Fabaceae database and results were compiled with a cut off e value of 1e-5 because protein sequences are better conserved evolutionarily than nucleotide sequences.

2.5 Identification of polymorphic SSR markers

A total of 648 primers were designed from 635 EST sequences and used for screening 6 genotypes 10, 15, 16, 18, 20 and 25 using Applied Biosystems (ABI) 3730xl DNA Analyzer. The 5 μl reaction mix has the following reagents in it

5 μl reaction:

1.0 μl DNA (@ 10ng/ul)

0.8 μl FAM labeled primer (IDT 16854154)

0.5 µl 10X Titanium Taq PCR Buffer (BD Biosciences, S1793)

0.04 µl Titanium Taq DNA Polymerase (BD Biosciences, S1792)

0.04 µl dNTP mix (25mM, Promega)

0.8 µl SSR specific primer

1.82 µl water

The PCR protocol has 10 steps and are Step1: 95°C for 3min, Step2: 95°C for 1min, Step3: 60°C for 1min, Step 4: GOTO Step 2 for1 time,   Step 5: 95 ^{o}C for 30 sec, Step 6: 60°C for 30sec, Step 7: 68°C for 30sec, Step 8: GOTO Step 5 for 26 times, Step9: 68°C for 4min, Step 10: 4°C forever. The resulting chromatograms were analysed using Genemapper (version 4.0) software (ABI, Life Technologies Corporation, California) and the amplified fragments displaying distinct sizes (peaks) were considered as different alleles. The allelic and genotypic frequencies for each locus were calculated using the TFPGA software (Miller, 1997). The number of alleles for locus, expected heterozygosity and observed heterozygosity were estimated by UPIC software (Renee et al., 2009). A measure of allelic diversity at a given locus (PIC) was calculated for all the loci (Wang et al., 2008).

3. Results and Discussion

3.1 Analysis of ESTs and their putative identities

The cDNA libraries from three genotypes generated a total of 54, 165 (Table1) ESTs. There are 44 EST sequences with less than 50 base pair in length not included for SSR as well as Blast analysis from three libraries (LI_10, LI_15 and LI_20). The average sequence length of the high quality ESTs range from 890-946 in three libraries. The length of the shortest sequence is 53bp (LI_15) and the longest sequence is 7496 (LI_10). When the libraries were blasted one on the other, there were 10,965 sequences from LI_15 hitting on sequences in LI_10 and 12,818 sequences from LI_10 were hitting on sequences in LI_15. There were 5643 sequences unique to genotype #10 not showing hits to LI_15 and 5179 unique to genotype #15 without any similar hits in LI_10.

The blast analysis against Fabaceae database revealed that there are 5260, 9882 and 16,714 sequences unique to tepary bean from libraries 10, 15 and 20 respectively showing no functional similarities to the known genes in Fabaceae. About 99.85% of the hits were showing significant similarities (eValue>= -05with Bit Score >50) to Fabaceae database. The top hit species distribution and direct GO count for known genes were presented in figures 1 to 6. The crop species identified with majority of hits to tepary bean ESTs is peas, cowpea, soyabean and alfalfa confirming the evolutionary relationship due to the conserved genes among the Fabaceae members.

It revealed highest homology with the sequences from *Glycine max, Vitis vinifera, Ricinus communis, Populus trichocarpa, Medicago truncatuala* and other crop species (Figure 2, 4 and 6). The mapping and annotation steps resulted in GO IDs (Gene Ontology Identifiers) to individual EST data sets of three libraries. A total of 13,263 (71.6%) of 18,523 ESTs were annotated with non-redundant (nr) GO IDs in the resistant line 10 while 6308 (40%) of 16190 ESTs in susceptible line 15 and 2738 (14.1 %) of 19, 452 ESTs were annotated with nr GO IDs. The remaining ESTs with unknown functions can be used in near future to identify the homologies that were conserved in related species and to identify various seed traits that could be used for grain quality improvement.

3.2 Functional classification of ESTs based on GO IDs and Fold enrichment scores

The ESTs were classified into subsets based on molecular function (level 2) and is represented as a pie diagram for each library (Figure 7, 8, 9). A major subset of ESTs (83%) across all libraries was linked to binding and catalytic activity, where the remaining groups involved in transporter (5%), transcription regulator activity (4%), molecular transducer activity (2%) and structural molecular activity (2%), electron career activity (1.8%), enzyme regulator activity (1%) and antioxidant activities (0.4%). In the Biological Process category, a major proportion of the genes were involved in cellular metabolic processes across all libraries and were involved mainly in protein , energy and organic acid metabolisms. ESTs were also grouped into different cellular components such as ribo-nucleoprotein complex, organelle part, cytoplasm and intracellular part and macromolecular complex based on fold enrichment (≥ 1). However, significant proportions of the expressed genes ($\geq 10\%$) were stress responsive and were conserved..

3.3 Genes involved in water stress tolerance

The putative genes related to various cellular and metabolic functions identified in tepary bean transcriptome were listed in Table 2. The number of stress responsive genes identified were abundant in genotype 10 compared to 15 and 20. The ESTs showing functional hits to proteins that probably function in stress tolerance (functional proteins) are water channel, transporters LEA proteins, chaperones and key enzymes for osmolyte biosynthesis (proline, sugars) and the second group contains protein factors involved in further regulation of signal transduction and gene expression that probably function in stress response (regulatory proteins) like transcription factors, protein kinases, phospholipid metabolism and ABA biosynthesis.

From the EST data sets,over expression of genes (115 GO terms) was clearly observed in LI_10 compared to those in LI_15. Using the Blast2Go tool, statistical analysis related to over expression of functional categories was performed based on Fisher Exact test. There were 1442 stress responsive genes significantly over expressed in genotype 10 compared to 15 (GO: 0006950) while 2407 to response to stimulus (GO: 0050896). The over expression of genes was also observed (286 Go terms) when combined files of LI_10 and LI_20 compared against LI_15.

The metabolic genes identified in the transcriptome are mostly involved in amino acid and

fatty acid metabolism (Table 2). The Cytochrome P450, Anthranilate, lipoxynase were highly expressed among the total number of metabolism related genes identified in line 10 compared to the other lines. They were identified as drought inducible genes (Bray, 2002). The cinnamoyl CoA reductase (CCR) activity, is over expressed in library 10 and 20 compared to 15 (GO:0016621), a key lignin gene, affect physical properties of the secondary cell wall such as stiffness and strength (Bala et al., 2005). The stress tolerance in genotype 10 could be due to pubescent leaf and stem surface, more leaves per plant and strong stem.

There were only few genes identified for energy metabolism in response to water stress (Table 2) to recognize the decrease in photosynthetic carbon and energy metabolism under water stress (Bray, 2002). The expression of oxygen evolving complex (GO: 0009654) was almost same level in all the genotypes under water stress. The photosynthesis genes, (GO: 0009768; GO: 0019684; GO: 0031978) light harvesting Photosystem I (light reaction), peroxidase (GO: 0004601) and oxido-reductase (GO: 0016491) activity were over expressed in genotype 10 and 20 compared to 15.

Transcription factors induced in response drought stress in tepary bean were WRKY, ethylene response, 14-3-3 like protein and MIP-PIP sub family (Table 2). A stress induced histone was also identified which helps maintaining chromatin structure during water deficit (Bray, 2002, Scippa et al., 2000). Aquaporins believed to play an important role in water homeostasis. Apparently the aquaporins are well represented from genotype 10 over the other two (Table 2). In higher plants aquaporin-like proteins, also called major intrinsic proteins (MIPs), are divided into five subfamilies. The MIP-PIP subfamily was over expressed in genotype 10 and was supported by previous reports in Arabidopsis (Alexandersson, 2010). The plasma membrane intrinsic protein (PIP) subfamily generally upregulated in leaves upon drought stress. The overall drought regulation of PIPs was a general and fundamental physiological response and was well conserved in genotype 10. These are similar to stress inducible transcription factors documented in Arabidopsis (Bray, 2002).

The genes participating in receptor signaling protein activity (GO: 0005057), receptor signaling protein serine/threonine kinase activity (GO: 0004702), red or far-red light signaling pathway (GO: 0010017) were over expressed in genotype 10 compared to 15. This could be due to enhanced ROS production under water stress. The ROS signaling is likely to be involved both upstream and downstream of the ABA-dependent signaling pathways under drought stress (Maria, 2008).

Water stress may cause degradation or malformation of proteins which lead to new protein or new genes induced. The mRNAs that code for ubiquitin, proline, leucine rich, zinc finger, early response to drought (erd) and heat shock proteins (hsp) were abundant in genotype 10 compared to 15 and 20 under water stress (Table 2). It was supplemented with the previous reports of rise in ubiquitin under stress (Caplan et al., 1989). Heat will denature the nascent polypeptides and cause the increased production of hsps involved in the correct protein folding (Caplan et al., 1989). It was revealed that drought tolerance of tepary bean genotype 10 could be due to biochemical processes related to proline metabolic enzymes beacausean increase in proline oxidase activity was observed under drought stress (Camacho, 1998)

compared to common bean. It was also supported by the electrophoretic patterns of matured seeds showing more proteins in tepary beans than in common beans with limited irrigation (Camacho, 1998). An abundant production of hydroxyl proline rich glycoproteins (HPRG's) which strengthen the cell wall was observed in genotype 10 in response to drought and was supported by previous reports (Showalter and Varner, 1989).

The gene expression to defense response is low (GO: 0050832) in genotype 15 during water stress compared to 10 (124 ESTs). The genes coding for chitinase and thioredoxin were more in genotype 10 compared to other two genotypes (Table 2)

Though 20 was a resistant line as per field observations, the gene expression data is not comparable with the resistant line under water stress conditions. Therefore based on the current gene expression studies, it was concluded that the genotype 10 is more tolerant to water stress compared to 20. It was revealed that the deep root penetration and sensitive stomata could cope with severe water stress in tepary bean genotype 10 and was supported (Mohamed, 2002) in two drought tolerant tepary bean (*Phaseolus acutifolius*) lines (NE # 8A and 19) for the responses to drought due to their deep root system.

3.4 Microsatellite containing ESTs

There were 1187 microsatellites found in 2329 ESTs containing SSRs out of 11,523 unique ESTs of LI_10 . The major types were 990 tri-nucleotide followed by 125 di, 36 tetra, 14 penta and 21 hexa nucleotide repeats. A total of 1087 microsatellites were found out of 2012 ESTs containing SSRs found in 16190 Unique ESTs of LI_15. The majority were Tri (901) followed by di (116), tetra (39), pent (11), hexa (14), hepta (4) and 24 (1) nucleotide repeats. There are 635 EST sequences containing SSRs in LI_20 with 648 micorsatellites. There are about a total of 3397 repeats containg di (93), tri (371), tetra (108), penta (10), hexa (49), hepta (1), octa (1), nano (8), deca 91) and dodeca (2) nucleotide repeats which were used for designing these 648 primers. A total of 2922 primer pairs were designed from three libraries for future microsatellite based mapping studies.

3.5. Identification of polymorphic EST-SSR markers

Out of 648 primer pairs tested, 104 primer pairs were showed no amplification on six genotypes selected for drought stress studies. All the amplified markers were polymorphic and 3866 alleles were found.

4. Conclusion

Our results documented the informative transcriptome for functional genes involved in drought stress tolerance and cellular metabolism. The information related to the effect of water stress on enzymes involved in individual metabolic pathways is lacking, might be due to lack of publicly available data for the genes involved, during the time of blast analysis. Further, bioinformatics screening of the generated ESTs against current database would help to find some more genes coding for enzymes involved in nutritional as well as structural, functional and metabolic pathways, and could be a value to the future crop improvement efforts at VSU's Agricultural Research Station. Being a nutritionally valuable crop, a

successful engineering of physiological and metabolic pathways for a number of genes involved or induced under water stress will resolve the problems of farming community in dryland areas. Therefore, future studies with continued funding could be directed towards 1) in depth functional analysis of genomic data generated for the tepary bean to date, 2) testing the genotypes selected for various agronomic, nutritional and anti-nutritional traits of interest utilizing the crop specific and cross species SSR markers developed to identify nutritionally valuable water stress tolerant genotypes for cultivation.

4.1 Accession numbers

The ESTs generated from this study have been deposited in Gen Bank and were assigned with accession numbers from HO774968-HO793490 (Library 10) , HO793491-HO809680 (Library 15) with db EST _ Id 71386542 to 71421254 and from HO 224465 to HO 243917 with db EST _ Id 70648090 for library 20.

List of abbreviations

BLAST: Basic Local Alignment Search Tool; bp: base pairs; EST: expressed sequence tag; GO: gene ontology; cDNA: Complementary DNA. LI: Library; SSR: Simple Sequence Repeat

Contribution of authors

Harbans Bhardwaj selected the genotypes for drought stress based on field experimentation and supplied all the materials required to complete the experiments. Brian E Scheffler provided the facilities as needed. Satya Narina raised the plants in green house, isolated the RNAs, performed the bioinformatic analysis of the ESTs, analyzed the results and drafted the manuscript. Sheron Simpson technically assisted in genotyping analysis. Linda Ballard performed the blast analysis and reviewed the manuscript. Ramesh Buyyarapu, Rao, K K., Harbans L Bhardwaj and Brian E Scheffler reviewed the manuscript. All the authors read and accepted for final submission.

Acknowledgements

References

Alexandersson, E., Danielson, J.A., Råde, J., Moparthi, V.K., Fontes, M., Kjellbom, P., & Johanson, U. (2010). Transcriptional regulation of aquaporins in accessions of Arabidopsis in response to drought stress, *61*(4), 650-60.

Bala, R.T., Maureen, F. N., Robert, E., & Gavin, F. M. (2005). Polymorphisms in Cinnamoyl CoA Reductase (CCR) Are Associated With Variation in Microfibril Angle in Eucalyptus spp.

Genetics Society of America. http://dx.doi.org/ 10.1534/genetics.105.042028.

Bhardwaj, H. L., Rangappa, M., & Hamama, A. A. (2002).Planting date and genotype effects on teparybean productivity Hort Science, *37*(2), 317-318

Bhardwaj, H. L., & Hamama, A. A. (2004). Protein and mineral composition of tepary bean seed Hort Science, *39*(6), 1363-1365

Bhardwaj, H. L. (2013). Preliminary evaluation of tepary bean (*Phaseolus acutifoilus* A Gray) as a forage crop. *Journal of Agricultural Science, 5*(7), 160-66. http://dx.doi.org/10.5539/jas.v5n7p160

Blair, M. W., Caldas, G. V., Munoz, C., & Bett, K. (2010). Evaluation of Condensed Tannins in Teparybean genotypes, 130-131. http://naldc.nal.usda.gov/download/IND44060984/PDF

Blum, A. (2011).Chapter 2: Plant water relations, plant stress and plant production. Book: Plant breeding for water limited environments. Vol 8:1-52, pages 255. ISBN 978-1-4419-7490-7/ http:www.springer.com/978-1-4419-7490-7.DOI 10.1007/978-1-4419-7491-4_2, © Springer Science+Business Media, LLC 2011

Bray, E. A. (2002). Classification of genes differentially expressed during water-deficit stress in Arabidopsis thaliana: an analysis using microarray and differential expression data Annals of Botany, 89:803-811.

Camacho, B. M., & González de Mejía, E. (1998). Comparative study of enzymes related to proline metabolism in tepary bean (*Phaseolus acutifolius*) and common bean (*Phaseolus vulgaris*) under drought and irrigated conditions, and various urea concentrations. Plant Foods Hum Nutr, *52*(2), 119-32.

Caplan, A Claes,B., Dekeyser, R., & Van Montagu,M.,(1989). Book Chapter: Salinity and drought stress in rice. Edited by Sanwan, R.S., Sangwan-Norreel, B.S. (1989) The Impact of Biotechnology in Agriculture. (Ed) Current Plant Science and Biotechnology in Agriculture, Kluwer Academic Publishers, PO Box 17,3300 AA Dordrecht, The Netherlands.P391

Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., & Robles M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics, *21*(18):3674-3676.

Cortés, A. J., Monserrate, F. A., Ramírez-Villegas, J., Madriñán, S., & Blair, M. W. (2013. Drought Tolerance in Wild Plant Populations: The Case of Common Beans (*Phaseolus vulgaris* L.). PLoS ONE 8(5): e62898. http://dx.doi.org/10.1371/journal.pone.0062898

Gupta, P. K., & Varshney, R. K. (2000). The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica, *113,* 163-185.

Hamama, A. A., & Bhardwaj, H. L. (2002). Tepary bean: A short duration summer crop in Virginia. In: J. Janick and A. Whipkey (eds.), Trends in new crops and new uses. ASHS Press, Alexandria, VA., 429–431.

Jonas, Q., Wang, S., Ping, L., Abernathy, J., Peatman, E., Wang, Y., Wang, L., Shi, Y., Wallace, R., Guo, X., & Zhanjiang, L. (2007). Generation and analysis of ESTs from the eastern oyster, Crassostrea virginica Gmelin and identification of microsatellite and SNP markers. BMC Genomics, *8,* 157.

Maria Helena Cruz de Carvalho (2008). Drought stress and reactive oxygen species Production, scavenging and signaling Plant Signal Behav, *3*(3), 156–165.

Mohamed, F. M., Norbert, K., Azza, A. T., & Georg, N. (2002). Dehydration-avoidance responses of tepary bean lines differing in drought resistance *Journal of plant physiology, 159*(1), 31-38.

Mohamed, M. F., Keutgen, N., Tawfika, A. A., & Noga, G. (2005). Comparative drought postponing and tolerance potentials of two tepary bean lines in relation to seed yield. *African Crop Science Journal, 13*(1), 49-60.

Montero-Tavera, V., Ruiz-Medrano, R., & Xoconostle-Cázares, B. (2008). Systemic nature of drought-tolerance in common bean. *Plant Signaling & Behavior, 3*(9), 663–666.

Narina, S. S., Xu,Y., Hamama, A. A., Phatak, S. C., & Bhardwaj, H. L.(2012) Effect of cultivar and planting time on resistant starch accumulation in pigeonpea grown in Virginia. ISRN Agronomy, vol. 2012, Article ID 576471, 4 pages, 2012. http://dx.doi.org/10.5402/2012/576471

Narina, S. S. S., Bhardwaj, H., Hamama, A. A., Burke, J. J., Phatak, S. C., & Xu, Y. (2014). Seed Protein and Starch Qualities of Drought Tolerant Pigeonpea and Native Tepary Beans. Journal of Agricultural Science, [S.l.], *6*(11), 247. ISSN 1916-9760

Narina, S. S. S., Bhardwaj, H. L., Burke, J. J., & Parris, J. W. (2015). Anti-nutrients and bioactive compounds in drought responsive native teparybeans. *IJAPES, 5*(1), 62-67.

Renee, S. A., Linda, B., & Brian, E. S. (2009). UPIC: how many SSR markers to run Perl Script to identify informative markers. Genomics and Bioinformatics Unit, USDA-ARS-MSA. American Society of Horticultural Science meeting, Saint Loius, IA, July, 2009.

Ricardo, A., Antonio, F., Paolo, V., & Maarten, J. C. (2006). Drought, Abscisic Acid and Transpiration Rate Effects on the Regulation of PIP Aquaporin Gene Expression and Abundance in Phaseolus vulgaris Plants. Ann Bot. December, *98*(6), 1301–1310.

Scheerens, J. C., Tinsley, A. M., Abbas, I. R., Weber, C. W., & Berry, J. W. (1983). Desert plants. The Nutritional Significance of tepary bean consumption. University of Arizona Agricultural Experiemnt Station paper 3719, *5*(1), 65-56
http://arizona.openrepository.com/arizona/bitstream/10150/552174/1/dp_05_01-011-056.pdf

Showalter, A. M., & Varner, J. E. (1989). Plant hydroxyproline-rich glycoproteins. In A Marcus, ed, The Biochemistry of Plants: A Comprehensive Treatise, Vol 15. Academic Press, New York, 485–520.

SSR finder software (2003). Steven Schroeder: University of Missouri-Columbia, Maize Mapping Project 2003-01-01
http://www.maizemap.org/bioinformatics/SSRFINDER/README.ssrfinder

Varshney, R. K., Mohana, S. M., Gaura, P. M.,Gangarao, N. V. P. R., ect.(2013). Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics Biotechnology advances, *31*(8), 1120-1134

Urrea, C. A., & Porch, T. G. (2009). Phenotypic evaluation of a subset of the *Phaseolus vulgaris* core collections and the *P. acutifolius* germplasm collection, and advanced common bean lines for drought tolerance in Nebraska. *Bean Improvement Cooperative Annual Report, 52,* 104-105.
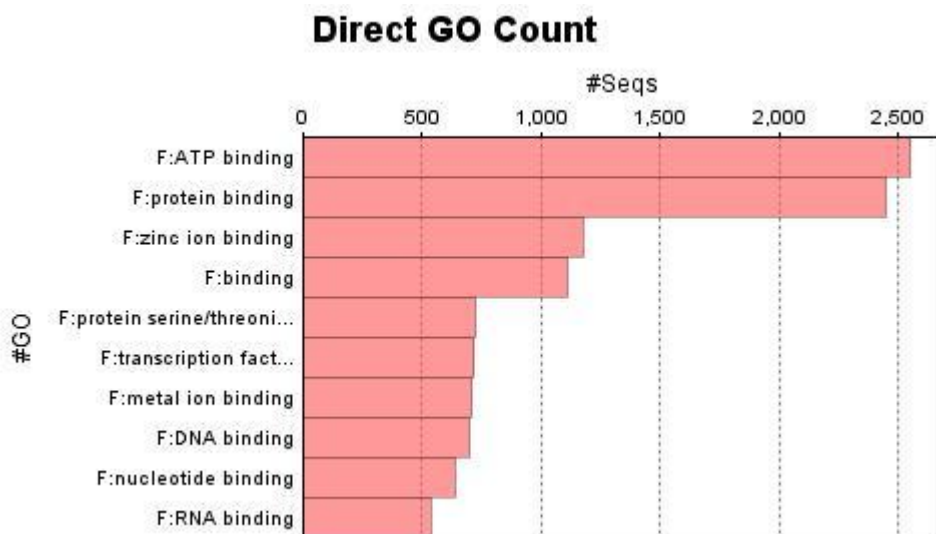
Figure1. Direct GO count for Putative ESTs from library TB10
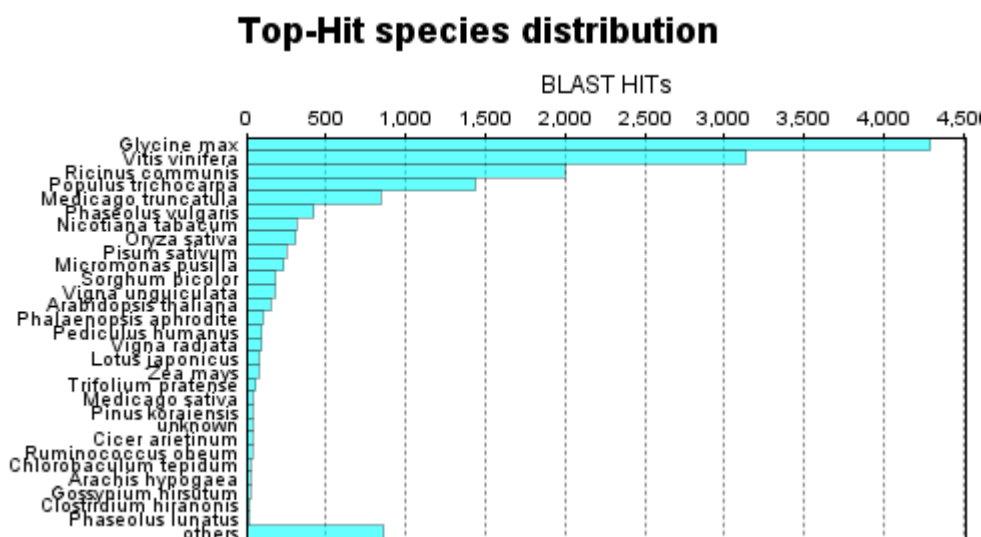
Figure 2: Top blast hit homologies identified from library TB 10
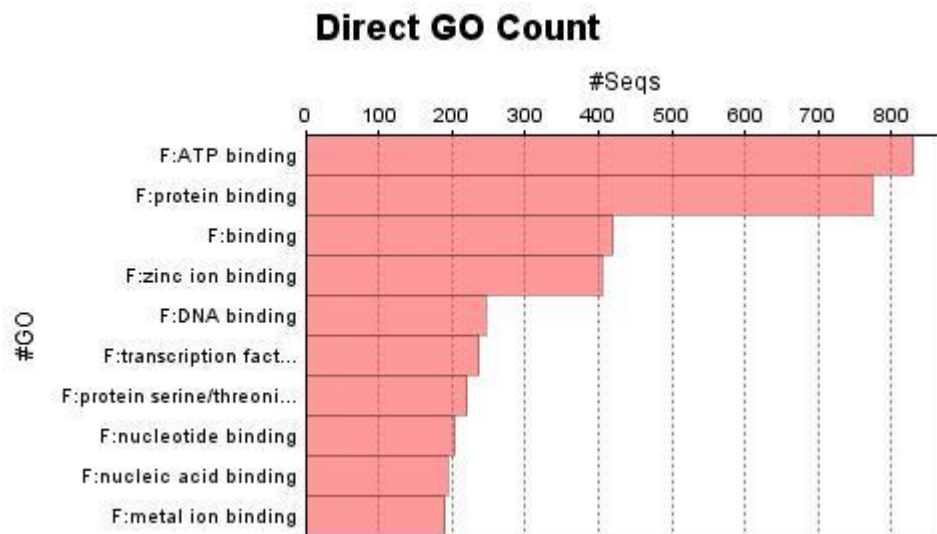
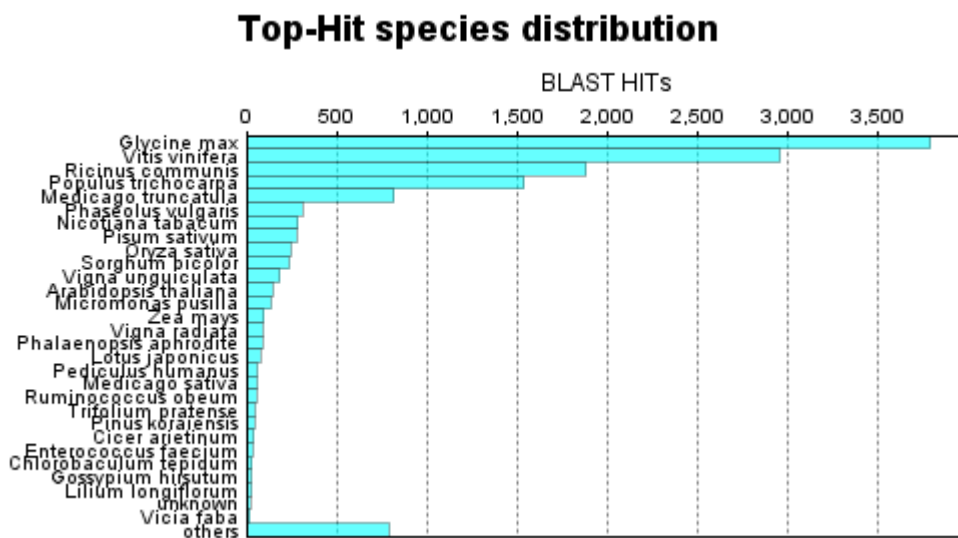**Figure 3**: Direct GO count for Putative ESTs from library TB15

Figure 4: Top blast hit homologies identified from library TB 15
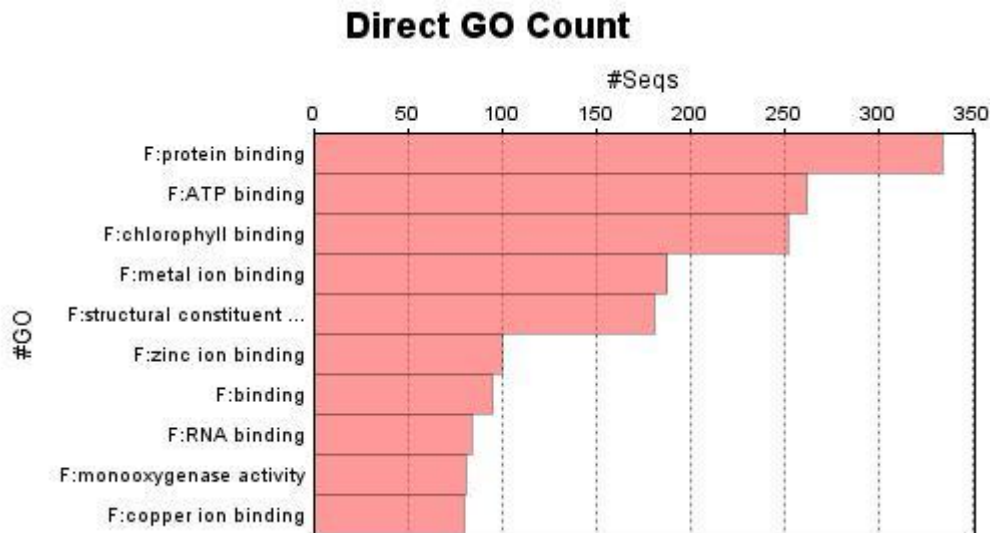
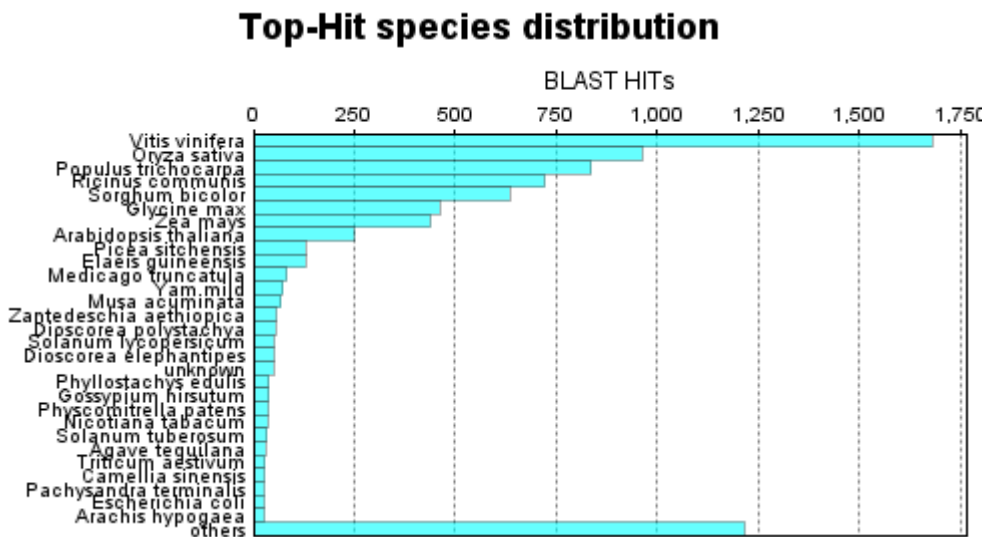**Figure 5**: Direct GO count for Putative ESTs from library TB20

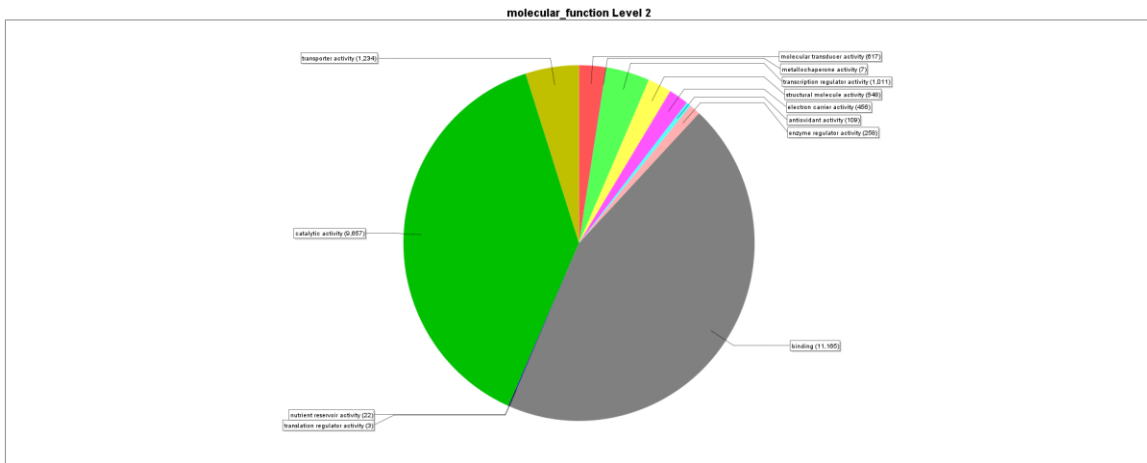**Figure 6**: Top blast hit homologies identified from library TB 20

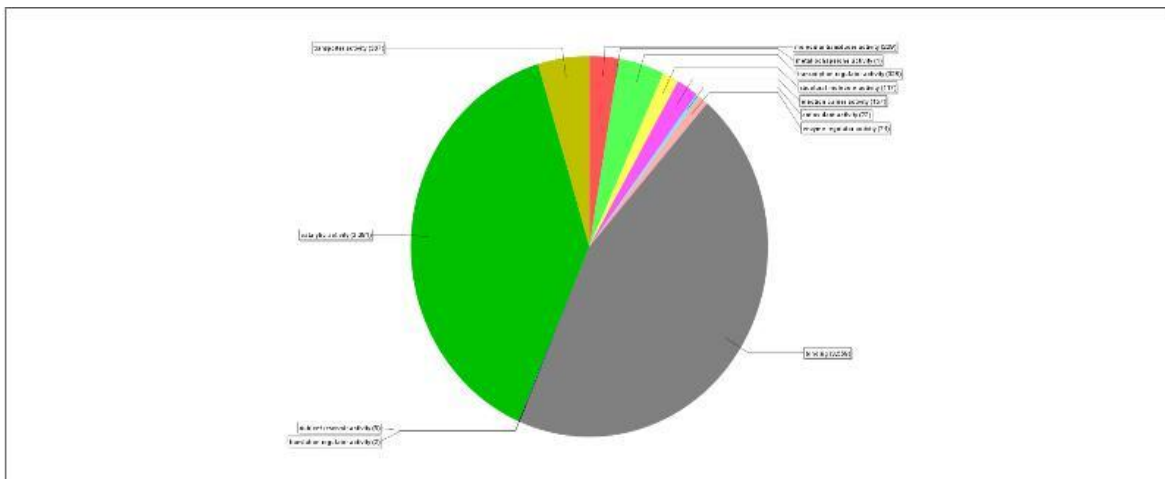**Figure 7**: GO distribution in TB10 based on molecular function.

Figure 8: GO distribution in TB15 based on molecular function.
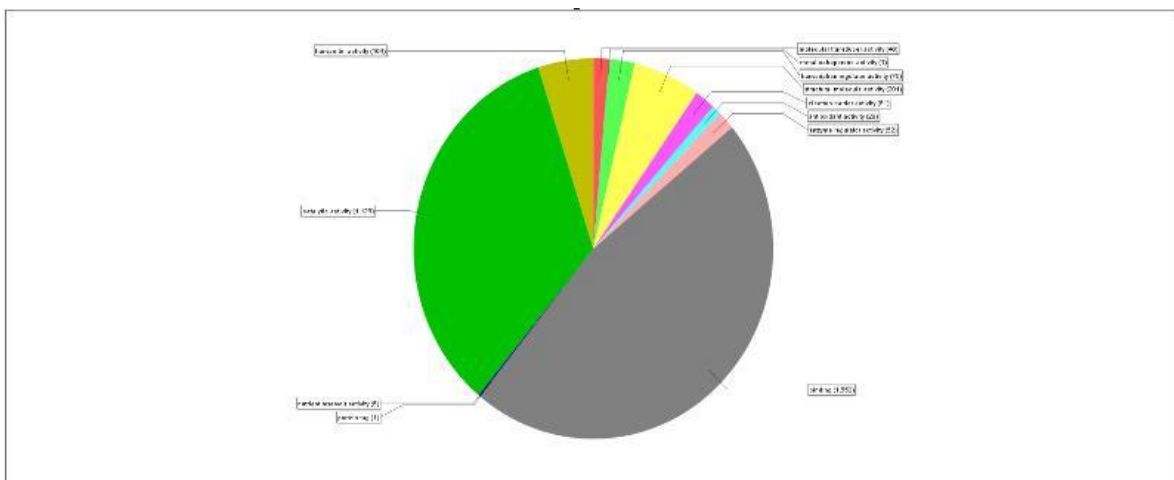

Figure 9: GO distribution in TB20 based on molecular function.

Table 1. Summary on genotypes, number of ESTs and SSR markers generated in tepary bean.

Genotype	Special characters	Total number of high quality ESTs (>50bp)	Number of trinucleotide repeats found	Number of ESTs containing SSR sequences
TB10	Resistant to Drought, profusely branching type vine with more number of broader trifoliate leaves at each node with less intermodal space and early germinating variety with green stem and deep root system.	18523	990	2329
TB15	Susceptible to Drought, slow growing type vine with few leaves and late germinating variety with red stem	16190	901	2012
TB20	Resistant to Drought, slow growing type vine with few narrow leaves with greater internodal distance and late germinating variety with read stem	19,452	371	635

Table 2. List of putative genes identified from cDNA libraries generated in three tepary bean genotypes.

Putative water stress responsive genes	Drought resistant tepary bean lines		Drought susceptible tepary bean line
	Library10	Library 20	Library 15
Metabolism			
Cytochrome P450	131	8	40
Hydroxynitrile lyase	1	0	0
Enolase	3	4	2
Amino acid biosynthesis and degradation			
Anthranilate	11	4	6
Tryptophan synthase (beta-subunit)	3	0	0
Lactoylglutathione lyase	5	1	2
Chorismate mutase	0	2	0
Aromatic metabolism			
4-Coumarate ligase	5	1	5
Cinnamyl alcohol dehydrogenase	11	0	1
Chalcone synthase	1	2	2
Cinnamoyl reductase	5	4	4
Phenylalanine ammonia lyase	2	0	1
0-methyltransferase protein	8	3	2
Fatty acid multifunctional protein			
Hydroperoxide lyase	2	0	0
Epoxide Hydrolase	4	1	3
Omega-3 fatty acid desaturase	7	0	1
Lipoxygenase	52	2	1
Allene oxide synthase/cyclase	6	0	1
Energy			
Oxygen evolving complex	4	5	3
Cell Growth			
Nitrilase	1	1	0
Transcription			
Ethylene response	1	0	0
14-3-3 protein	11	4	1
WRKY Transcription factor	10	2	6
mip pip subfamily	12	0	3
Protein synthesis/destination			
Proline rich protein like/proline transporter	18	7	7
protein kinase (calcium dependent)	207(12)	27(0)	130(5)
Ubiquitin	107	27	49
Homeobox -leucine rich/ zipper protein	62	7	43

Zink Finger Protein	110	19	57
Heat Shock Protein (hsp)	94	19	36
Early response to drought/dehydration (erd)	5	0	3
Transport			
Aquaporin	4	3	1
Transporter (Sugar)	229(8)	29(1)	117(6)
Cell rescue/defence			
<u>Oxidative stress</u>			
Glutathione-S-transferase	14	8	1
Glutathione peroxidase	6	3	2
Catalase	10	1	1
Thioredoxin	32	8	13
Peroxiredoxin	6	2	0
<u>Pathogenesis related</u>			
Chitinase	37	7	1
beta Glucanase	7	3	4
Metallothionein	7	3	6
Putative lectin	5	8	6

Copyright Disclaimer