International Journal of Agricultural Sciences and Natural Resources
2015; 2(4): 113-134
Published online August30, 2015 (http://www.aascit.org/journal/ijasnr)
ISSN: 2375-3773

AASCIT American Association for Science and Technology

Keywords
ESTs,
Legumes,
Abiotic,
Biotic,
GC Profile,
ka and ks

Received: July 27, 2015
Revised: August 6, 2015
Accepted: August 7, 2015

Comprehensive EST Based Analysis of Differentially Expressed Stress Genes/Factors in Comparative Mode Among Four Legumes

Sabeen Fatma[1, *], Jitender Singh[2], D.V. Rai[1], Mohd Uruj Jaleel[3]

[1]Department of Bioinformatics, Faculty of Biological Engineering, Shobhit University, Meerut, India
[2]College of Biotechnology, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut, India
[3]School of CS & IT, Singhania University, Jhunjhunu, Rajasthan, India

Email address
sabeenfatima16@gmail.com (S. Fatma)

Citation
Sabeen Fatma, Jitender Singh, D.V. Rai, Mohd Uruj Jaleel. Comprehensive EST Based Analysis of Differentially Expressed Stress Genes/Factors in Comparative Mode Among Four Legumes. *International Journal of Agricultural Sciences and Natural Resources*. Vol. 2, No. 4, 2015, pp. 113-134.

Abstract
Background: Legume is one of the world's most important crop, it is consumed worldwide and play a significant role in the economy of producing countries. Legume crop productivity is severely affected due to various stress factors both biotic and abiotic(e.g. drought, salinity, cold, ethylene temperature, kinase cascade) these stress factors both biotic and abiotic significantly impact the production of crop in response to these, plants activates a number of defense mechanism that function to increase tolerance in adverse conditions. Therefore tracking the candidate genes responsible for stress tolerance through sequence similarity and functional studies is becoming increasing important for breeding and will act as useful resource for comparative genomics and can be further used as molecular markers or for genetic transformation to develop desired. Thus, research initiatives have been launch to produce genomic and transcriptomic data about legumes species(*Cajanus cajan*, *Cicer arietinum*, *Pisum sativum* and *Lens culinaris*). *Results:* Assembling the ESTs of *Cajanus cajan*, *Cicer arietinum*, *Pisum sativum*, *Lens culinaris* and by applying various bioinformatics tools, the detection of GC content through GC profile, prediction of ka/ks values using MEGA 5.0, identified legume genes under purifying and positive selection and phylogenetic relationship among various stress factors(biotic and abiotic) in four legume species. On these result we emphasize gene annotations and discuss various stress factors based on the categories defined. *Conclusion:* Identifying and mining genes involved in stress response represent a key step to unraveling and manipulating stress tolerance in legumes. Comparative analysis among the legumes within the same species and between species will enable us to identify species specific genes underlying stress response. Despite knowing that comparisons between these legumes species data should be carefully inspected, our initiative established possible transcriptome elements that could guide the legumes specific community in unraveling the molecular mechanism that distinguish these four extremely important legumes species. In addition, the annotation of legumes-specific/stress prominent genes adds new element to genomic initiatives that our searching for traits (factors) that could differentiate legume each species from other. We believe that such data are a valuable aid to the interpretation of legume development,

providing insight that could help in legumes reading program and indicating potentialtargets for functional analysis and biotechnology products of such socially and economically important legume species.

1. Background

Legumes are valuable agricultural and commercial crops that serve as important nutrient sources for human diet and animal feed. The legume family is wide and incredibly diverse (Doyle and Luckow, 2003). It is the vital class of plant that provides protein in diet for significant proportion of human population. As for other protein-rich legume crop species, rapid genetic improvement of these species is needed to meet the increasing demand for protein food sources in the world. Legume constitute an important component of human diet in developing nations,the crop under study includes*Cicer arietinum*, *Pisum sativum*and *Lens culinaris* Mof Hologalegina class (Cool Season Legume) and also include *Cajanus cajan* of Phaseolid class(Warm Season Legume). Of all the above include legume species, despite having an important role in food security majority of these legumes crop demonstrate low productivity due to various stress factors both biotic(eg bacteria. Fungi nematodes viruses& insects) and abiotic(eg drought, salinity, cold, extreme temperature, kinases cascades, accumulation of hormones such as ethylene(ET, Abscissic acid). These stress factors both abiotic & biotic significantly impact the production of crop in response to these, plants activates a number of defense mechanism that function to increase tolerance to adverse condition. Therefore, tracking the candidate genes responsible for stress tolerance through sequence similarity and functional studies is becoming increasingly important for marker-assisted breeding. Such candidate genes serve as useful resources for comparative genomics and can be further used as molecular markers or for genetic transformation to develop desired cultivars. Recent advancement in the technology has allowed development of various genetic tools and genomic approaches to identify genomic regions and genes/quantitative trait loci (QTLs) underlying plant stress response in many crop species (Varshney *et al*., 2005). Expressed Sequence Tags (ESTs) are sequenced portions of messenger RNA and provide a source for the discovery of new genes for comparative analyses between organisms. Many ESTs sequencing efforts have successfully provided insight into crop plant development (Marques *et al.,* 2009). EST sequencing allows qualitative expression analyses by correlated EST frequency with desirable traits of plant species. It also constitutes an interesting tool for detection of tissues/stress specific promoters and genetic variation that may account for specific characteristics. Furthermore, EST analyses can provide targets for transgenes is, an interesting tool for genetic improvement of such a crop. In recent years, EST projects have been initiated for numerous plant and animal species, and have generated a vast amount of sequence information that can be used for gene discovery, functional genetic studies, andmarker development (Pashley*et al.,* 2006).The availability of EST data as on dated(01/11/2014) prompted us to perform a wide bioinformatics analysis.

2. Material and Methods

ESTs sequences of *Cajanus cajan, Cicer arietinum, Pisum satvium, Lens culinaris*were retrieved from NCBI home pagewww.ncbi.nlm.nih.gov. EGassembler(http://www.genome.jp/tools/egassembler/) is applied for contigs and singleton sequences generation in order to make searches for stress responsive factors (Ali Masoudi-Nejad *et al.,* 2006).

2.1. Sequence Analysis

Standalone BLAST package, ncbi-blast-2.2.25+ from NCBI was used for homologs search analysis. All stress related genes were compared against all four legume's CDS using BlastN program. Further, homology hits were filtered on criteria on >70% identity and E value cut-off ≤1E-50 using in-house perl script. Finally genes were selected by using a criteria of presence of homologs in at least three out of four legumes species (Table:2).

All selected stress factors were subjected for ka, ks value and ka/ks value prediction by MEGA 5.0 using Neighbor-joining, Maximum likelihood and maximum parisomy method with 1000 replicates in Boot step test(Tamura *et at.,*2011). All position containing gap and missing sequences were not calculated. Sequence distances, ka and ks were calculated between all gene pair in each stress related gene by MEGA 5, sequences distances which indicate the extend of similarity between homologs (including orthologs and paralogs) was calculated by number of base substitution per site.

2.2. Function Annotation Through Fast annotator

The potential function of gene transcript, fast annotator site was applied, fast annotator pipeline to provide automatic annotation of nucleotide sequences via a web interface. The output file of fast annotator include the best hit in the NCBI non redundant database, GO term, EC number and Domain identity (Punta *et al*., 2012; Pruitt *et al*.,2007). Fast annotator allows users to assign protein functions, cellular location, enzyme activity and function domain to query sequence through an easy to use interface (Chen *et al*., 2012).

2.3. GC Profile

To evaluate the GC content of the stress responsive factors (genes), GC profile was applied (http://tubic.tju.edu.cn/GC-Profile/). GC profile, an interactive web based software system available as a public resource dedicated to analyze the compositional heterogeneity of DNA sequences (Gao and Zhang *et al.* 2006).

3. Result and Discussion

ESTs sequences available at NCBI EST database were used for identification of stress responsive genes. A total of ~25,578 *Cajanus cajan*, ~46,441 *Cicer arietinum*, ~21,838 *Pisum sativum* and ~10,341 *Lens culinaris* M were downloaded from NCBI database. These ESTs sequences show high redundancy. Using EGassembler software, these ESTs were assembled into 1402 contigs, 14,889 singletons of *Cajanus cajan*, 4451 contigs, 11,672 singletons of *Cicer arietinum*, 3451 contigs, 7794 singletons of *Pisum sativum* and 1071 contigs, 3818 singletones of *Lens culinaris* M(Fig:1).The average length of *Cajanus cajan*, *Cicer arietinum*, *Pisum sativum* and *Lens culinaris* were predicted through Geneious software (Table:1).These contigs were then used for similarity search against known candidate genes for stress responsive factor (abiotic & biotic) using BLASTN. After manually blast, 31 stress responsive factors (Genes) were selected for further analysis (Table:2)

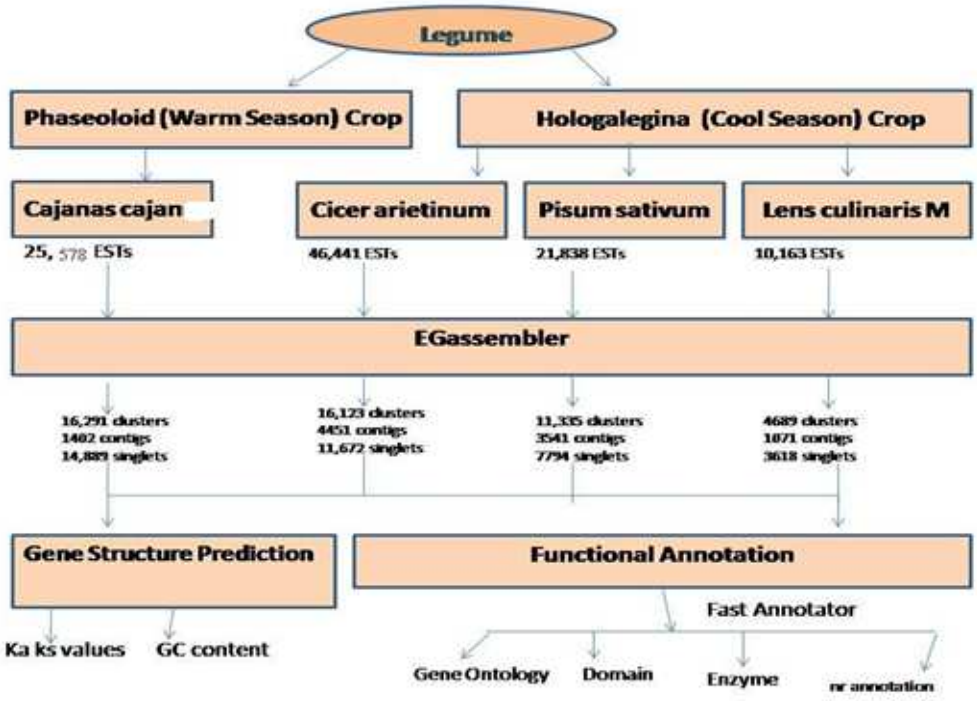


Fig. 1. *Flow Diagram of bioinformatics procedure applied in four legumes species.*

3.1. Prediction of ka and ks Values and Phylogenetic Relationship

Sequence distances, ka aswell as ks values were calculated for all possible 31 stress responsive genes in legume under the present study surveyed. It is well known that ka is smaller than ks in natural evolution because of conservation of functional coding genes therefore non-synonymous change was less frequent in mutation of nucleotides during evolution(Hurst, 2002;Nekrutenko *et al.,* 2002) for protein coding sequences, the synonymous rate ks is often regarded as a major of underline mutation rate(Miyata *et al.,*1980), through it may be influenced by other factors (Williams and Hurst, 2002).

To gain insight into the molecular and phylogenetic evolution of stress responsive genes in the four legumes species analyzed, estimated the rate of synonymous(ks, silent mutation) and non-synonymous(ka, amino acid altering mutation) substitution, generated by MEGA 5 analysis and performed the ka/ks test for positive selection of each gene. ka/ks is a good indicator of selective pressure at the sequence level. Theoretically, a ka/ks>1 indicates the rate of evolution is higher than neutral rate. Conversely a gene with ka/ks<1 has a rate of evolution less than neutral rate (Yang and Bielawski, 2000).

The estimated ka and ks for these stress responsive genes as the coalescence time for different stress responsive gene in this linage is quite similar, differences in relative rates between proteins should reflect in selection or mutations rates than time elapsed. The correlation between these ka and ks values is estimated (Fig:2).

The ka/ks analysis allowed the detection of genes with low ka/ks ratio, such as those in coding protein H3, H2, Lipoxygenase, Stress induced protein, Aquaporin. The majority of these proteins have been shown to be highly conserved and to suffer strong positive selection (Roth and Liberles, 2006). The highest ka/ks includefactor/gene Glyceraldehyde-3 phosphate Dehydrogenase, Glycine dehydrogenase, AP2/ERBP and Dehydrin.These result are in accordance with previous reports which shows that genes acting in response to stress are often positively selected for diversification due to competition with the evolving effectors proteins of pathogens(Roth and Liberles, 2006; Stukenbrock

et al.,2009). This study once again underlines the importance and significance of the stress responsive genes for legumes species. Because of this reason, these genes have remained more conserved in speciation and rearrangement during the evolution of legumes species.

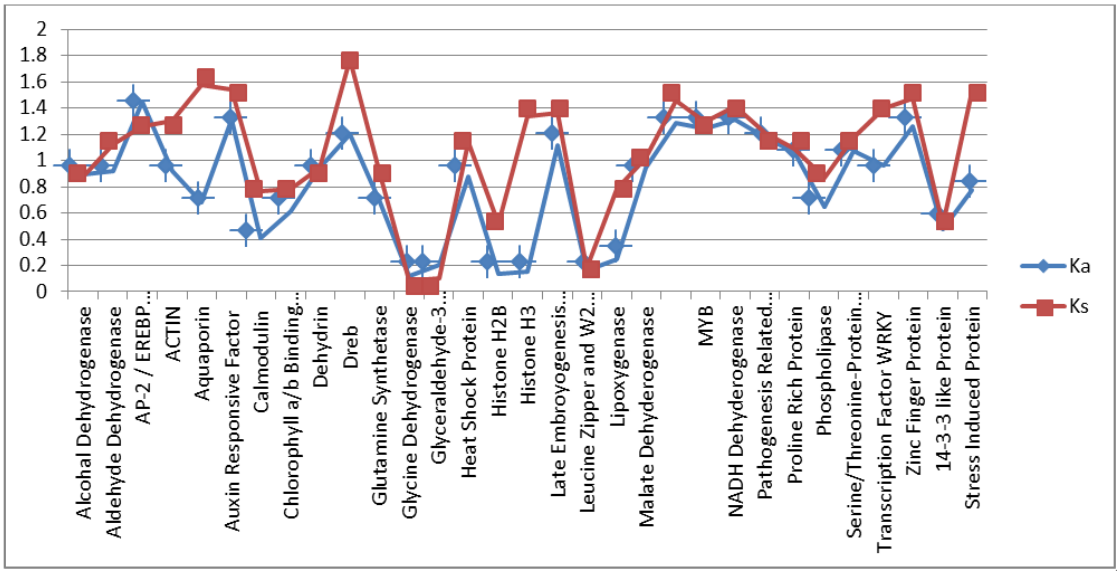


Fig. 2. *The ka and ks values of 31 Stress Responsive Factors Using MEGA 5.0.*

3.2. Evolution of GC Content

Inorder to understand the evolution structure and function of the genes under study, it was important to know the GC content of the sequences. The GC content of all the 31 stress responsive factors were 44.03 in *Cajanus cajan*, 41.21 in *Cicer arietinum*, 42.30 in *Pisum sativum* and 42.05 in *Lens culinaris* as shown graphically in Fig:3

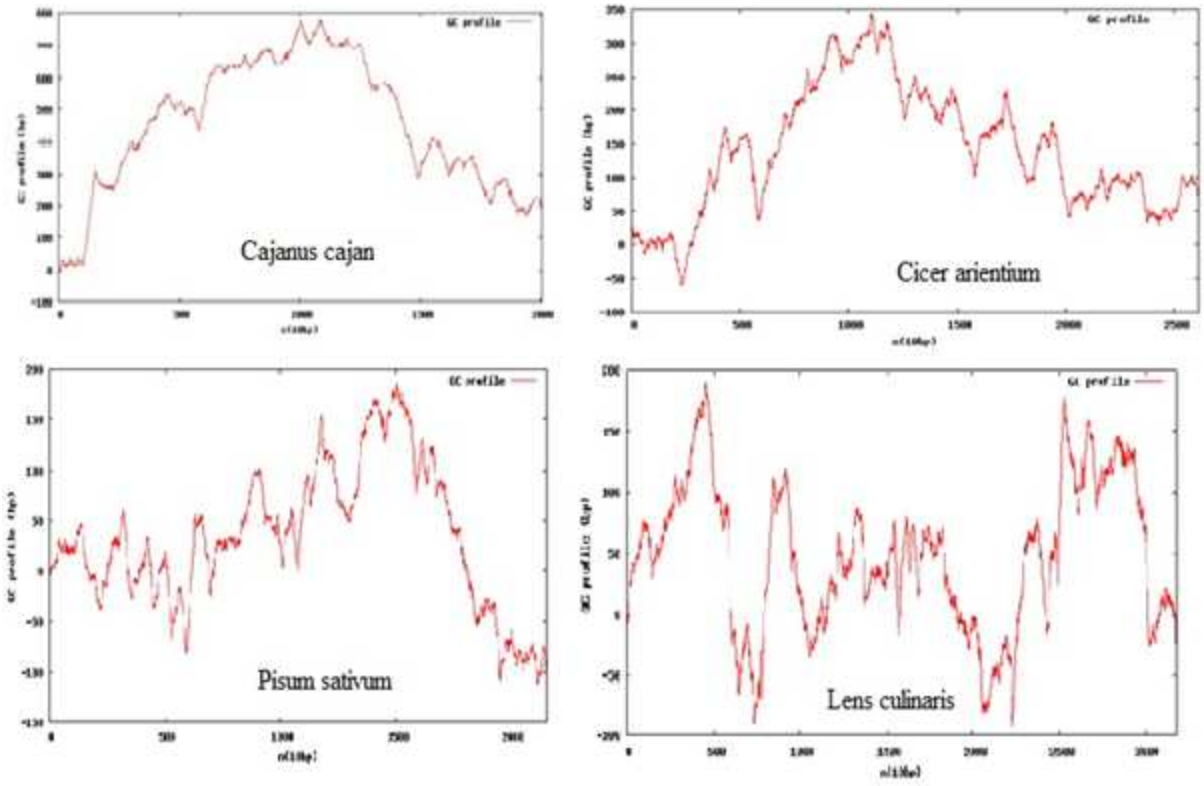
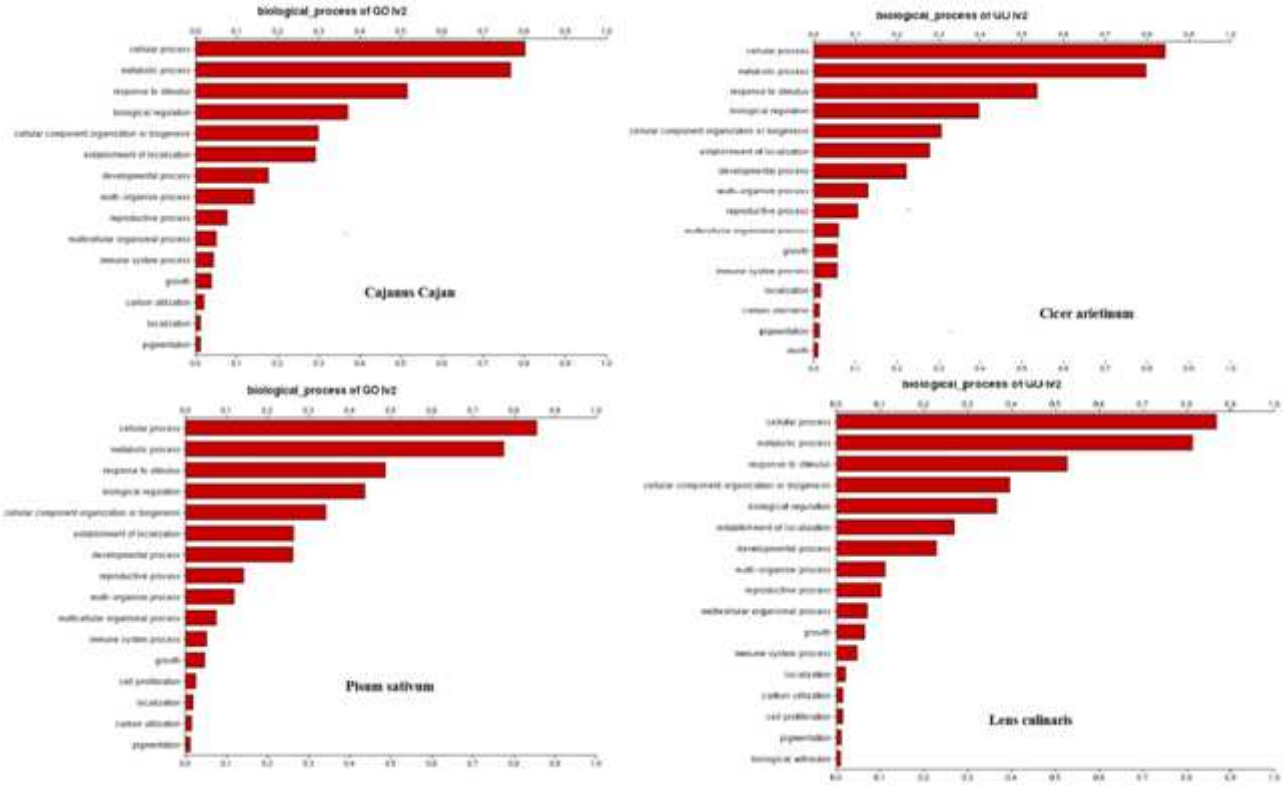
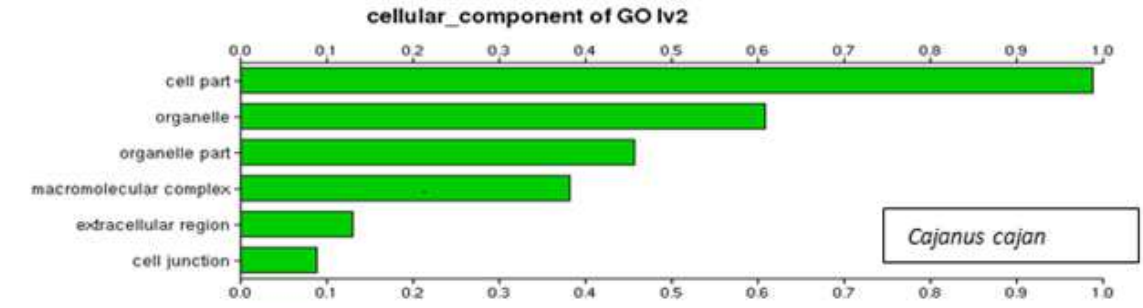


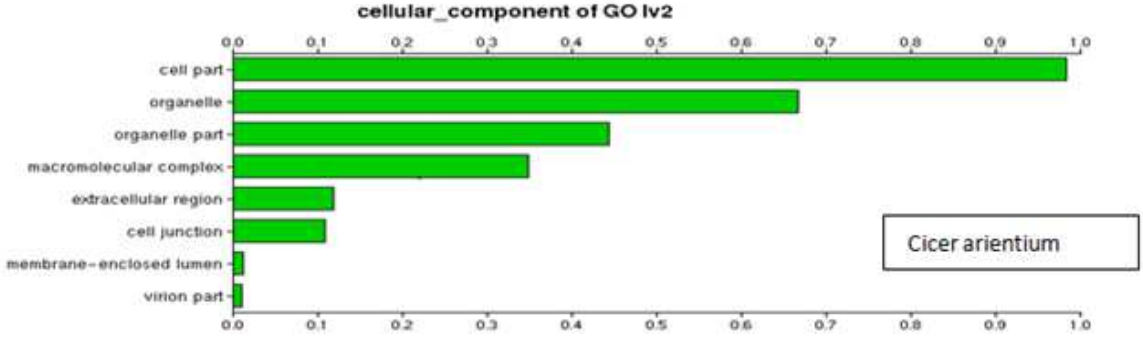
Fig. 3. *The GC Content of All Four Legumes Cajanus cajan, Cicer arientium, Pisum sativum and Lens culinaris Using GC Profile.*

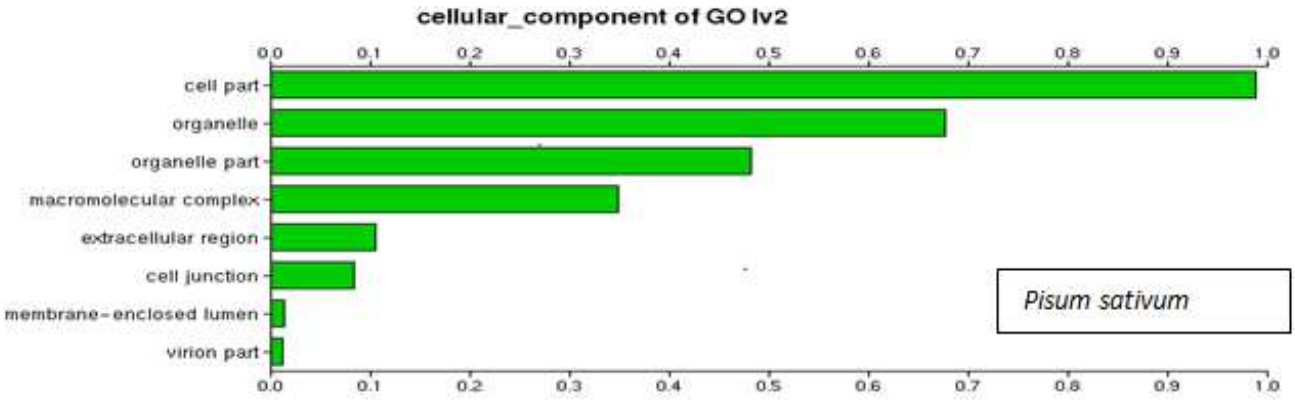
3.3. GO Annotation

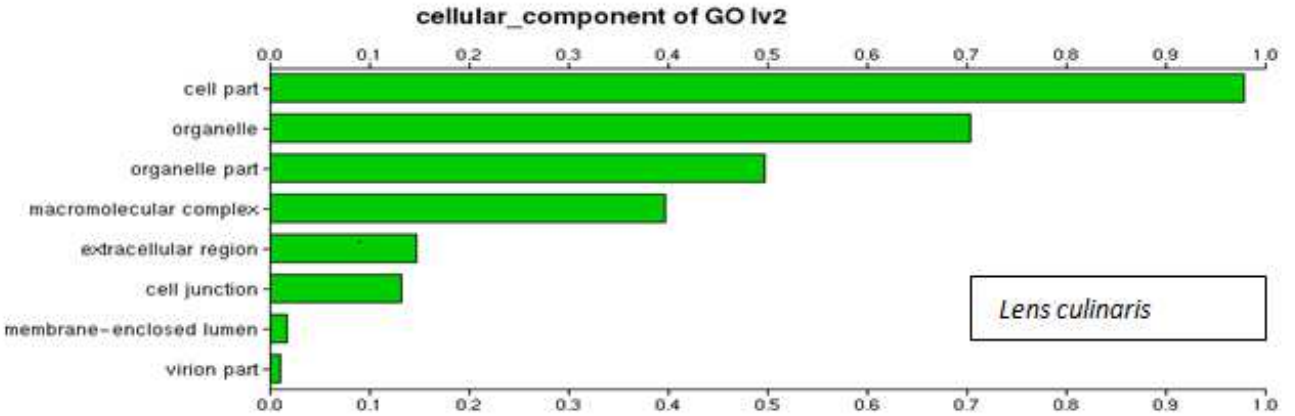
GO annotation is the most commonly used and well established functional annotation scheme. Here in GO annotation results are grouped according to biological processes, cellular component and molecular functions are plotted in three separated horizontal bar chart as shown in Fig:4

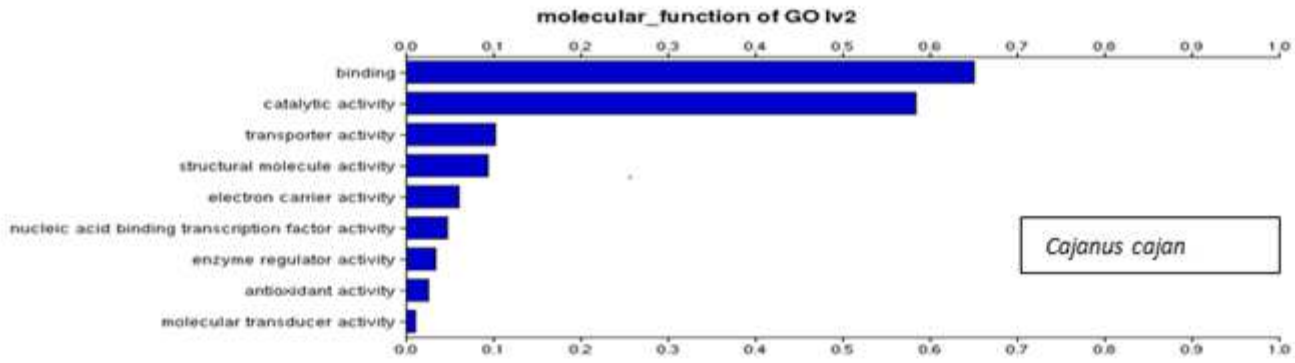


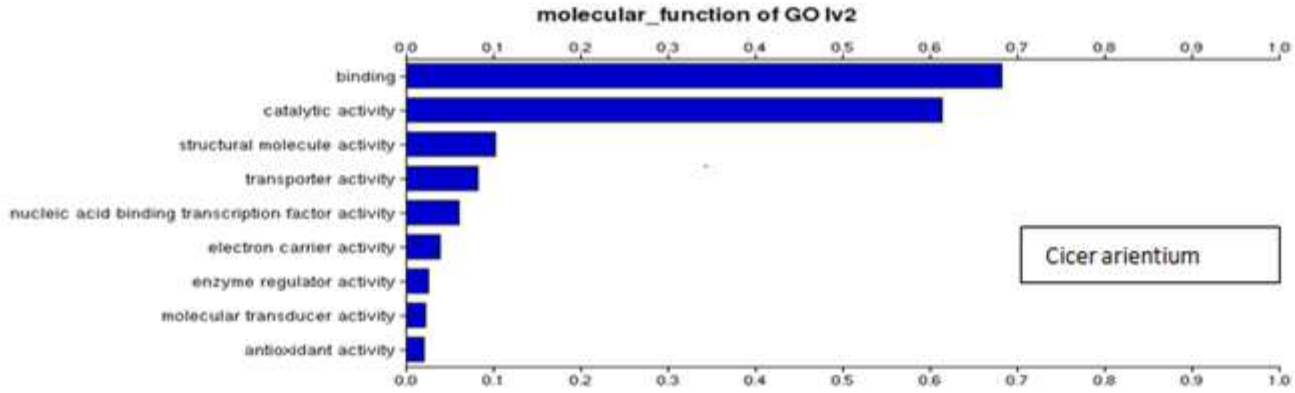





cellular_component of GO lv2
cell part
organelle
organelle part
macromolecular complex
extracellular region
cell junction
membrane-enclosed lumen
virion part
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Pisum sativum


cellular_component of GO lv2
cell part
organelle
organelle part
macromolecular complex
extracellular region
cell junction
membrane-enclosed lumen
virion part
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Lens culinaris


molecular_function of GO lv2
binding
catalytic activity
transporter activity
structural molecule activity
electron carrier activity
nucleic acid binding transcription factor activity
enzyme regulator activity
antioxidant activity
molecular transducer activity
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Cajanus cajan


molecular_function of GO lv2
binding
catalytic activity
structural molecule activity
transporter activity
nucleic acid binding transcription factor activity
electron carrier activity
enzyme regulator activity
molecular transducer activity
antioxidant activity
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Cicer arientium

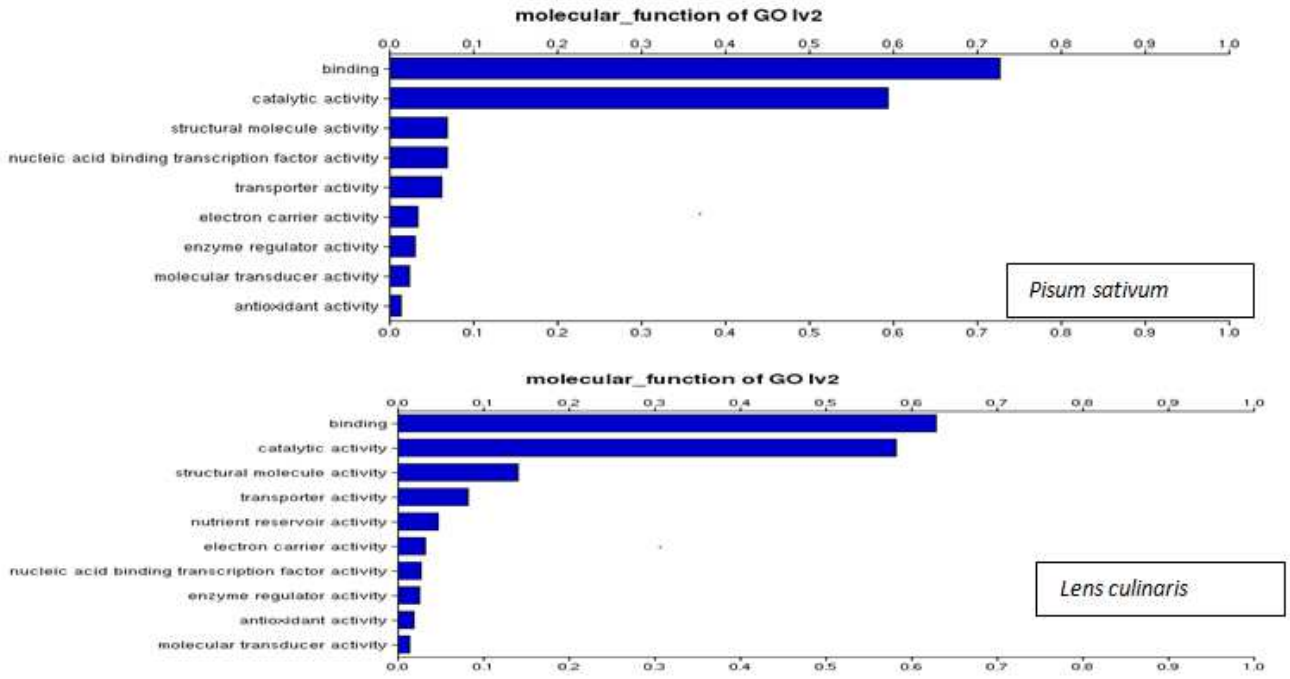


Fig. 4. *GO Annotation Results are Grouped According to Biological Processes, Cellular Component and Molecular Functions.*

Table 1. *Summary of four legumes species through Geneious Software.*

Legumes	Contigs	Average Contig Length	Singlets	Average Singlet Length	Clusters
Cajanus cajan	1402	597.9	14889	459.2	16291
Cicer arietinum	4451	685.6	11672	468.3	16123
Pisum sativum	3541	699.9	7794	492	11385
Lens culinaris M	1071	968.9	3618	691.0	4689

Table 2. *List of 31 Stress Responsive Genes Analyzed in Four Legume Species.*

S.No	Stress Responsive Factor	*Cajanus cajan*	*Cicer arietinum*	*Pisum sativum*	*Lens culinaris*
1.	Alcohal Dehydrogenase	AF079499.1 FJ914863.1 XM_003602082.1	NM_001255479.2	XM_003630356.1	XM_003602086.1 FJ914863.1 AF079499.1
2.	Aldehyde Dehydrogenase	NM_001249884.1	XM_003630762.1	XM_003595260.1	XM_003595260.1 NM_001249884.1
3.	AP-2 / EREB P transcription factor	NM_001148270.2	JN595892.1	XM_003608516.1	XM_003611426.1
4.	Actin	XM_003637460.1 NM_001249519.2	NM_001279094.1 JQ028731.1 X67666.1	NM_001289231.1	JQ028731.1 NM_001279094.1 X67666.1
5.	Aquaporin	NM_001254121.2 XM_003600815.1 KF770828.1	XM_003591968.1	XM_003624529.1 NM_001255806.1	XM_003629815.1 NM_001255076.2 AJ243308.1
6.	Auxin Responsive factor	XM_003624415.1	XM_003593821.1	XM_003612885.1	XM_003624415.1
7.	Calmodulin	NM_001250954.1 AF030034.1	XM_004502222.1	XM_003617094.1 NM_001250954.1	NM_001279096.1 NM_001250954.1

S.No	Stress Responsive Factor	*Cajanus cajan*	*Cicer arietinum*	*Pisum sativum*	*Lens culinaris*
8.	Chlorophyll a/b binding protein	NM_001279096.1		JX869966.1	AF030034.1
		NM_001279182.1	EF488077.1	X81962.1	EU041720.1
		EU041720.1	XM_003596442.1	XM_003637291.1	NM_001279182.1
		XM_003618434.1	NM_001253951.2		XM_003618434.1
9.	Dehydrin	NM_001253177.2	U91970.1	EU121850.1	Z14145.1
		U54703.1		KM652424.1	
10.	DREB	HE647689.1	FJ223566.1	----------	XM_003588416.1
11.	Glutamine Synthetase	XM_003600654.1	XM_003600654.1	XM_003592590.1	M20663.1
		L20248.1	M20663.1		XM_003600654.1
		M10159.1	AF301590.1		X04001.1
		AM238618.1			
12.	Glycine Dehydrogenase	XM_003588952.1	XM_003588952.1	XM_003588952.1	XM_003588952.1
13.	Glyceraldehyde-3 phosphate Dehydrogenase	NM_001253117.2	XM_003601780.1	XM_003608827.1	XM_003601780.1
		XM_003601780.1	NM_001253117.2	NM_001249200.2	NM_001253117.2
			L07500.1	L07500.1	
14.	Heat Shock Protein	NM_001249683.1	KJ808806.1	U08820.1	XM_003621821.1
		XM_003617828.1	XM_003621838.1	XM_003621821.1	KJ808806.1
15.	Histone H2B	XM_010111630.1	XM_003606579.1	XM_003606962.1	XM_003625462.1
16.	Histone H3	XM_003608487.1	XM_003608487.1	XM_003628643.1	XM_003592420.1
17.	Late Embroyogenesis Abundant	----------	FJ226587.1	XM_003590357.1	XM_003604675.1
18.	Leucine zipper and w2 Domain Containing Protein	XM_003624134.1	XM_003624134.1	XM_003624134.1	XM_003624134.1
		NM_001253130.1	NM_001253130.1	NM_001253130.1	NM_001253130.1
19.	Lipoxygenase	NM_001250409.1	AJ276265.1	XM_003627154.1	X07807.1
		XM_003627139.1	U84198.1	U84198.1	XM_003597511.1
			XM_003627154.1	S76063.1	NM_001248454.1
			GU942746.1		AF204210.2
20.	Malate Dehyderogenase	XM_003638117.1	NM_001249732.1	NM_001289254.1	XM_003589977.1
		X74507.1	NM_001279132.1		NM_001279132.1
			XM_003589977.1		NM_001256362.1
21.	Mitogen Activated Protein Kinase	XM_002306571.2	XM_003608053.1	XM_003594049.1	XM_003604087.1
			NM_001248497.2		
			X70703.1		
22.	MYB	AB108649.1	DQ822949.1	XM_003607152.1	XM_003548266.2
23.	NADH Dehyderogenase	AF451596.1	NM_001253992.2	XM_004505646.1	XM_003603962.1
		FJ695012.1			
		AF451594.1			
		AF451592.1			

S.No	Stress Responsive Factor	*Cajanus cajan*	*Cicer arietinum*	*Pisum sativum*	*Lens culinaris*
24.	Pathogenesis Related Protein	JX475048.1	JX475048.1	JX475045.1 AF137351.1	AF115574.1
25.	Proline Rich Protein	NM_001251664.1	AJ489610.1	XM_003601427.1 X67427.1	XM_003608694.1
26.	Phospholipase	XM_003597846.1	XM_003597846.1	Y15253.2 XM_003600732.1 NM_001248395.1	XM_003597846.1
27.	Serine/threonine-Protein Kinase	XM_006597089.1	XM_003624327.1	XM_003611756.1	XM_003602843.1
28.	Transcription Factor WRKY	EU019586.1	EU049488.1 EU019582.1	NM_001250742.1	----------
29.	Zinc Finger Protein	XM_003638479.1	XM_003591078.1	XM_003630801.1	XM_003625476.1
30.	14-3-3 like Protein	U15036.1 NM_001252761.1 XM_003615112.1	NM_001255222.2 XM_003629715.1 FJ225662.1	XM_003629715.1 NM_001248750.1	XM_003607752.1 NM_001251494.1
31.	Stress Induced Protein	NM_001249109.2	XM_003618154.1	XM_003591163.1	XM_003591493.1 NM_001253157.1

3.4. Functional Classification of Differentially Expressed Stress Responsive Genes and Prevalent Protein Families in Legumes

Based on the result of functional annotation and GO analysis, we established functional categories to various stress responsive factors (genes) and its consequences in development, dehydration and environmental adaptation.

3.4.1. Genes Involved in Regulatory Pathway (Cell Signaling and Transcription)

Genes involved in cell signaling and transcription processing are known to regulate many cellular responses in an organism. Calcium is an element that is crucial for numerous biological functions. The calmodulin (CAM) family is a major class of sensors protein which collectively play a crucial role in cellular signaling cascades through the regulations of numerous target proteins. CAM is one of the most conserved proteins in all eukaryotes (Benoit *et al*., 2006). The CAM gene have been found in *Cajanus cajan* (CcContig169), *Cicer arietinum* (CaContig3373), *Pisum sativum* (PsContig2187) and *Lens culinaris* (Lc Contig891) and ka/ks valuefor CAM is 0.533.The serine/threonine protein kinase are involved in disease response via a signaling cascade presumably conserved in plants (Zhang *et al*,. 2001) while somatic embryogenesis related kinases are associated with brassinosteroid signaling pathway that plays a key role in plant defense (Karlova *et al*., 2006).The ser/thr protein kinase have been found in *Cajanus cajan* (CcContig1092), *Cicer arietinum* (CaContig654), *Pisum sativum* (PsContig2310) and *Lens culinaris* (LcContig640) and predicted ka/ks value is 0.910.14-3-3 proteins a large family of protein found in all eukaryotes (Aitken, 2006 and Roberts, 2000). 14-3-3s plays a vital role in regulation of nitrate reductase, protein related to the plant growth and brassinolide signaling (Roberts, 2000) and the proton ATPase (Jahn *et al*., 1997). The interaction of these 14-3-3s with many proteins involved in K, N and S metabolism and show that 14-3-3 protein interaction alters the activities of target proteins in legumes. Thus it suggests a role for 14-3-3s in regulating multiple steps in N and S metabolism. The 14-3-3 protein have been found in *Cajanus cajan*(CcContig789), *Cicer arietinum*(CaContig193), *Pisum sativum*(PsContig2284) and *Lens culinaris*(LcContig51) and ka/ks value is 0.941.MAPs are involved in cell differentiation, division and stress response(Robinson and Cobb,1997) in plant a number of study have demonstrated that MAPKs signal both abiotic and biotic stresses, including cold and drought(Jonak *etal.*, 1996) wounding(Zhang and Klessig, 1998), hormoneaction (Ligterink and Hirt, 2001) and plant pathogen attack. The MAP kinase have been found in *Cajanus cajan* (CcContig432), *Cicer arietinum*(CaContig532), *Pisum sativum*(PsContig1630) and *Lens culinaris*(LcContig1067) and ka/ks value is 0.883.Actin is also involved in several basic developmental processes including the establishment of cell polarity, cell division, cell wall deposition and cell elongation(Mascarenhas,1993). These functions are of crucial significance to plant development as they play a pivotal role in Tyr kinase signaling, integrin signaling, TGFβ signaling, transforming growth factor and G-protein coupled reaction. The Actin have been found in *Cajanus cajan*(CcContig327),

Cicer arietinum(CaContig2044), *Pisum sativum*(PsContig3144) and *Lens culinaris*(LcContig10) and ka/ks value is 0.705.

3.4.2. Transcription Factors(TF)Acting in Stress

The families of transcription factors including MYB, AP2/EREBP, Leucine Zipper and W2 domain containing protein, Zinc finger and WRKY.

MYB an important TFs that participates in plant responses to a stress belongs to the large MYB family, which contain functionally diverse protein(Dubos*etal.,* 2010). MYB is transcriptionally induced by dehydration and this upregulation reversed upon rehydration (Urao*etal.,* 1993). The MYB have been found in *Cajanus cajan*(CcContig304), *Cicer arietinum*(CaContig2842), *Pisum sativum*(PsContig444) and *Lens culinaris*(LcContig257) and ka/ks value is 0.964.AP2/EREBP TF- A2(APETALA2) and EREBPs(ethylene-responsive element binding protein) are genes form a large number of multigene family and that play a variety of role throughout the plant life cycle. It plays a key regulator of several developmental processes, like floral organ identity determination or control of leaf epidermal cell identity, to forming part of mechanism used by biotic and environmental stress. On the basis of sequence similarity AP2/EREBP-TFhave been found in *Cajanus cajan*(CcContig864), *Cicer arietinum*(CaContig534), *Pisum sativum*(PsContig2064) and *Lens culinaris*(LcContig673) and ka/ks value is 1.150.Leucine Zipper and W2 domain containing protein play a pivotal role in the response to drought,in modulation of plant growth and response to environmental signals. These genes also involved in several abiotic stress responses, meristem regulation, photomorphogenesis and root development(Ariel*et al.,*2007). The Leucine Zipper and W2 domain containing proteinhave been found in *Cajanus cajan* (CcContig654), *Cicer arietinum*(CaContig2447), *Pisum sativum*(PsContig1902) and *Lens culinaris*(LcContig243) and ka/ks value is 0.915.Zinc finger proteins are required for key functionsincluding cellular processes, transcriptional regulation development,pathogen defense and stress responses(Ciftci-Yilmaz and Mittler,2008). The Zinc finger have been found in *Cajanus cajan*(CcContig980), *Cicer arietinum*(CaContig555), *Pisum sativum*(PsContig17) and *Lens culinaris*(LcContig1070) and ka/ks value is 0.855.WRKY TF- WRKY genes super family, one of the largest TF gene families has suggested to play important roles in the regulation of transcriptional programming associated with plant stress response(Rushton*et al,*.2010). WRKY protein function via interaction with a diverse array of a protein patners including MAP kinase, 14-3-3 protein, calmodulin, histone deacetylases, resistance proteins and other WRKY TF.The WRKY TF have been found in *Cajanus cajan*(CcContig610), *Cicer arietinum*(CaContig934), *Pisum sativum*(PsContig1426) and ka/ks value is 0.686. This data suggest that interplay of the broad spectrum of TF possibly regulates multiple signaling cascades during immune stress.

3.4.3.Gene Involved in Metabolism

Developmental changes and stress response are often correlated with or result in adjustment in various metabolic pathways. The genes present in this class represented several biochemical pathways such as carbohydrates, fatty acid energy and N_2 metabolism. The genes involved in Carbohydrate metabolism comprised ADH, GAPDH, Malate dehydrogenase, NADH were present.

ADH enzyme activity can be used as a bio indicator for water logging stress in legume plant and there by playing an important role in flooding stress tolerance and signaling(El-Enany *et al.,*2014). ADH activity is critical for the recycling of NADH and thus continuation of glycolytic pathway(Johnson *et al.,*1994). Plant cell have the physiological potential to trigger the mechanism under anoxia which maintains satisfactory energy loads under stress(Rawyler *et al.,* 2002). Therefore plant expose to the flooding normally are exposed to the fermentation process as a secondary route available in plant metabolism for energy production. ADH, NADH activities increases in flooding, is resulting in the induction of anaerobic metabolism in many higher plants. ADH and NADH have been found in Contig *Cajanus cajan*(CcContig836,CcContig697), *Cicer arietinum*(CaContig81, CaContig3416), *Pisum sativum*(PsContig430, PsContig1897) and *Lens culinaris*(LcContig2, LcContig605) and ka/ks value are 1.014 and 0.927.Glyceraldehyde-3 Phosphate Dehydrogenase(GAPDH)- plants must constantly respond to various environmental stresses during growth and development. The GAPDH genes that encodes the enzyme in the glycolytic pathway (Russell and Sachs,1989; Yang Y*et al.,*1993).The GAPDH have been found in *Cajanus cajan*(CcContig655), *Cicer arietinum*(CaContig665), *Pisum sativum*(PsContig750) and *Lens culinaris*(LcContig465) and ka/ks value is 1.980.Malate dehydrogenase(MDH) that catalysis the inter conversion of malate and oxaloacetate(OAA) coupled to the reduction or the oxidation of NAD pool. MDH also operate in three different pathway in plant(i)TCA (ii) Conversion of Gly to Ser, by reducing OAA to malate and proving a supply of NAD^+ for GLDH(Journet *et al.,*1981).(iii)In the photorespiration in C4 plant. Hence MDH plays a vital role in metabolism and in many pathways. The MDHgenehave been found in *Cajanus cajan*(CcContig937), *Cicer arietinum*(CaContig3128), *Pisum sativum*(PsContig529) and *Lens culinaris*(LcContig508) and ka/ks value is 0.922.Most of these enzymes are known to altered the expression in response to the stress(Laxalt*et al.,*1996 and Cushman*et al.,*1989) many of these enzymes are involved in maintaining the metabolite pool that may drive the metabolic processes to overcome such stresses.

(i). TheGenes Involved in Lipid Biosynthesis

In plants, oxidative stress is one of the major causes of damage as a result of various environmental stresses. Oxidative stress is thought to be one of the major causes of cellular damage and cell death(Bartel,2001).Aldehydes comprises a major portion of the lipid peroxidation products

and are toxic because of their chemical reactivity. Increase aldehyde dehydrogenase activities represent one defense strategy in the detoxification of aldehyde(Yoshida*et al.,* 1998). The over expression of ALDH improves stress tolerance most likely by scavenging toxic aldehydes and thus reducing lipid peroxidation. The ALDH have been found in *Cajanus cajan*(CcContig847), *Cicer arietinum*(CaContig3404), *Pisum sativum*(PsContig2732) and *Lens culinaris*(LcContig58) and ka/ks value is 0.814. The Activation of phospholipases is at the origin of the production of important defense signaling molecules such as oxylipins and jasmonate as well as the potent second messenger.It has found that phospholipase A and phospholipase C play an important role in the control of plant defense response to the attack of invading pathogen(Munnik*et al.,* 2009 and Laxalt *et al.,* 2002). The reports identification of phospholipases allows a better understanding of its function during plant pathogen interactions.The Phospholipase have been found in *Cajanus cajan*(CcContig549), *Cicer arietinum*(CaContig336), *Pisum sativum*(PsContig10) and *Lens culinaris*(LcContig4333) and ka/ks value is 0.750.

(ii). Amino Acid and Nitrogen Metabolism

Glutamine synthase(GS) is the key enzyme involved in the assimilation of NH_3 derived either from nitrate reduction, N_2 fixation, photorespiration, or asparagine breakdown. GS increases photosynthetic productivitiesunder N-deficiency(Fuentes*et al.,*2001). The GS have been found in *Cajanus cajan*(CcContig624), *Cicer arietinum*(CaContig2545), *Pisum sativum*(PsContig750) and *Lens culinaris*(LcContig465) and ka/ks value is 0.757. Proline rich protein-proline is a multifunction amino acid and plays a role in carbon and nitrogen metabolism, cell signaling, nutrient adaption, and protection against osmotic and oxidative stresses(Lehmann*et al.,*2010). The plant used increased proline content for biosynthesis of physiological specific proteins or stress proteins(proline rich protein) (Ashraf and Foolad,2007). Thereforerapid accumulation of free proline in plants is typical response to a wide range of environmental stresses(Pavlikova *et al.,*2008). The proline have been found in *Cajanus cajan*(CcContig550), *Cicer arietinum*(CaContig3402), *Pisum sativum*(PsContig163) and *Lens culinaris*(LcContig117) and ka/ks value is 0.972.Glycine dehydrogenasealso known as glycine decarboxylase(GLDC), is a multi protein complex, that is responsible for an important reaction of primary metabolism in all organism including plants. Glycine itself is a precursors for chlorophyll, glutathione, tryptophan, phosphatidyl choline and related phospholipid and ethanolamine. The role of GLDC in all the organism is to interconnect the metabolism of one, two and three carbon compounds(Douce*et al.,*2001). Therefore a malfunction of GLDC results in serious metabolic consequences. Plants are not able to perform oxygenic photosynthesis without GLDC and with reduced activity of these enzymes, will usually show severe growth retardation(Somerville,2001; Heineke *et al.,*2001). The GLDC have been found in *Cajanus cajan*(CcContig94), *Cicer arietinum*(CaContig168), *Pisum sativum*(PsContig3046) and *Lens culinaris*(LcContig74) and ka/ks value is 1.463.

3.4.4. Genes Related to Heat Stress

Plant must cope with heat stress for survival, so they develop different mechanism including the maintenance of cell membrane, stability, capturing the reactive oxygen species(ROS), synthesis of antioxidants, accumulation and osmoregulations of osmoticum, induction of some kinases that respond to stress, calcium dependent kinase proteins, and enhancing the transcription and signal transfer of chaperones(Wahid *et al.,*2007). The role of HSPs and stress induced protein in the folding of other proteins is important. (Morimoto and Santoro, 1998) indicated that HSPs protects cells from the injury and facilitate recovery and survival after a return to normal growth conditions. It has been found that the role of HSPs is to act as a molecular chaperones(Timperio*et al.,*2008). Regulating the folding and accumulation of protein as well as localization and degradation in all plants and animal species(Gupta*et al.,*2010). These proteins as chaperones, prevents the irreversible aggregation of other proteins and participates in refolding proteins during heat stress conditions(Tripp*et al.,*2009). HSP and stress induced proteins have been found in Contig *Cajanus cajan*(CcContig684,CcContig622), *Cicer arietinum*(CaContig69,CaContig750), *Pisum sativum*(PsContig1869,PsContig262) and *Lens culinaris*(LcContig426,LcContig805) and ka/ksvalue are 0.766 and 0.511.

3.4.5. Phytohormone

Auxin signaling is key to many plant growth and developmental processes from embryogenesis to senescence. The auxin pathway constitutes an essential component of plant biotic and abiotic stress tolerance mechanism. Auxin as an indicator of environmental signal in plant root development, emerging evidence also implicates auxin as an integral part of plants overall biotic and abiotic stress tolerance mechanism(Zhang *et al.,*2012; and Stirnberg *et al.,*2012). The Auxin have been found in *Cajanus cajan*(CcContig1046), *Cicer arietinum*(CaContig2021), *Pisum sativum*(PsContig404) and *Lens culinaris*(LcContig958) and ka/ks value is 0.851.

3.4.6. Defense and Stress Responsive Gene

Pathogen attack is often accompanied by the accumulation of elevated levels of transcript of diseases related protein, the PR gene. The involved of these proteins in plant defense responses is well known(Yasuda *et al.,*2001). PR are the heterogeneous group of plant proteins, inducible by biotic stresses(Van Loon *et al.,*2006; Sels *et al.,*2008). Some of these proteins are effectors against pathogens and insects, while other are involved in the reestablishing homeostasis after the stress(Van Loon *et al.,*2006). Many PRs are induced not only by pathogens but also by abiotic stresses, developmentally regulated are constituently expressed in some tissues(Van Loon *et al.,*2006). Many other proteins for

example universal stress protein, aquaporin, cold acclimation responsive protein were also identified, implicating their involvement in pathostress response despite their role in abiotic and other stresses.Aquaporin are membrane intrinsic proteins that faciliate and regulate the passive movement of water molecule down a water potential gradient (Maurel *et al.,*2008). Aquaporin seems to play a specifically important role in controlling transcellular water transport and hydraulic conductance in plant tissue(Javot and Maurel, 2002). Aquaporin also play a role in nitrogen metabolism. PR gene and Aquaporin have been found in Contig *Cajanus cajan*(CcContig830,CcContig256), *Cicer arietinum*(CaContig2090,CaContig1741), *Pisum sativum*(PsContig2873, PsContig709) and *Lens culinaris*(LcContig1024, LcContig988) and ka/ksvalue are 0.972 and 0.430.H2B has been reported to affect leaf and root growth and seed dormancy(Fleury *et al.,*2007). H2B is also involved in the regulating the expression of key flowering times and plant development(Xu *et al.,*2009; Schmitz*et al.,*2009). Defense responses in the plants have been demonstrated to required H2B for both positive and negative regulations(Devoto *et al.,*2003; Trujillo and Shirasu,2010). In Arabidopsis H2B is a regulatory component of plant defense against necro tropic fungal pathogen(Dhawan *et al.,*2009). These results suggest that H2B is involved in multiple developmental processes and responses to biotic stress in plants(Hu *et al.,*2014). H2B have been found in Contig *Cajanus cajan*(CcContig924), *Cicer arietinum*(CaContig2294),*Pisum sativum*(PsContig2173) and *Lens culinaris*(LcContig28) and ka/ks value are 0.259.Lipoxygenase(LOXs) LOXs are normally present in the seeds of plants(Siedow,1991). LOXs are almost ubiquitous and involved in various physiological processes. Therefore induction of the LOX genes during plant-pathogen interactions have been reported in several species and confirming, the function of LOXs in the defense against pest seems to be related to the synthesis of number of different compounds with signaling functions(Creelman and Mullet,1997). LOXs have been found in Contig *Cajanus cajan*(CcContig463), *Cicer arietinum*(CaContig2149),*Pisum sativum*(PsContig2444) and *Lens culinaris*(LcContig302) and ka/ks value are 0.312.

3.4.7. Genes Related to Abiotic Stress and Detoxification

Genes related to the abiotic stresses are potentially important in the recent scenario of harsh environmental changes such as the increase of extreme temperature and drought periods. This category includesLate Embroyogenesis Abundant(LEA), Dehydrin, *DREB*,Histone H3, Chlorophyll a/b binding protein(CAB).

LEA genes codes for a diverse group of proteins that accumulate to high levels in seed development(Baker*et al.,*1988). The expression of LEA genes can be induced by the application of absiccsic acid and by various abiotic stresses such as dehydration, osmoticstress, or cold in bothreproductive and vegetative tissues. LEA proteins have been detected in different stages of plant development in response to different water deficient conditions(salinity, freezing and drought). LEA have been found in *Cicer arietinum*(CaContig593), *Pisum sativum*(PsContig2980) and *Lens culinaris*(LcContig1) and ka/ks value is 0.819.Dehydrin which also belong to group of LEA proteins, are a family of instrincally unstructured plant proteins that accumulate during the late stages of embryogenesis, and play an important role in plant responses and adaptation to stress, dehydrin proteins are major components of response to several abiotic stress such as cold or drought in plant kingdom. Dehydrin have been found in Contig *Cajanus cajan*(CcContig469), *Cicer arietinum*(CaContig2990),*Pisum sativum*(PsContig3478) and *Lens culinaris*(LcContig357) and ka/ks value are 1.002.*DREB* proteins are important TFs in regulating abiotic stress related genes and play a crucial role in imparting stress tolerance to plant. Dreb genes are strongly and transiently induced by low temperature stresses(Thomashow,2002). The Dreb regulons can thus we used to improve the tolerance of various kinds of agriculturally important crop plant to drought high salinity and freezing stress by gene transfer(Charu and Prasad,2011).*DREB* have been found in Contig *Cajanus cajan*(CcContig825), *Cicer arietinum*(CaContig525), and*Lens culinaris*(LcContig804)and ka/ks value are 0.677.Histone H3 are the main protein component of chromatin they undergo extensive post translation modifications particularly acetylation, methylation, phosphorylation, ubiquition, ADP ribosylation which modify the structural,functional properties of chromatin.(Kim*et al.,*2008). H3 have been found in Contig *Cajanus cajan*(CcContig145), *Cicer arietinum*(CaContig1323),*Pisum sativum*(PsContig758) and *Lens culinaris*(LcContig440) and ka/ks value are 0.114.Chlorophyll a/b binding protein (CAB)-different stressful environment, includes salinity, drought and heat causes generally a considerable reduction in content of important photosynthetic pigments, chlorophyll a and chlorophyll b. The extend of reduction in the content depends on the species, variety, duration of plant exposure and tolerance of the stress. Reduction in the photosynthetic pigment whether through the impairment in pigment biosynthesis or destruction of pigment may lead to impairment in electron transport and hence reduced photosynthetic capacity in most plant(Ashraf and Harris,2013) with all these explanation, Chlorophyll a/b binding protein acting as a potential stress responsive factors. CAB have been found in Contig *Cajanus cajan*(CcContig314), *Cicer arietinum*(CaContig240),*Pisum sativum*(PsContig2360) and *Lens culinaris*(LcContig380) and ka/ks value are 0.784.

4. Conclusion

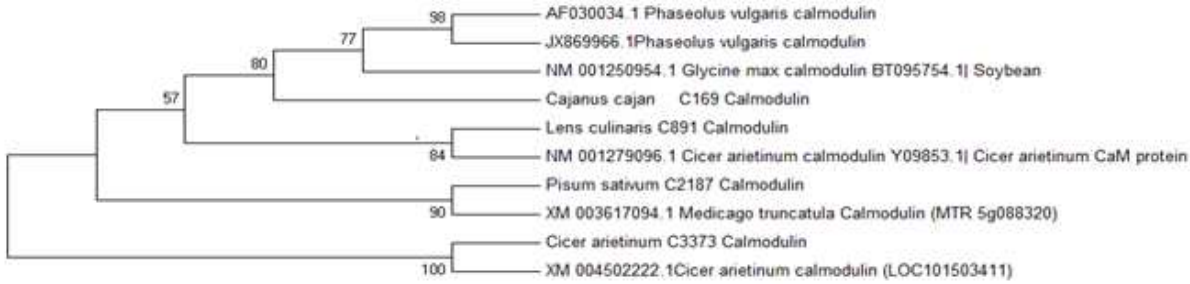
We assembled ESTs from *Cajanus cajan*, *Cicer arietinum*, *Pisum sativum*, *Lens culinaris* and applied diverse array of bioinformatics tools to extract information about gene content features, transcriptome changes and novel stress responsive genes both abiotic and biotic factors. The results

concerning the prevalence of protein related to various stress related genes in all four crops under study. Identifying and mining genes involved in stress response represent a key step to unraveling and manipulating stress tolerance in legumes. Comparative analysis among the legumes within the same species and between species will enable us to identify species specific genes underlying stress response. Despite knowing that comparisons between these legumes species data should be carefully inspected, our initiative established possible transcriptome elements that could guide the legumes specific community in unraveling the molecular mechanism that distinguish these four extremely important legumes species. In addition, the annotation of legumes-specific/stress prominent genes adds new element to genomic initiatives that our searching for traits (factors) that could differentiate legume each species from other. We believe that such data are a valuable aid to the interpretation of legume development, providing insight that could help in legumes reading program and indicating potential targets for functional analysis and biotechnology products of such socially and economically important legume species.

Appendix

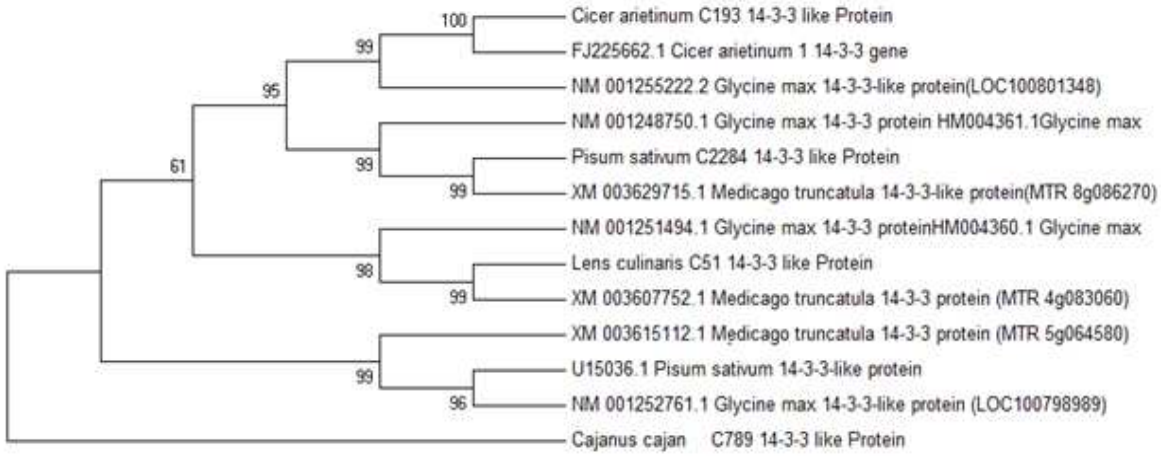
1. Phylogenetic Tree of Genes Involved in Regulatory Pathway (Cell Signaling and Transcription)

a. Calmodulin

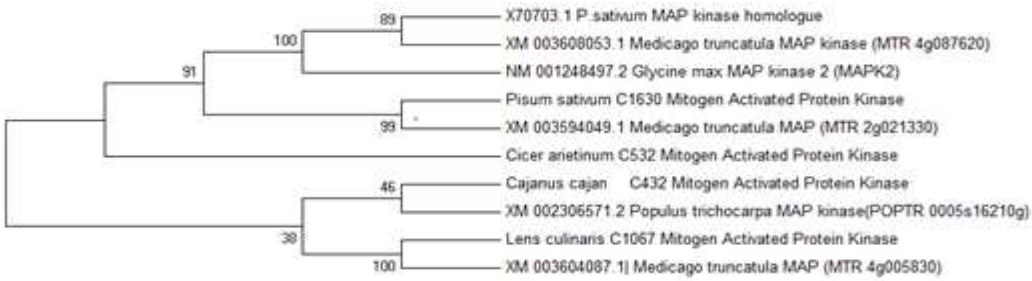


b. Serine/threonine- Protein Kinase

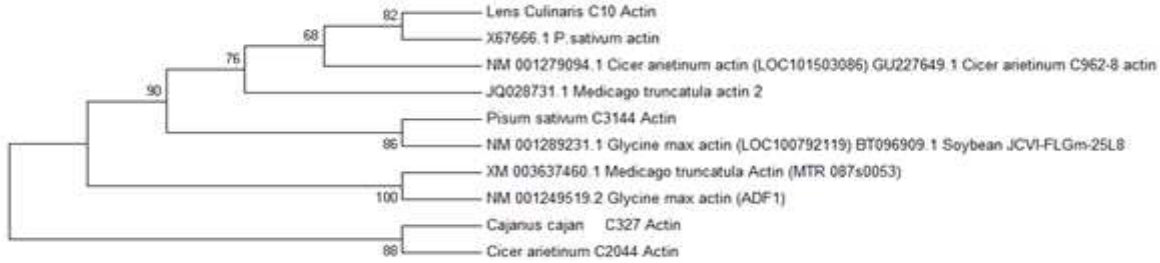
c. 14-3-3 like protein



d. MAP Kinase

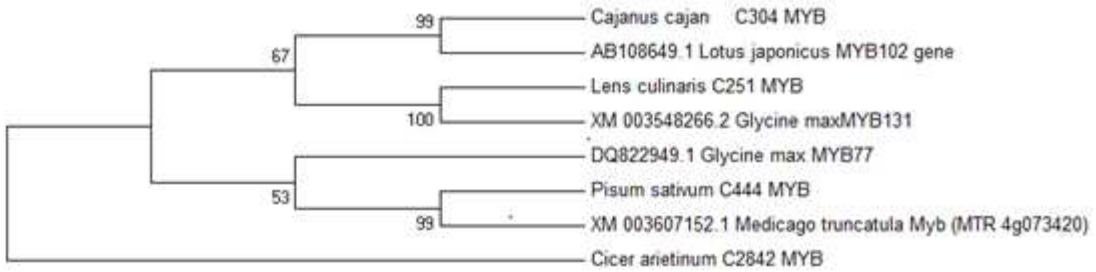


e. Actin

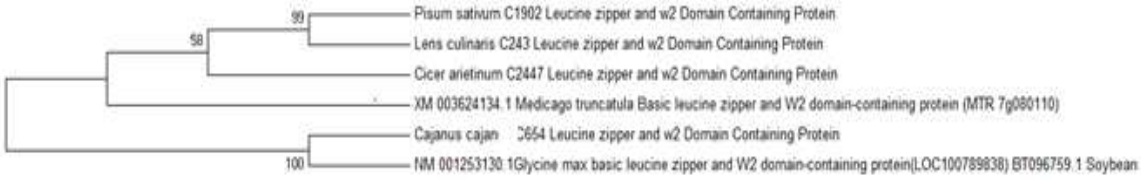


2. Phylogenetic Tree of Transcription Factors (TF) Acting in Stress

a. MYB



b. Leucine Zipper and W2 Domain Containing Protein



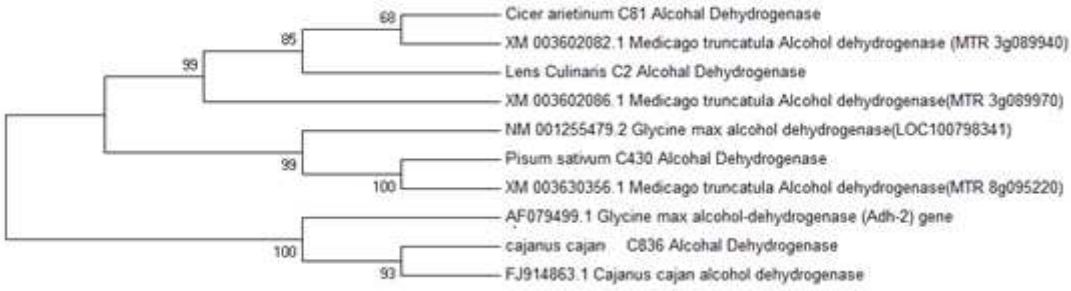
c. Zinc Finger Protein

d. AP-2/EREBP

e. WRKY

3. Phylogenetic Tree of Gene Involved in Metabolism

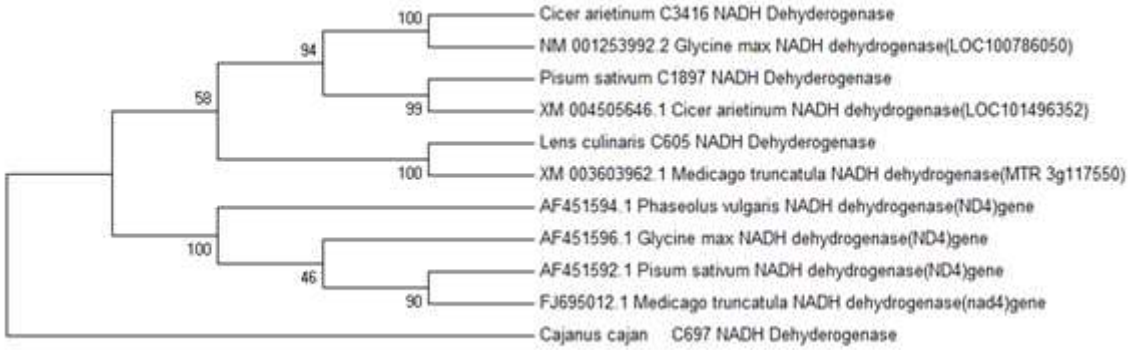
a. Alcohal Dehydrogenase



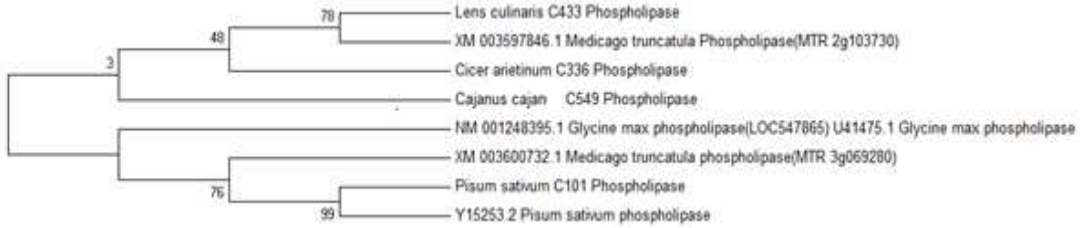
b. GAPDH

c. Malate Dehyderogenase

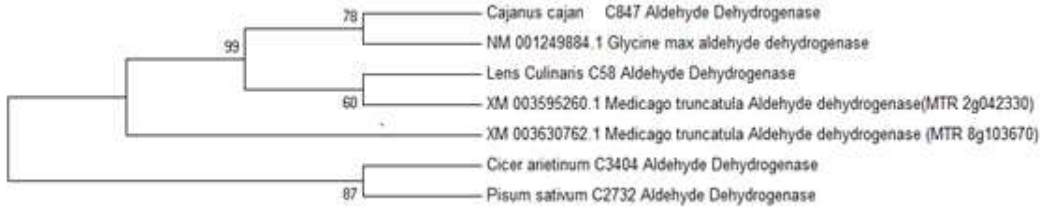
d. NADH Dehyderogenase



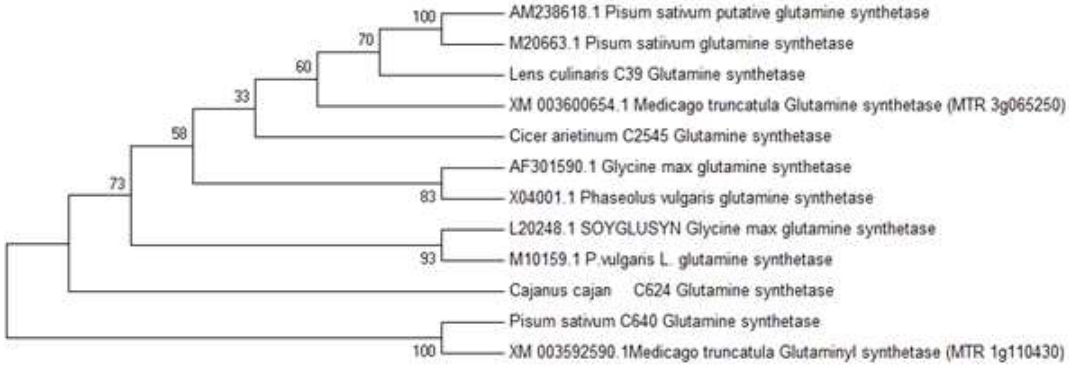
e. Phospholipase



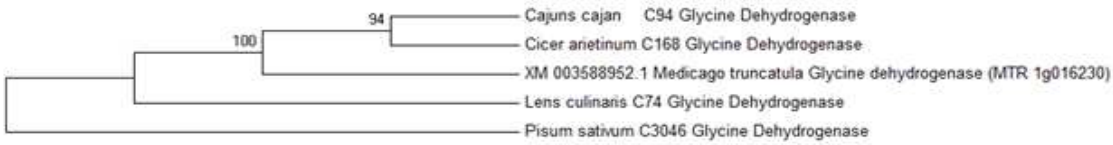
f. Aldehyde Dehydrogenase



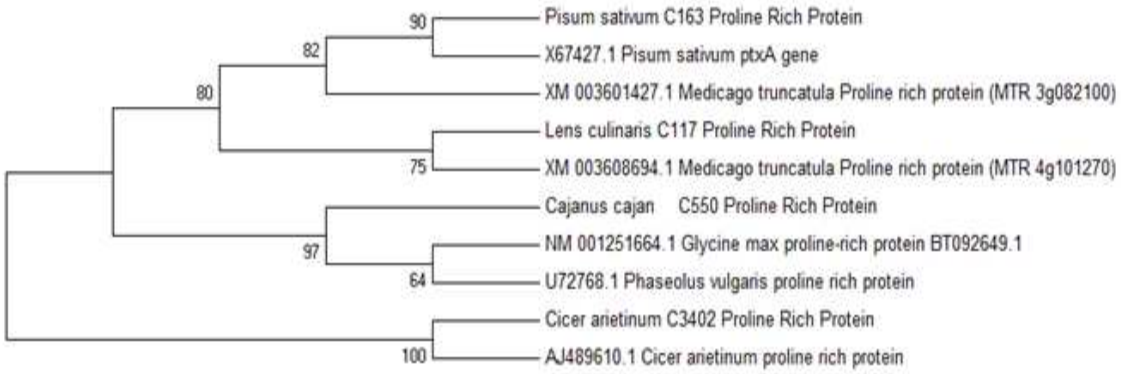
g. Glutamine synthetase



h. Glycine Dehydrogenase

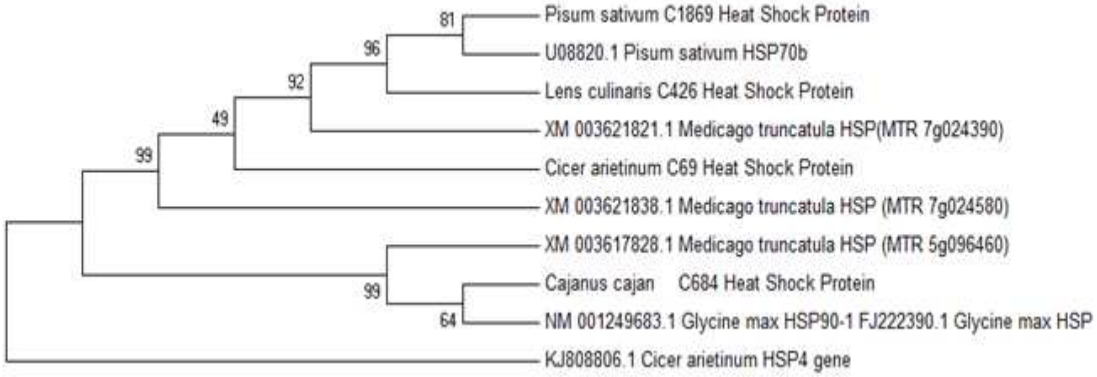


i. Proline Rich Protein



4. Phylogenetic Tree of Genes Related to Heat Stress

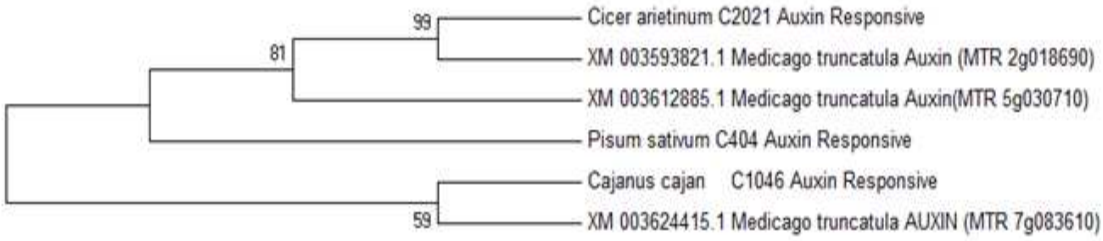
a. Heat Shock Protein



b. Stress Induced Protein

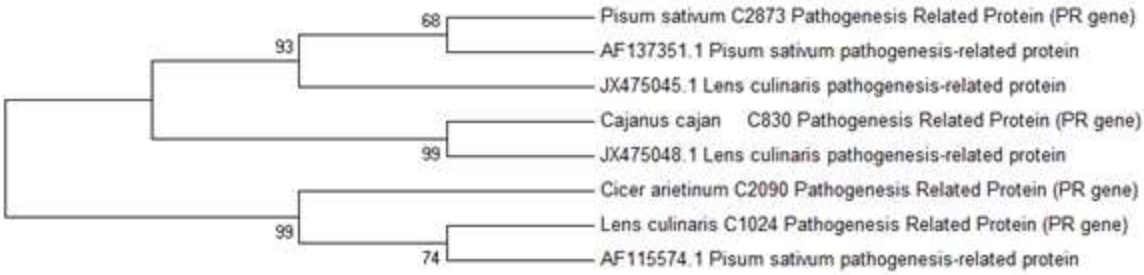
5. Phylogenetic Tree of Phytohormone

a. Auxin Responsive

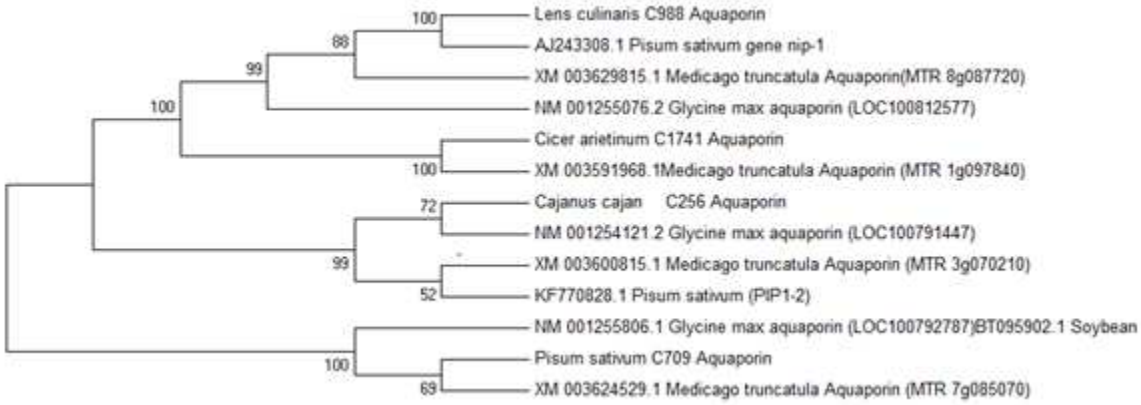


6. Phylogenetic Tree of Defense and Stress Responsive gene

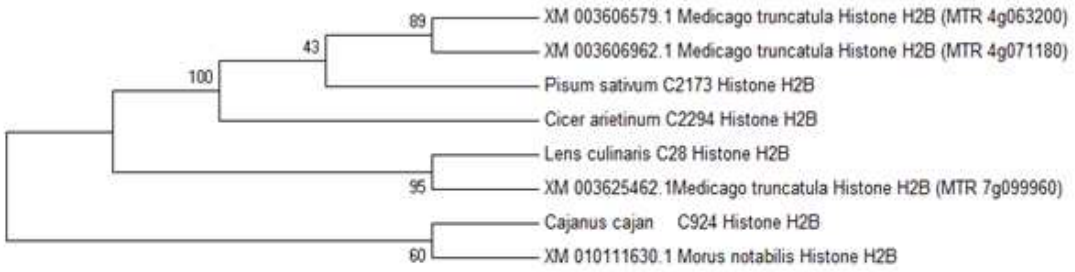
a. PR Gene



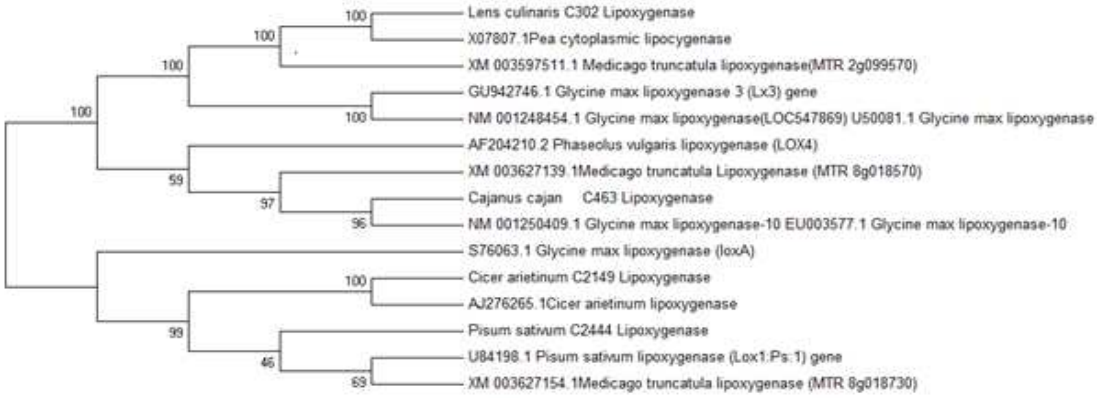
b. Aquaporin



c. H2B

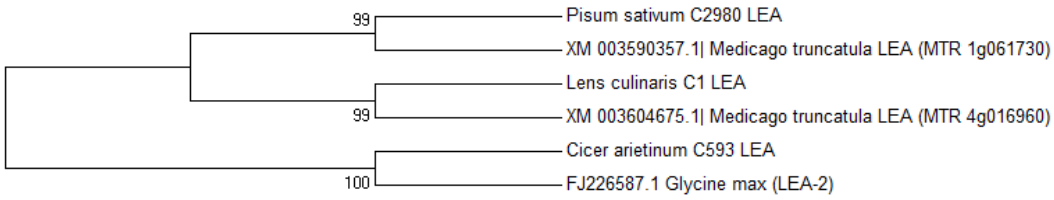


d. Lipoxygenase

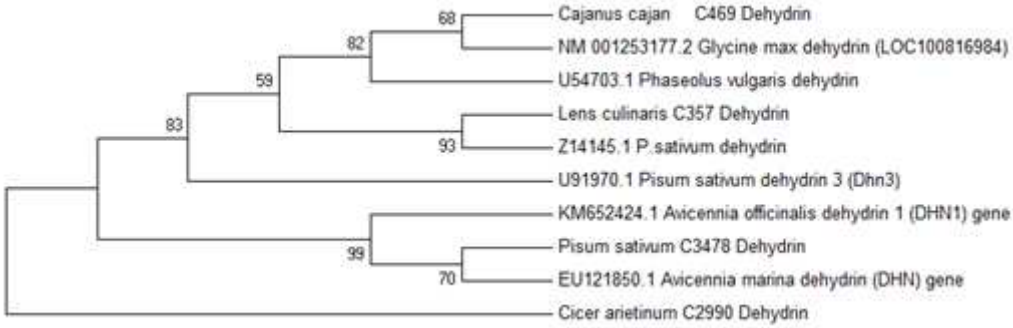


7. Phylogenetic Tree of Genes Related to Abiotic Stress and Detoxification

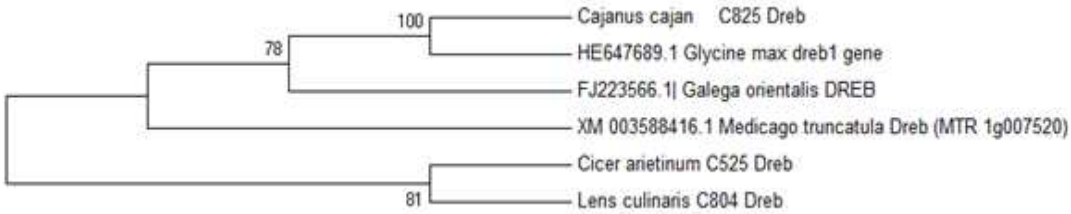
a. LEA



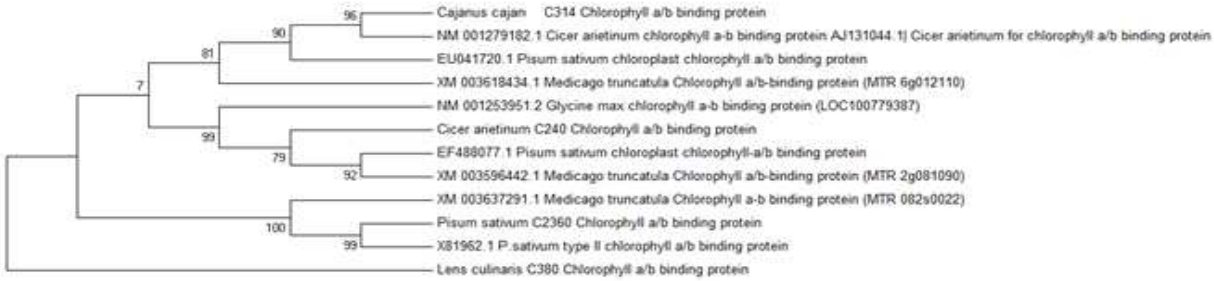
b. Dehydrin



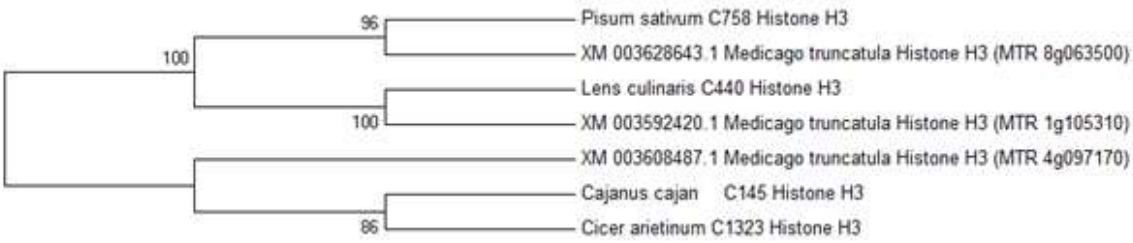
c. DREB



d. Chlorophyll a/b binding protein



e. Histone H3



References

[1] Aitken, A. (2006) 14-3-3 proteins: a historic overview. Semin. Cancer Biol. 16, 162–172.

[2] Ali Masoudi-Nejad, Koichiro Tonomura, Shuichi Kawashima, Yuki Moriya, Masanori Suzuki, Masumi Itoh, Minoru Kanehisa, Takashi Endo and Susumu Goto, EGassemb `CV1WFBler: online bioinformatics service for large-scale processing, clustering and assembling ESTs and genomic DNA fragments. Nucleic Acids Research, 2006, Vol. 34.

[3] Ariel, A.D., Manavella, P.A., Dezar, C.A. and Chan, R.C. (2007) The true story of the HD-Zip family. Trends Plant Sci., 12, 419–426.

[4] Ashraf M., Foolad M.R. (2007): Roles of glycine betaine and proline in improving plant abiotic stress resistance. Environmental and Experimental Botany, 59: 206–216.

[5] Ashraf M., Harris P.J.C., (2013). Photosynthesis under stressful environments, Photosynthetica. 51 (2): 163-190.

[6] Baker J, Steele C, LIII D (1988) Sequence and characterization of 6 Lea proteins and their genes from cotton. Plant Molecular Biology 11: 277–291.

[7] Bartel, D.,2001. Targeting detoxification pathways: An efficient approach to obtain plant with multiple stress tolerance?.Trends Plant. Sci 6,284-286.

[8] Benoît Ranty et al. (2006). Review Plant Calmodulins and Calmodulin-Related Proteins Multifaceted Relays to Decode Calcium Signals. Plant Signaling & Behavior 1:3, 96-104, May/June 2006]; ©2006 Landes Bioscience.

[9] Charu Lata and Manoj Prasad, Role of DREBs in regulation of abiotic stress responses in plants. Journal of Experimental Botany, Vol. 62, No. 14, pp. 4731–4748, 2011.

[10] Ciftci-Yilmaz, S and Mittler, 2008. The Zinc finger networks of plants. Cell Mol Life Sci. 65:1150-1160.

[11] Creelman RA, Mullet JE (1997) Biosynthesis and action of jasmonates in plants. Annu Rev Plant Physiol Plant Mol Biol 48: 355–381.

[12] Cushman JC, Meyer G, Michalowski CB, Schmitt JM, Bohnert HJ: Saltstress leads to differential expression of two isogenes ofphosphoenolpyruvate carboxylase during Crassulacean acidmetabolism induction in the common ice plant. Plant cell 1989, 7:715-725.

[13] De Smet I, Voss U, Lau S, et al. 2010. Unraveling the evolution of auxin signaling. Plant Physiology 155: 209–221.

[14] Devoto A, Muskett PR, Shirasu K (2003) Role of ubiquitination in the regulation of plant defence against pathogens. Curr Opin Plant Biol 6: 307–311.

[15] Dhawan R, Luo HL, FoersterAM, Abuqamar S, Du HN, Briggs SD, Mittelsten Scheid O, Mengiste T (2009) HISTONE MONOUBIQUITINATION1 interacts with a subunit of the mediator complex and regulates defense against necrotrophic fungal pathogens in Arabidopsis. Plant Cell 21: 1000–1019.

[16] Douce R, Bourguignon J, Neuburger M, Rebeille F. 2001. The glycine decarboxylase system: a fascinating complex. Trends in Plant Science 6, 167±176.

[17] Doyle JJ, Luckow MA (2003) The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. Plant Physiol 131(900):910.

[18] Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. Trends Plant Sci. 15: 573–581.

[19] El-Enany A.E,Morsy F and Dief N IMPACT OF WATER STRESS ON GROWTH CRITERIA AND ALCOHOL DEHYDROGENASE ACTIVITY OF THREE LEGUME PLANTS. Minia Science Bulletin Volume 25 (1), pp. 29-51, 2014.

[20] Feng Gao and Chun-Ting Zhang, GC-Profile: a web-based tool for visualizing and analyzing the variation of GC content in genomic sequences. Nucleic Acids Research, 2006, Vol. 34, W686–W691.

[21] Fleury D, Himanen K, Cnops G, Nelissen H, Boccardi TM, Maere S, Beemster GT, Neyt P, Anami S, Robles P, et al (2007) The Arabidopsis thaliana homolog of yeast BRE1 has a function in cell cycle regulation during early leaf and root growth. Plant Cell 19: 417–432.

[22] Fowler S, Thomashow MF. 2002. Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. The Plant Cell 14, 1675–1690.

[23] Fuentes S.I., Allen D.J., Ortiz-López A. & Hernández G. (2001) Over-expression of cytosolic glutamine synthetase increases photosynthesis and growth at low nitrogen concentrations. Journal of Experimental Botany 52, 1071–1081.

[24] Gupta, S.C., Sharma, A., Mishra, M., Mishra, R., Chowdhuri, D.K., 2010. Heat shock proteins in toxicology: how close and how far? Life Sci. 86, 377–384.

[25] Heineke D, Bykova N, GardestroÈm P, Bauwe H. 2001. Metabolic response of potato plants to an antisense reduction of the Pprotein of glycine decarboxylase. Planta 212, 880±887.

[26] Hu, M., Pei, B.L., Zhang, L.F. and Li, Y.Z., 2014. Histone H2B Monoubiquitination Is Involved in Regulating the Dynamics of Microtubules during the Defense Response to Verticillium dahliae Toxins in Arabidopsis. Plant Physiology,164,1857–1865.

[27] Hurst, L.D., 2002. The Ka/Ks ratio: diagnosing the form of sequence evolution. Trends. Genet. 02, 02722-1.

[28] Jahn, T. et al. (1997) The 14-3-3 protein interacts directly with the C-terminal region of the plant plasma membrane H+-ATPase. Plant Cell 9, 1805–1814.

[29] Javot, H., and Maurel, C. 2002. The role of aquaporins in root water uptake. Ann. Bot. 90:301-313.

[30] Johnson J.R, Cobb B.G., Drew M.C., (1994).Hypoxic induction of anoxia tolerance in roots of Zea mays, Plant Physiol. 105 :61-67.

[31] Jonak C, Kiegerl S, Ligterink W, Barker PJ, Huskisson NS, Hirt H. 1996. Stress signalling in plants: A mitogen-activated protein kinase pathway is activated by cold and drought. Proceedings of the National Academy of Sciences, USA 93, 11274–11279.

[32] Journet EP, Neuburger M, Douce R (1981) Role of glutamate-oxaloacetate transaminase and malate dehydrogenase in the regeneration of NAD for glycine oxidation by spinach leaf mitochondria. Plant Physiol 67: 467–469.

[33] Karlova R, Boeren S, Russinova E, Aker J, Vervoort J, de Vries S (2006) The Arabidopsis somatic embryogenesis receptor-like kinase1 protein complex includes brassinosteroid-insensitive. Plant Cell 18: 625–638.

[34] Kim J.M., To T.K., Ishida J., Morosawa T., Kawashima M., Matsui A., Toyoda T., Kimura H., Shinozaki K. & Seki M. (2008) Alterations of lysine modifications on the histone H3 N-tail under drought stress conditions in Arabidopsis thaliana. Plant CellPhysiology 49, 1580–1588.

[35] Laxalt A, Munnik T. Phospholipid signaling in plant defence. Curr Opin Plant Biol 2002; 5:332-8.

[36] Laxalt AM, Cassia RO, Sanllorenti PM, Madrid EA, Andreu AB, Daleo GR, Conde RD, Lamattina L: Accumulation of cytosolic glyceraldehyde-3-phosphate dehydrogenase RNA under biologicalstress conditions and elicitor treatments in potato. Plant MolBiol 1996, 30:961-972.

[37] Lehmann S, Funck D, Szabados L, Rentsch D (2010) Proline metabolism and transport in plant development. Amino Acid. 39(4):949-962.

[38] Ligterink W, Hirt H. 2001. Mitogen-activated protein (MAP) kinase pathways in plants: versatile signalling tools. International Review of Cytology 201, 209–215.

[39] Marques MC, Alonso-Cantabrana H, Forment J, Arribas R, Alamar S, Conejero V, Perez-Amador MA: A new set of ESTs and cDNA clones from full-length and normalized libraries for gene discovery and functional characterization in citrus. BMC Genomics 2009, 10:428.

[40] Mascarenhas, J.P., 1993 Molecularm mechanisms of pollen tube growth and differentiation. Plant Cell 5: 1303-1314.

[41] Maurel, C., Verdoucq, L., Luu, D. T., and Santoni, V. 2008. Plant aquaporins: Membrane channels with multiple integrated functions. Annu. Rev. Plant Biol. 59:595-624.

[42] Meagher, B., 1991 Divergence and differential expression of actin gene families in higher plants. Int. Rev. Cytol. 125 139-163.

[43] Miyata, T., T. Yasunaga, and T. Nishida. 1980. Nucleotide sequence divergence and functional constraint in mRNA evolution. Proc. Natl. Acad. Sci. USA 77:7328–7332.

[44] Morimoto, R.I., Santoro, M.G., 1998. Stress-inducible responses and heat shock proteins: new pharmacologic targets for cytoprotection. Nat. Biotechnol. 16, 833–838.

[45] Munnik T, Testerink C. Plant phospholipid signaling: “in a nutshell”. J Lipid Res 2009; 50:260-5.

[46] Nekrutenko, A., Makova, K. D. and Li, W.H., 2002. The Ka/Ks ratio test for assessing the protein-coding potential of genomic regions: an empirical and simulation study. Genome Res. 12, 198–202.

[47] Pashley CH, Ellis JR, McCauley DE, Burke JM: EST databases as a source for molecular markers: lessons from Helianthus. Journalof Heredity 2006, 97(4):381-388.

[48] Pavlíková D., Pavlík M., Staszkova L., Motyka V., Száková J., Tlustoš P., Balík J. (2008): Glutamate kinase as a potential biomarker of heavy metal stress in plants. Ecotoxicology and Environmental Safety, 70: 223–230.

[49] Pruitt KD, Tatusova T, Maglott DR: NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 2007, 35 Database: D61-65.

[50] Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, et al: The Pfam protein families database. Nucleic Acids Res 2012, 40 Database: D290-301.

[51] Rawyler A., Arpagaus S., Braendle R., (2002). Impact of oxygen stress and energy availability on membrane stability of plant cells. Ann. Bot. 90: 499- 507.

[52] Roberts, M.R. (2000) Regulatory 14-3-3 protein–protein interactions in plant cells. Curr. Opin. Plant Biol. 3, 400–405.

[53] Robinson MJ, Cobb MH. 1997. Mitogen-activated protein kinase pathways. Current Opinion in Cell Biology 9, 180–186.

[54] Roth C, Liberles DA: A systematic search for positive selection in higher plants (Embryophytes). BMC Plant Biol 2006, 6:12.

[55] Rushton PJ, Somssich IE, Ringler P, Shen QJ. 2010. WRKY transcription factors. Trends in Plant Science 15, 247–258.

[56] Russell RA, Sachs MM (1989) Differential expression and sequence analysis of the maize glyceraldehydes 3-phosphate dehydrogenase gene faxnily. Plant Cell 1: 793-803.

[57] Schmitz RJ, Tamada Y, Doyle MR, Zhang XY, Amasino RM (2009) Histone H2B deubiquitination is required for transcriptional activation of FLOWERING LOCUS C and for proper control of flowering in Arabidopsis. Plant Physiol 149: 1196–1204.

[58] Sels J, Mathys J, De Coninck BM, Cammue BP, De Bolle MF: Plant pathogenesis-related (PR) proteins: a focus on PR peptides. Plant Physiol Biochem 2008, 46(11):941-950.

[59] Seo S, Okamoto M, Seto H, Ishizuka K, Sano H, Ohashi Y. 1995. Tobacco MAPK kinase: a possible mediator in wound signal transduction pathways. Science 270, 1988–1992.

[60] Siedow JN (1991) Plant lipoxygenase: structure and function. Annu Rev Plant Physiol Plant Mol Biol 42: 145–188.

[61] Somerville CR. 2001. An early Arabidopsis demonstration resolving a few issues concerning photorespiration. Plant Physiology 125, 20±24.

[62] Stirnberg P, Zhao S,Williamson L,Ward S, Leyser O. 2012. FHY3 promotes shoot branching and stress tolerance in Arabidopsis in an AXR1-dependent manner. Plant Journal 71: 907–920.

[63] Stukenbrock EH, McDonald BA: Population genetics of fungal and oomycete effectors involved in gene-for-gene interactions. Mol Plant Microbe Interact 2009, 22(4):371-380.

[64] Tamura,K.,Peterson,D.,Peterson,N., Stecher,G.,Nei,M.,andKumar, S. (2011).MEGA5:molecular evolutionary genetics analysis usingmaximum likely hood, evolutionarydistance, and maximum parsimonymethods. Mol. Biol.Evol. 28, 2731–2739.doi: 10.1093/molbev/msr121.

[65] Timperio, A.M., Egidi, M.G., Zolla, L., 2008. Proteomics applied on plant abiotic stresses: role of heat shock proteins (HSP). J. Proteomics 71, 391–411.

[66] Ting-Wen Chen, Ruei-Chi Richie Gan, Timothy H Wu4, Po-Jung Huang, Cheng-Yang Lee, Yi-Ywan M Chen, Che-Chun Chen, Petrus Tang FastAnnotator- an efficient transcript annotation web tool. BMC Genomics 2012.

[67] Tripp, J., Mishra, S.K., Scharf, K.-D., 2009. Functional dissection of the cytosolic chaperone network in tomato mesophyll protoplasts. Plant Cell Environ. 32, 123–133.

[68] Trujillo M, Shirasu K (2010) Ubiquitination in plant immunity. Curr Opin Plant Biol 13: 402–408.

[69] Urao, T., Yamaguchi-Shinozaki, K., Urao, S., and Shinozaki, K. (1993). An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. Plant Cell 5: 1529–1539.

[70] Van Loon LC, Rep M, Pieterse CM: Significance of inducible defenserelated proteins in infected plants. Annu Rev Phytopathol 2006, 44:135-162.

[71] Varshney RK et al. Trends Plant Sci. 2005 10: 621 [PMID: 16290213]

[72] Wahid, A., Gelani, S., Ashraf, M., Foolad, M.R., 2007. Heat tolerance in plants: an overview. Environ. Exp. Bot. 61, 199–223.

[73] Williams, E. J. D., and L. D. Hurst. 2002. Is the synonymous substitution rate in mammals gene-specific? Mol. Biol. Evol. 19:1395–1398.

[74] Xu L, Ménard R, Berr A, Fuchs J, Cognat V,Meyer D, Shen WH (2009) The E2 ubiquitin-conjugating enzymes, AtUBC1 and AtUBC2, play redundant roles and are involved in activation of FLC expression and repression of flowering in Arabidopsis thaliana. Plant J 57: 279–288.

[75] Yang Y, Kwon H-B, Peng H-P, Shih M-C (1993) Stre!is responses and metabolic regulation of glyceraldehyde-3-phosphate dehydrogenase genes in Arabidopsis. Plant Physiol 101: 209-216.

[76] Yang Z, Bielawski JP: Statistical methods for detecting molecular adaptation. Trends Ecol Evol 2000, 15(12):496-503.

[77] Yasuda H, Nakajima M, Ito T, Ohwada T, Masuda H: Partial characterization of genes whose transcripts accumulate preferentially in cell clusters at the earliest stage of carrot somatic embryogenesis. Plant Mol Biol 2001, 45(6):705-712.

[78] Yoshida, A., et al.,1998. Human aldehyde dehydrogenase family. Eur.J. Biochem.51,549-557.

[79] Zhang S, Klessig DF. 1998. The tobacco wounding-activated mitogen-activated protein kinase is encoded by SIMK. Proceedings of the National Academy of Sciences, USA 95, 7225–7230.

[80] Zhang XS, Choi JH (2001) Molecular evolution of calmodulin-like domain protein kinases (CDPKs) in plants and protists. J Mol Evol 53: 214–224.

[81] ZhangQ, Li J,ZhangW, et al. 2012.The putative auxin efflux carrier OsPIN3t is involved in the drought stress response and drought tolerance. Plant Journal 72: 805–816.