ISSN print: 1224-5984
ISSN online: 2248-3942

Rom Biotechnol Lett. 2021; 26(2): 2384-2395
doi: 10.25083/rbl/26.2/2384.2395

Received for publication, July, 21, 2020
Accepted, November, 24, 2020

Original paper

In silico identification of cis-Regulatory elements and their functional annotations from assembled ESTs of *Artemisia annua L.* involved in abiotic stress signaling

PRAVEJ ALAM[1]*, THAMER AL BALAWI[1]

[1]Department of Biology, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, 11942, Saudi Arabia

Abstract

Salt stress is a common side-effect in plants impacted on plant growth, metabolism and productivity. *A. annua* L. is one of the well-known antimalarial plants, biosynthesized artemisinin in its leaf, now introduced in all-over the world. In this article, we have analyzed the *A. annua* L. ESTs under salt stress and predicted cis-regulatory elements, roles in abiotic stress signaling. Further, the predicted abiotic stress responsive factors were analyzed in order to their function annotations as compare to the genome of *Arabidopsis thaliana*. 11 EST-contigs assembled from 127 were 29 signals elements were identified by CAP3 program. In order to evaluate accuracy of the identified factors, gene ontology functions were performed. GOBP analysis enriched the genes (85.71%) as the response to abiotic signaling. The co-expression analysis was revealed by gene investigators and String 10.0, these factors-oriented genes had at least 0.40 correlations and 0.7 mutual connection. In projected PPI network, the recognized factors belong to plant hormone signaling and diterpene pathways. These factors (ABF1, APX CCC1, CPK6, JAZ1, MYC2) introduced as candidate genes responsive factors could be overexpressed in *A. annua* L. plants either alone or in a shuttle may led the good metabolism and higher artemisinin content.

Keywords Artemisinin, ESTs, salt stress, cis-regulatory elements, GOBP, Co-Expression.

To cite this article: ALAM P, BALAWI AT. In silico identification of Cis-Regulatory elements and their functional annotations from assembled ests of artemisia *Annua L.* involved in abiotic stress signaling. *Rom Biotechnol Lett*. 2021; 26(2): 2384-2395. DOI: 10.25083/rbl/26.2/2384.2395

*Corresponding author: PRAVEJ ALAM, Department of Biology, College of Science and Humanities, Prince Sattam bin Abdulaziz University, 11942, Al-Kharj, Kingdom of Saudi Arabia (KSA)
E-mail: alamprez@gmail.com; ps.alam@psau.edu.sa

Introduction

The plant growth is stunted by wide spectrum of abiotic (Salinity, drought heavy metals) and biotic factors (Fungus, bacteria; JENKS and HASEGAWA, 2005). Due to stress conditions especially abiotic, 50% crop productivity loss in recent past (HAGGAG et al, 2015; PELEG et al, 2011). The abiotic stress responses downstream the plant growth and development process (hormonal regulation) and also affect the changes in the molecular and biochemical phenomenon (RAMAKRISHNA and RAVISHANKAR, 2011; SHESHADRI et al, 2016). Moreover, the researcher proved that these changes are governed by gene (s) or combination of cis-regulatory elements (TFs) especially those associated with transcriptional regulation of associated genes (SHESHADRI et al, 2016; JOSHI et al, 2016). These elements bind to their specific recognition sites of stress responsive genes and affects the primary and secondary metabolites networks in the plants systems under abiotic conditions (DURAISAMY et al, 2016; CARETTO et al, 2015; OSAKABE et al, 2014; SCHULTER et al, 2013).

Artemisinin (sesquiterpenes) is biosynthesized and accumulate in the leaf of *A. annua* L. plants and strictly controlled by spatial and temporal manner and affected by several abiotic factors responsive elements led the low artemisinin content (~1.0%) is reported in *A. annua* L. plants wordwide (SHEN et al, 2019; ALAM et al, 2016; ALAM and ABDIN, 2012; ALAM et al, 2010; MA et al, 2009).

These spatiotemporal elements play an important role in metabolic pathways to up-regulate the metabolic fluxes by interacting the signature sequences i.e. regulatory region in plant genome of the targeted gene (s) and modulate the transcription levels (BUSSEMAKER et al, 2001). Despite the remarkable progresses have been achieved in artemisinin production either though understand the role of cis-elements or genetic engineering approaches in plants and microbes (ABDIN and ALAM, 2015; PADDON et al, 2013), *A. annua* is the only valuable source for artemisinin (PEPLOW, 2016; ABDIN and ALAM, 2015). However, the transcriptional regulation of artemisinin pathway in *A. annua* is not yet well established. More recently, some transcription factors (AabZIP9, AabZIP1, AaABF3, AabZIP1) are characterized and targeted as transcriptional regulators those playing a significant role in *Artemisia ann*ua L. plant development, metabolism under abiotic stress conditions (CHEN et al, 2016; TANG et al, 2012; OSTERLUND et al, 2000; ANG et al, 1998).

Hence, we, emphasized that the cis-elements is the key factors to play an important role in artemisinin biosynthesis under the stress conditions. In this article, we predicted the cis-elements from *Artemisia annua* L. ESTs retrieve from NCBI db to assess their biological role associated with particular gene (s) by comparing the genome of *Arabidopsis thaliana* (TAIR data base) under salt condition. The functional analysis; Gene Ontology (GO) enrichment analysis of the predicted unigenes were also assessed their role or resemble function in term of cellular biological and molecular.

Materials and Methods

In this study, 127 ESTs of *A. annua* L. were retrieved from NCBI under salt stress. In order to analysis the similarity-based clustering of the ESTs for its enhance effectiveness were screened to abolish the vector sequences caused inappropriate clustering through *Vec screen* on line tool from NCBI (https://www.ncbi.nlm.nih.gov/tools/vecscreen/). The trimmed ESTs sequences were assembled into clusters constructed on the basis of sequences resemblance to obtained the constant clusters using CAP3 online program (http://doua.prabi.fr/software/cap3), with default (HUANG and MADAN, 1999). Each cluster were separately the with overlap identity as cut off range showed the minimum sum of nucleotides. Ungrouped sequences, with low identity recorded as singletons may be a formed of genes or may be an outcome of contamination, were not counted for abiotic stress-responding elements and functional annotations.

Identification of abiotic stress-responding elements and its functional annotation

PlantCare, Plant PAN were used for transcriptions start site (TSS) and the promoter region for all of these genes were obtained. The *cis*-regulatory elements of abiotic signaling related gene were predicted from Plant CARE (LESCOT et al, 2002), and PlantPAN3.0 (CHOW et al, 2019) to find out an occurrence of motifs at the promoter regions. The presence of abiotic stress responsive factors annotation at the promoter region of a gene was considered as a marker for determining the putative abiotic stress responsive role for that gene. This consideration was based on the ration that a large portion of the gene expression control takes place at the transcriptional level. After identification of abiotic stress-responsive genes, functional analysis of the assembled contigs was performed, Gene Ontology (GO) enrichment analysis for stress-responsive genes were carried out to determine their role in biological processes, molecular function, and cellular component using GO vocabularies by online Gene Ontology tool (goProfiler), Blast2go and STRING 10.0 (http://string-db.org).

Results

Contigs prediction: EST assembly

Identification of putative abiotic stress responsive genes of *A. annua* L. were predicted by analyses of ESTs, under salinity. A total of 127 ESTs of *A. annua* L. were obtained from NCBI GenBank. Out of 127 EST sequences, 2 sequences were trimmed and rest of 125 sequences were assembled and predicted 11 contigs through the CAP3 program (http://doua.prabi.fr/software/cap3). Assembled ESTs interpretation for solitary 8.66% of the size of total number of ESTs. Less abundant or deprived ESTs were not assembled and known as singletons (26).

Identification of TFs

In addition, cis-regulatory motifs were predicted by using the PlantCARE and PlantPAN 3.0 data base from 11 assembled contigs of *A. annua* L. plants. These 11 assembled contigs were predicted, cis-regulatory motifs involved in cold, light, salt, abscisic acid etc. responses. Of 127

collected unigenes under salinity condition, at least more than one cis-regulatory motifs were found at the promoter region of 11 contigs to 29 signals elements. Since many of the genes had unknown protein annotation, the accession number and description were obtained from PlantPAN 3.0, TAIR, UniProtKB for cis-regulatory motifs (Table 1).

Table 1. Prediction of cis-motif from ESTs-contigs *Artemisia annua* L. under salinity

Signals	ID	Functions	Sequence
ABRE	Contigs1 Contigs3 Contigs9 Contigs11	Abscisic acid responsive element-binding factor 1	ACGTG
CAAT-box	Contigs1 Contigs2 Contigs3 Contigs4 Contigs5 Contigs6 Contigs7 Contigs8 Contigs9 Contigs10 Contigs11	Encodes a homeodomain leucine zipper class I (HD-Zip I) meristem identity regulation	CCAAT
DRE core	Contigs1 Contigs5 Contigs11	Cold tolerance in maize	GCCGAC
G-Box	Contigs1 Contigs3 Contigs11	Phytochrome interacting factor 3-light responsive	CACGTT
GATA-motif	Contigs1 Contigs4 Contigs7 Contigs10	GATA type zinc finger transcription factor family protein-light	GATAGGA
MYB	Contigs1 Contigs2 Contigs3 Contigs4 Contigs5 Contigs8 Contigs9 Contigs11	Homeodomain-like superfamily protein-response to salt stress	TAACTG
P-box	Contigs1 Contigs5	Gibberellin-responsive element-rice	CCTTTTG
STRE	Contigs1 Contigs2 Contigs4 Contigs5 Contigs6 Contigs7 Contigs8 Contigs11	Salt tolerance zinc finger	AGGGG
Sp1	Contigs1 Contigs8	Phosphotyrosine protein phosphatases superfamily protein-light	GGGCGG
TCA-element	Contigs1 Contigs2 Contigs4 Contigs5 Contigs11	Abscisic acid responsive element-binding factor 1, is-acting element involved in salicylic acid responsiveness	CCATCTTTTT
CGTCA-motif	Contigs2 Contigs3 Contigs6	Cis-acting regulatory element involved in the MeJA-responsiveness,Hordeum vulgare	CGTCA
MYC	Contigs2 Contigs3 Contigs4 Contigs5 Contigs11	Jasmonate-zim-domain protein 1	CAATTG
TGACG-motif	Contigs2 Contigs3 Contigs6	Cis-acting regulatory element involved in the MeJA-responsiveness, basic leucine zipper transcription factor involved in the activation of SA-responsive genes	TGACG
as-1	Contigs2 Contigs3 Contigs6	salt stress	TGACG
ARE	Contigs3 Contigs4 Contigs8	Response to salt stress, response to heat	AAACCA
LTR	Contigs3	Cis-acting element involved in low-temperature responsiveness-Hordeum vulgare	CCGAAA
MBS	Contigs2 Contigs3 Contigs8	MYB binding site involved in drought-inducibility-Abscisic acid-activated signaling pathway	CAACTG
TCT-motif	Contigs4	Part of a light responsive element-Arabidopsis	TCTTAC
AP-1	Contigs5	Gibberellic acid biosynthetic pathway-light	TGAGTTAG
GARE-motif	Contigs5	Gibberellin-responsive element-Bracica oleracea	TCTGTTG
ERE	Contigs7	Ethylene responsive binding fator	ATTTTAAA
GT1-motif	Contigs7	Light responsive element-arabidosis thaliana	GGTTAA
TGA-element	Contigs8	Auxin-responsive element,gibberellin catabolic process-heat	AACGAC
W box	Contigs8	Homeodomain-like superfamily protein-salt stress	TTGACC

GC-motif	Contigs9	Enhancer-like element involved in anoxic specific inducibility-abscisic acid-activated signaling pathway	CCCCCG
box S	Contigs9	Cold tolerance -Homeodomain-like superfamily protein	GCCACT
CAT-box	Contigs11	Jasmonic acid mediated signaling pathway-cold	GCCACT
CCAAT-box	Contigs3 Contigs11	Member of Cation-chloride co-transporter family-salt stress	CAACGG
WRE3	Contigs1 Contigs9	wound-response element reported in *Pisum sativum*	CCACCT
BOX 4	Contigs4 Contigs7 Contigs10	Part of a conserved DNA module involved in light responsiveness in Petroselinum crispum	ATTAAT
TC-rich repeats	Contigs7	Cis-acting element involved in defense and stress responsiveness in Nicotiana tabacum	GTTTTCTTAC

Functional annotation of EST-contigs

Moreover, 11 contigs were functionally annotated and compared with the NCBI nr database by using the BLAST2GO OMICS box program. The BLAST results revealed above 50% sequences were conserved to *Arabidopsis thaliana* followed by *Lemna gibba* and *Nicotiana tabacum* as compared to others (Fig. 1-2). Further, these ESTs-contigs were carried out to gene ontology (GO) for their functional role. The available GO terms; only 4 EST-contigs. 27 GO terms repossessed, representing a usual of 2 GO terms/contig (minimum) and one contig for 18 GO terms (maximum) were recorded for this study (Fig. 3).

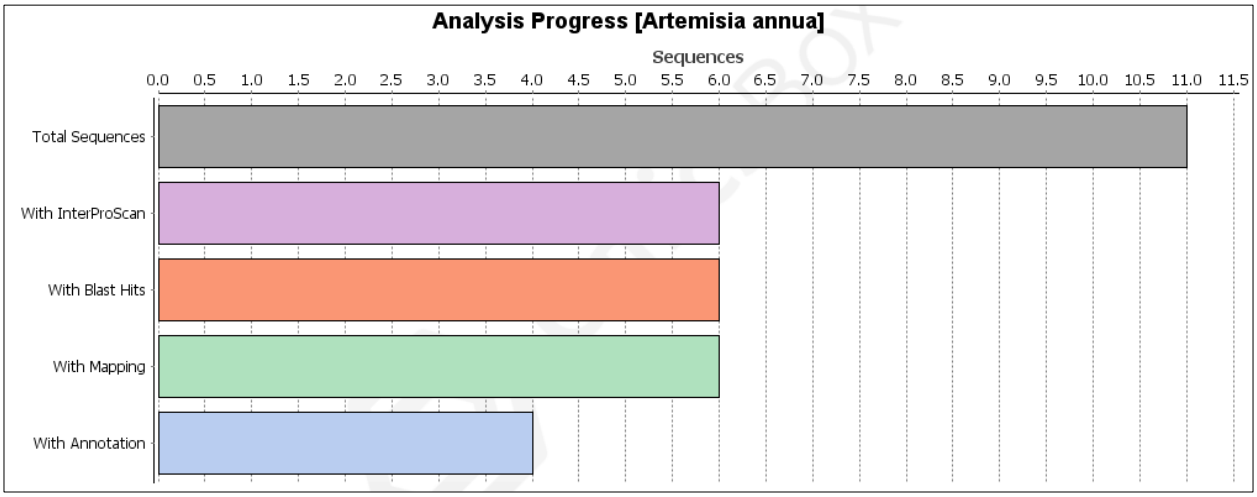

Figure 1. Total number of sequences (contigs) analysed through BLAST, InterproScan, mapping and annotation by OMICS BOX having blast2go program.

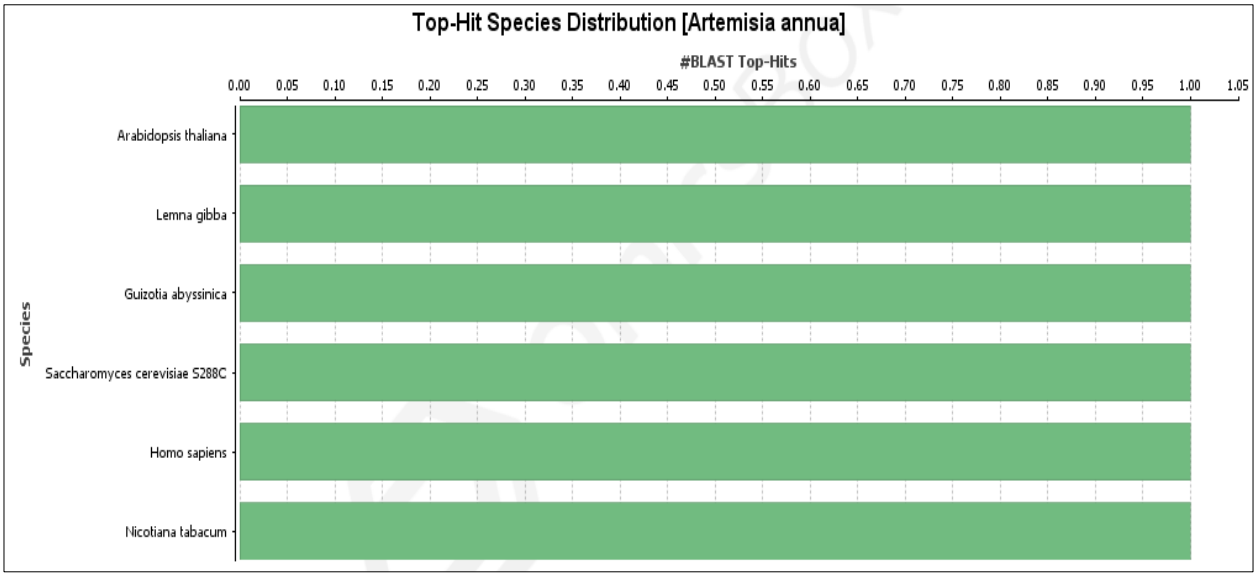

Figure 2. Top hit species distribution of assembled contigs from *A. annua* L. plants.

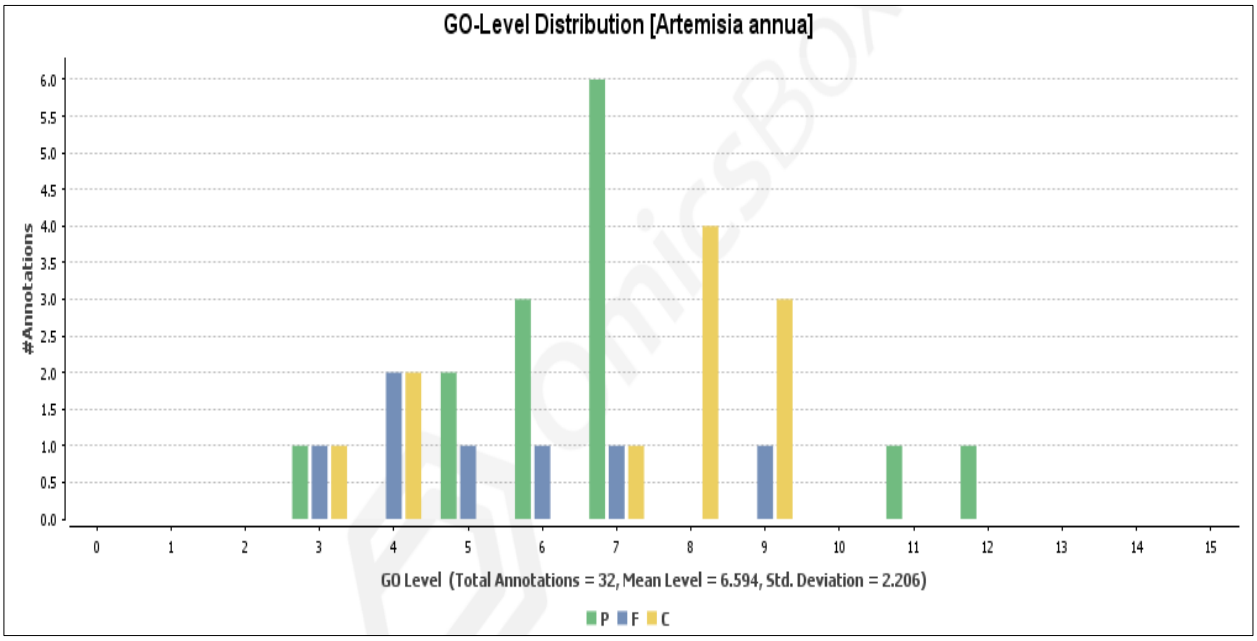

Figure 3. GO-level distribution graph produced by BLAST2GO for ESTs-contigs sequences of *Artemisia annua* L. P, biological process; F, molecular function; and C, cellular component.

Further, functional characterized of 11 ESTs-contigs were for further analyzed for biological process, cellular component and molecular function at level three (Fig. 4a, b, c). GO analysis in which whole genome annotation assumed as reference (*Arabidopsis thaliana*), demonstrated that 28 identified genes out of 31, were significantly enriched as abiotic signaling responsive genes at p-value of 0.05. According to Gene Ontology-Biological Process (GOBP) analysis, 31 enriched genes had hits with GOBP belongs to abiotic stress and diterpenoid biosynthesis signaling by using gProfilar program. The majority of genes were enriched as response to stimulus (cluster frequency of 29%) and (50%) response to abiotic stress signaling. Furthermore, few of gene enriched as a specific GOBP response to abscisic acid, jasmonic acid, auxins and gibberellic acid for response under salinity (Table S1). GO Slim-plant analysis enriched genes into response to stress with p-value of 1.16E-07 and 24/28 number of gene (cluster frequency of 85.71%) (Fig. 4a, b and c).

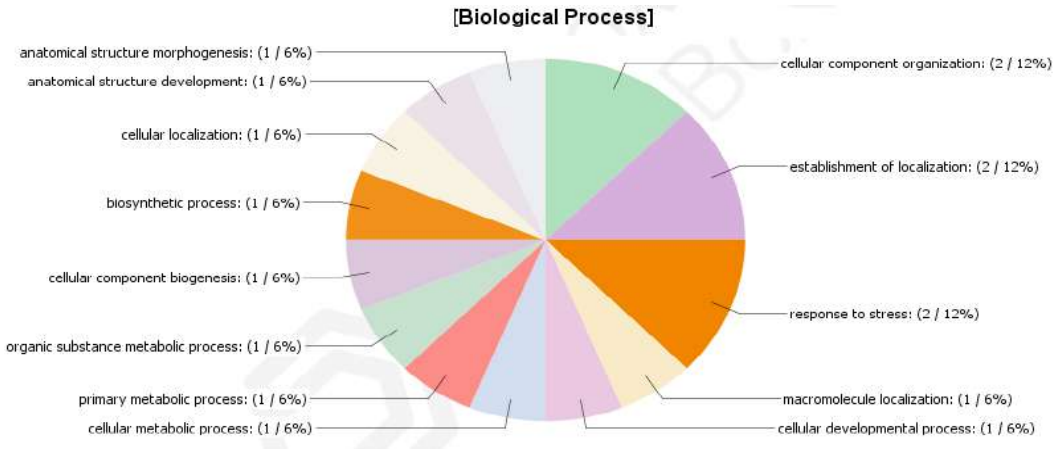

Fig. 4a

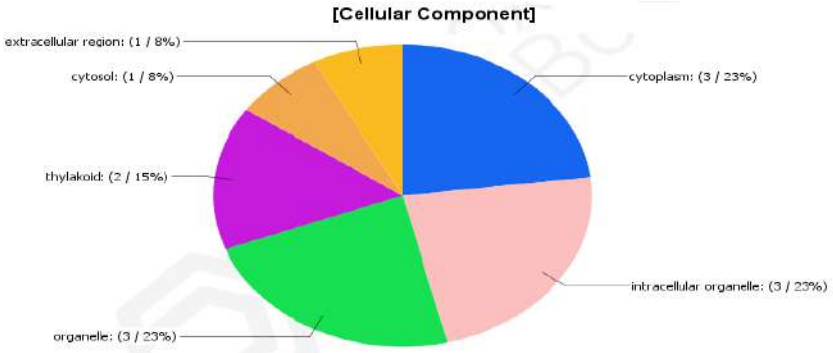

Fig. 4b

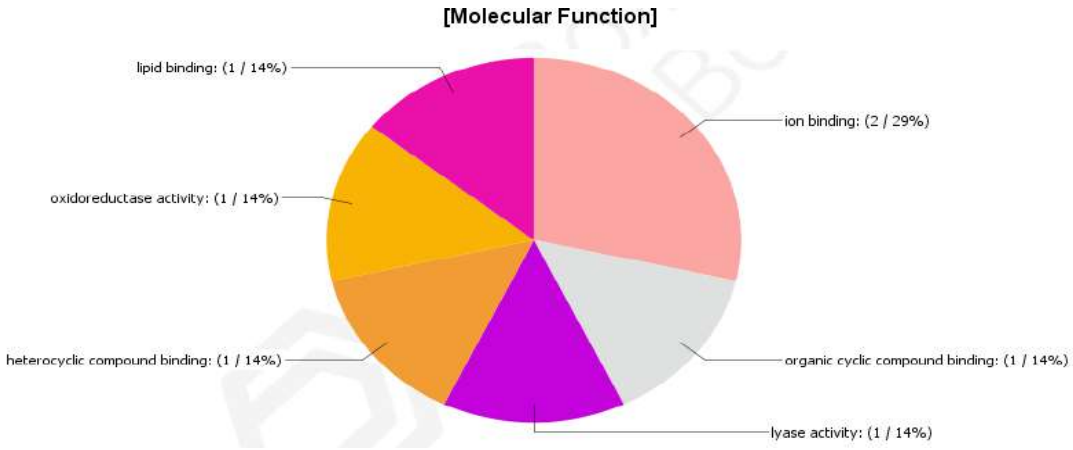

Fig. 4c

Figure 4. Gene ontology classification of EST-contig sequences of *Artemisia annua* L. (**a**): Distribution of GO terms in the biological process category. (**b**): Distribution of GO terms of in the cellular process category. (**c**): Distribution of GO terms of in the molecular function category.

Other genes were involved either in biotic process like wound healing, signal transduction, cell communications under salt stress (Table S2). The results of GO Slim-plant analysis putative stress responsive genes determined by cis-regulatory element detection method is further confirmed through the STRING 10.0 indicated the presence of specific domain known for stress response (Fig. 5). The results of STRING 10.0 based on PFAM Protein Domains PF13188, PF08446,PF,00360, PF08448 and PF01590) and INTERPRO Protein Domains (IPR0035965, IPR029016, IPR018467, IPR016132 and IPR013767 and PF00582) and Features databases indicated 5 domains belongs to abiotic universal stress protein family (Pathway ath04712, ath04075 and ath00904) were significantly enriched for proteins AT2G17290.1 and AT1G01560.2 (Fig. 5 Table S2). All identified putative abiotic stress responsive genes indicated correlation above 0.4 and mutual connection of at least 0.7. The results of co-expression analysis were confirmed using produced Protein-Protein Interaction (PPI) network by gene investigators (Fig. 5).

Discussion

The substantial progresses in artemisinin biosynthesis are achieved in heterologous hosts by using the different cutting-edge technology (ABDIN and ALAM, 2015; PADDON et al, 2013; ALAM and ABDIN, 2011). At this stage, *A. annua* is still the only reliable and effective resource for artemisinin production. It is because of artemisinin and its competitive pathways which played an important role to transfer the carbon fluxes towards artemisinin biosynthesis (PEPLOW, 2016; ABDIN and ALAM, 2015) in *A. annua* L. plants. Though, the transcriptional factors and its regulation controlled the artemisinin biosynthesis pathway in *A. annua* is not yet well recognized. But some factors e.g bZIP transcription, abscisic acid (ABA) response AabZIP1, and AaABF3 and WRKKY were studied by some investigators and play important role for plant growth development processes, metabolism, with under un favorable conditions i.e. abiotic or stress (ZHONG et al, 2018; CHEN et al, 2016; ZHANG et al, 2015; TANG et al, 2012; OSTERLUND et al, 2000; ANG et al. 1998) led increase artemisinin biosynthesis if over expressed in *A. annua* L. plants. In this study almost all identified factors (ABF1, APX CCC1, CPK6, JAZ1, MYC2 and JAZ5) with leading under salt stress when compared to *Arabidopsis thaliana* genome used as a reference (Table 1) while AT1G10588.1 (auxin response factor family) and AT1G59750. (Gibberellin-regulated protein family) did not show expression potential under salt stress in *A. annua* L. plants. (Fig. 5, THAKUR, 2010; MIURA and TADA 2014).

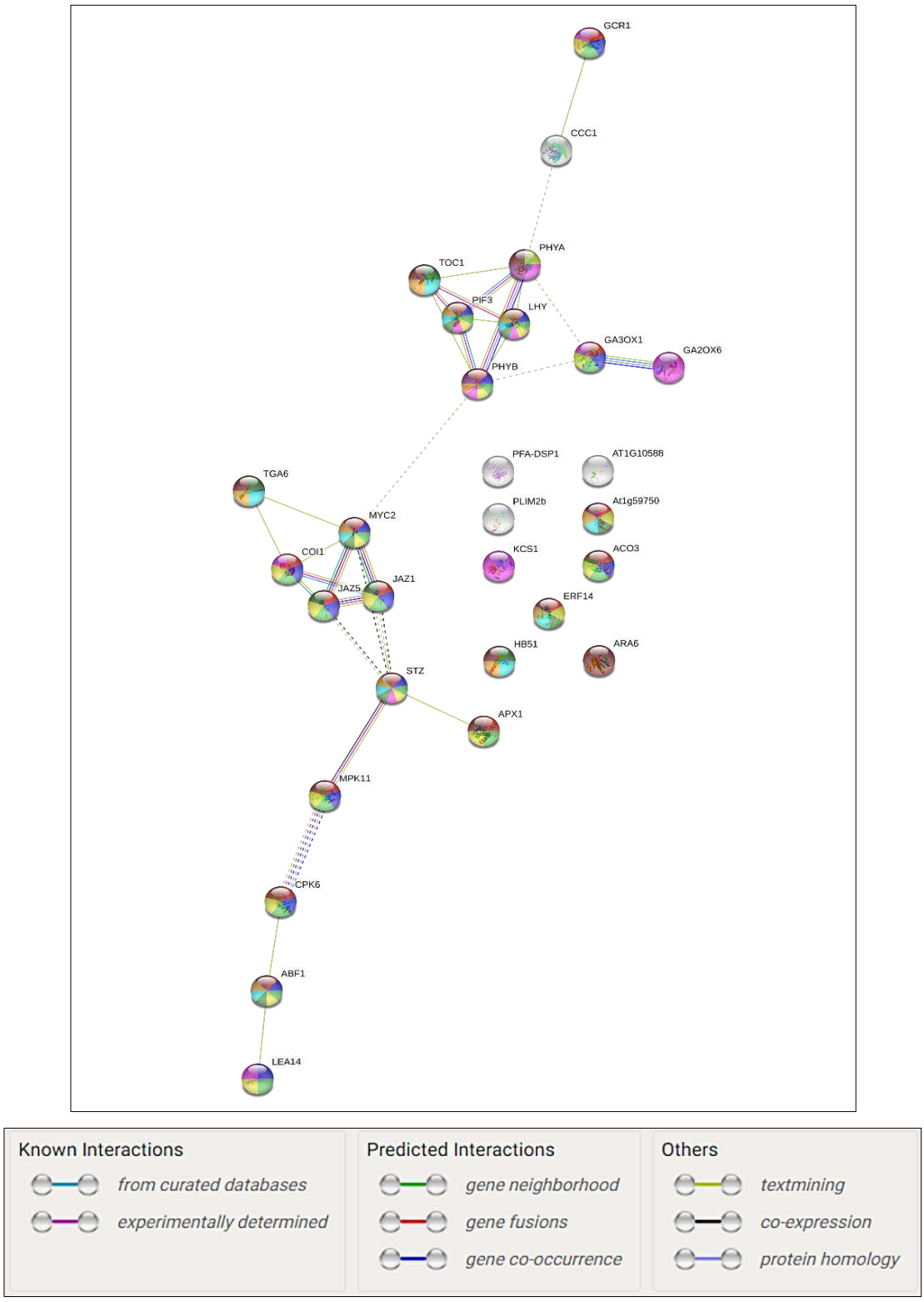

Figure 5. The interaction partners of the significant proteins network i.e. association of gene networking from STRING10.0. The inferences of different colorful outlines between the proteins are presented below the network map.

Most of identified genes from 11 unigenes of *A. annua* L. obtained through ESTs assembly corresponding proteins showed interaction between themselves. Different correlation calculations have been used for categorizing gene expression data. Co-expression analysis, indeed, reveals response of genes to a particular stress or more stresses with some similar consequences. Furthermore, in order to confirm co-expressed genes, their co-regulation might be evaluated for having similar cis-regulatory elements by String 10.0 (SHARMA, 2015; SHINOZALI et al, 2000).

Therefore, to check out assumed functionally related genes with same co-regulation pattern, comparing the results with *A. thaliana* genome data could be useful.

GO analysis enriched the major of identified genes (85.71%%) as responsive genes to stimuli and GO Slim-plant analysis enriched all genes into "response to stress". Salinity and ABA treatment mostly similar regulatory elements and signal transduction pathways with abscisic acid dependent pathways (SHINOZAKI et al, 2000). Major part of the genes showed potential of expression under salt, drought, Jasmonate and ABA treatment as well as similar co-expression pattern with correlation and mutual connection of more than 0.40 (Table S1).

According to the obtained results, all predicted genes for 11 assembled contigs represented acceptable correlation based on Pearson's correlation coefficient (Fig. 5). The co-expression analysis results demonstrated the accuracy of the cis-regulatory elements detection to confirm the functional relationship of proteins encoded with the identified genes in response to particular stress conditions. STRING 10.0 provides user friendly interface by which could easily study proteins of interest based on physical and functional association from various known predicted interactions scattered over databases (SZKLARCZYK et al, 2014). As shown in (Table S2), more than 63% of the identified genes from goProfiler showed interaction with each other in protein level. These data illustrated accordance of co-expression results with their functional property.

PPI network also indicated that proteins encoded by identified genes had interactions with each other in co-expression prediction (Fig. 5). In this study of *A. annua* L. unigene, these unigenes play important role in different abiotic responsive pathways consequence of salt, drought, and temperature stress; plant signaling hormone pathways (ABA-dependent pathway and jasmonate pathway) and diterpenoid pathway in which the central role of ABF1, APX CCC1, CPK6, JAZ1, MYC2 and JAZ5 transcription factors including well characterized (Table 1). The observed transcription factors have also been characterized in many plants and play an important role in cellular activities eg. transcriptional activators, RNA polymerase activation and initiation of transcription and DNA binding transcriptional factor activity under different environmental stress (KVINT et al, 2003; KERK et al, 2003; Fig. 4a-c, Fig. 5, Table S2). On the basis of our analyses we, it has been suggested that the *A. annua* L. plants ESTs based factors associated genes regulates the hormonal regulation activity and elucidate the biosynthetic pathway. Therefore, it is now confirmed that the identified factors predicted from assembled unigenes/contigs using known cis-regulatory elements play an important role in plant yields and metabolites production (XIONG, 2005; RIANO-PACHON, 2007; ZHANG et al, 2005; BRIVANLOU and DARNELL, 2002). In other study, Sharma et al. (2015) *in silico* demonstrated that those genes that are co-expressed under stress share similar cis-regulatory elements are co-regulated with the same regulatory system under abiotic stress conditions. The protective role of these elements has been indicated in response to salt, drought and osmotic stress induced by salt treatments (YOSHIDA et al, 2014; NAKASHIMA et al, 2013 FUJITA et al, 2011).

Conclusions

Stresses influenced the signaling process that affect the common regulatory system and cross-talk in term of molecular changes. These molecular changes in regulatory network of plants affect the biosynthesis of metabolites and ultimately lead to stunt the plant growth. In this study, we made the relationship of predicted unigenes/contigs by ESTs assembly involved in the processes of plant signaling and diterpenoids pathways observed by GO analyses. The assembled unigenes obtained from *Artemisia annua* L. plants ESTs might play important role in plant growth development and artemisinin production. Since co-expression and mutual connection of genes could be a reliable indicator of their involvement in certain environmental factors, these identified unigenes based transcription factor (specially ABF1 JAZ1 and JAZ5 MYC1 CPK6) if over-expressed in *A. annua* L. it may lead to enhanced the production of artemisinin and plant development.

Conflict of Interest

The authors have no conflict of interest to declare.

Acknowledgments/Funding

This project was supported by the Deanship of Scientific Research at Prince Sattam bin Abdulaziz University under the research project #2020/01/16362

References

1. ABDIN MZ, ALAM P (2015). Genetic engineering of artemisinin biosynthesis: prospects to improve its production. *Acta Physiol Plant* 37:33, doi.org/10.1007/s11738-015-1771-5.
2. ALAM P, ABDIN MZ (2011). Over-expression of HMG-CoA reductase and amorpha-4,11-diene synthase genes in *Artemisia annua* L. and its influence on artemisinin content. *Plant Cell Rep* 10:1919-1928.
3. ALAM P, KIRAN U, AHMAD MM, KAMALUDDIN KHAN MA, JHANWAR S, ABDIN MZ (2010). Isolation, characterization and structural features of amorpha – 4,11-diene synthase (ADS_{3963}) from *Artemisia annua* L. *Bioinformation* 4:421-429.
4. ALAM, P, KAMALUDDIN, SHARAF-ELDIN, MA et al. (2016). The effect of over-expression of rate limiting enzymes on the yield of artemisinin in *Artemisia annua*. *Rend. Fis. Acc. Lincei* 27:311-319.
5. ANG LH, CHATTOPADHYAY S, WEI N, OYAMA T, OKADA K, BATSCHAUER A et al. (1998) Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of *Arabidopsis* development. *Mol. Cell* 1:213-222.
6. BRIVANLOU AH, DARNELL JE (2002). Signal transduction and the control of gene expression. *Science*. 295: 813-818.
7. BUSSEMAKER HJ, LI H, SIGGIA ED (2001). Regulatory element detection using correlation with expression. *Nat Genet* 27: 167-171.
8. CARETTO S, LINSALATA V, COLELLA G, MITA G, LATTANZIO V (2015). Carbon Fluxes between Primary Metabolism and Phenolic Pathway in Plant Tissues under Stress. *Int. J. Mol. Sci.* 16: 26378-26394.

9. CHEN XB, YAO QF, GAO XH, JIANG CF, HARBERD NP, FU XD (2016). Shoot-to-root mobile transcription factor HY5 coordinates plant carbon and nitrogen acquisition. *Curr. Biol.* 26:640-646.
10. CHOW CN, LEE TY, HUNG YC, LI GZ, TSENG KC, LIU YH, KUO PL, ZHENG HQ, CHANG WC (2019). PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Res* 47(D1):D1155-D1163.
11. DURAISAMY GS, MISHRA AK, KOCABEK T, MATOUŠEK J (2016). Identification and characterization of promoters and cis-regulatory elements of genes involved in secondary metabolites production in hop (*Humulus lupulus*. L). *Comput. Biol. Chem* 64:346-352.
12. FUJITA Y, FUJITA M, SHINOZAKI K, YAMAGUCHI-SHINOZAKI K (2011). ABA-mediated transcriptional regulation in response to osmotic stress in plants. *Journal of plant research* 124: 509-525.
13. HAGGAG WH, ABOUZIENA HF, ABD-EL-KREEM F, EL HABBASHA S (2015). Agriculture biotechnology for management of multiple biotic and abiotic environmental stress in crops. *J. Chem. Pharm*. Res 7, 882-889.
14. HUANG X, MADAN A (1999*).* CAP3: A DNA sequence assembly program. *Genome* Res. 9:868-77.
15. JENKS MA, HASEGAWA PM (2005). Plant Abiotic Stress. (Blackwell, Oxford, 2005).
16. JOSHI R, WANI SH, SINGH B, BOHRA A, DAR ZA, LONE AA, PAREEK A, SINGLA-PAREEK SL (2016). Transcription Factors and Plants Response to Drought Stress: Current Understanding and Future Directions. *Front Plant Sci* 7:1029. doi: 10.3389/fpls.2016.01029
17. KERK D, BULGRIEN J, SMITH DW, GRIBSKOV M (2003). *Arabidopsis* proteins containing similarity to the universal stress protein domain of bacteria. *Plant Physiol* 131: 1209-1219.
18. KVINT K, NACHIN L, DIEZ A, NYSTROM (2003). The bacterial universal stress protein: function and regulation. *Current opinion in microbiology* 6: 140-145.
19. LESCOT M, DÉHAIS P, THIJS G, MARCHAL K, MOREAU Y, VAN DE PEER Y, ROUZÉ P, ROMBAUTS S (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 30:325-7.
20. MA D, PU G, LEI C, MA L, WANG H, GUO Y, CHEN J, DU Z, LI G, YE H, LIU B (2009b). Isolation and characterization of AaWRKY1, an *Artemisia annua* transcription factor that regulates the amorpha- 4,11-diene synthase gene, a key gene of artemisinin biosynthesis. *Plant Cell Physiol*. 50:2146-2161.
21. MIURA K, TADA Y (2014). Regulation of water, salinity, and cold stress responses by salicylic acid. *Frontiers in plant science*. 5: 4. doi.org/10.3389/fpls.2014.00004
22. NAKASHIMA K, YAMAGUCHI-SHINOZAKI K (2013) ABA signaling in stress-response and seed development. *Plant Cell Rep* 32: 959-970.
23. OSAKABE Y, OSAKABE K, SHINOZAKI K AND TRAN L-SP (2014). Response of plants to water stress. *Front. Plant Sci* 5:86. doi: 10.3389/fpls.2014.00086
24. OSTERLUND MT, HARDTKE CS, WEI N, DENG XW (2000). Targeted destabilization of HY5 during light-regulated development of *Arabidopsis*. *Nature* 405:462-466.
25. PADDON CJ, WESTFALL PJ, PITERA DJ, BENJAMIN K, FISHER K, MCPHEE D et al (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496:528-532.
26. PELEG Z, APSE MP, BLUMWALD E (2011). Engineering salinity and water-stress tolerance in crop plants: getting closer to the field. *Adv. Bot. Res* 57, 405-443.
27. PEPLOW M (2016). Synthetic biology's first malaria drug meets market resistance. *Nature* 530:389-390.
28. RAMAKRISHNA A, RAVISHANKAR GA (2011). Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signal Behav*. 6, 1720-1731.
29. RIANO-PACHON DM (2007). PlnTFDB: an integrative plant transcription factor database. *BMC bioinformatics* 8: 42. doi.org/10.1186/1471-2105-8-42
30. SCHLÜTER U, COLMSEE C, SCHOLZ U, BRÄUTIGAM A, WEBER AP, ZELLERHOFF N, BUCHER M, FAHNENSTICH H, SONNEWALD U (2013). Adaptation of maize source leaf metabolism to stress related disturbances in carbon, nitrogen and phosphorus balance. *BMC Genomics* 14:442. doi.org/10.1186/1471-2164-14-442
31. SHARMA A (2015). In silico identification of regulatory motifs in co-expressed genes under osmotic stress representing their co-regulation. *Plant Gene*. 1: 29-34.
32. SHEN Q, HUANG H, ZHAO Y, XIE L, HE Q, ZHONG Y, WANG Y, WANG Y, TANG K (2019). The Transcription Factor Aabzip9 Positively Regulates the Biosynthesis of Artemisinin in *Artemisia annua*. *Front Plant Sci*. 10:1294. doi:10.3389/fpls.2019.01294
33. SHESHADRI SA, NISHANTH MJ, SIMON B (2016). Stress-Mediated cis-Element Transcription Factor Interactions Interconnecting Primary and Specialized Metabolism in planta. *Front Plant Sci.* 7:1725.
34. SHINOZAKI K, YAMAGUCHI-SHINOZAKI K (2000). Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Current opinion in plant biology* 3: 217-223.
35. SZKLARCZYK D, FRANCESCHINI A, WYDER S, FORSLUND K, HELLER D, HUERTA-CEPAS J et al. (2014). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids res* 43: 447-452.
36. TANG N, ZHANG H, LI XH, XIAO JH, XIONG LZ (2012). Constitutive activation of transcription factor OsbZIP46 improves drought tolerance in rice. *Plant Physiol* 158:1755-1768.
37. THAKUR P (2010). Cold stress effects on reproductive development in grain crops: an overview. *Environmental and Experimental Botany* 67: 429-443.
38. XIONG Y (2005). Transcription factors in rice: a genome-wide comparative analysis between monocots and eudicots. *Plant mol. biol* 59: 191-203.
39. YOSHIDA T, MOGAMI J, YAMAGUCHI-SHINOZAKI K (2014). ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. *Current opinion in plant biology* 21: 133-139.
40. ZHANG F, FU X, LV Z, LU X, SHEN Q, ZHANG L et al. (2015). A basic leucine zipper transcription factor, AabZIP1, connects abscisic acid signaling with artemisinin biosynthesis in *Artemisia annua*. *Mol. Plant* 8:163-175.
41. ZHANG W, RUAN J, HO TH, YOU Y, YU T, QUATRANO RS (2005). Cis-regulatory element based targeted gene finding: genome-wide identification of abscisic acid-and abiotic stress-responsive genes in Arabidopsis thaliana. *Bioinformatics* 21: 3074-3081
42. ZHONG Y, LI L, HAO X, FU X, MA Y, XIE L et al. (2018). AaABF3, an abscisic acid–responsive transcription factor, positively regulates artemisinin biosynthesis in *Artemisia annua*. *Front. Plant Sci* 9:1777. doi: 10.3389/fpls.2018.01777

TableS1: co-expression of 11 ESTs interactions in abiotic stress signallings

Interactions				
node1	**node2**	**node1 accession**	**node2 accession**	**score**
ABF1	CPK6	AT1G49720.2	AT2G17290.1	0.547
ABF1	LEA14	AT1G49720.2	AT1G01470.1	0.518
APX1	STZ	AT1G07890.8	AT1G27730.1	0.58
CCC1	GCR1	AT1G30450.2	AT1G48270.1	0.483
CCC1	PHYA	AT1G30450.2	AT1G09570.1	0.41
COI1	JAZ1	AT2G39940.1	AT1G19180.1	0.998
COI1	JAZ5	AT2G39940.1	AT1G17380.1	0.839
COI1	MYC2	AT2G39940.1	AT1G32640.1	0.968
COI1	TGA6	AT2G39940.1	AT3G12250.4	0.461
CPK6	ABF1	AT2G17290.1	AT1G49720.2	0.547
CPK6	MPK11	AT2G17290.1	AT1G01560.2	0.416
GA2OX6	GA3OX1	AT1G02400.1	AT1G15550.1	0.928
GA3OX1	GA2OX6	AT1G15550.1	AT1G02400.1	0.928
GA3OX1	PHYA	AT1G15550.1	AT1G09570.1	0.523
GA3OX1	PHYB	AT1G15550.1	AT2G18790.1	0.624
GCR1	CCC1	AT1G48270.1	AT1G30450.2	0.483
JAZ1	COI1	AT1G19180.1	AT2G39940.1	0.998
JAZ1	JAZ5	AT1G19180.1	AT1G17380.1	0.935
JAZ1	MYC2	AT1G19180.1	AT1G32640.1	0.997
JAZ1	STZ	AT1G19180.1	AT1G27730.1	0.791
JAZ5	COI1	AT1G17380.1	AT2G39940.1	0.839
JAZ5	JAZ1	AT1G17380.1	AT1G19180.1	0.935
JAZ5	MYC2	AT1G17380.1	AT1G32640.1	0.971
JAZ5	STZ	AT1G17380.1	AT1G27730.1	0.41
LEA14	ABF1	AT1G01470.1	AT1G49720.2	0.518
LHY	PHYA	AT1G01060.2	AT1G09570.1	0.782
LHY	PHYB	AT1G01060.2	AT2G18790.1	0.834
LHY	PIF3	AT1G01060.2	AT1G09530.2	0.665
LHY	TOC1	AT1G01060.2	AT5G61380.1	0.995
MPK11	CPK6	AT1G01560.2	AT2G17290.1	0.416
MPK11	STZ	AT1G01560.2	AT1G27730.1	0.417
MYC2	COI1	AT1G32640.1	AT2G39940.1	0.968
MYC2	JAZ1	AT1G32640.1	AT1G19180.1	0.997
MYC2	JAZ5	AT1G32640.1	AT1G17380.1	0.971
MYC2	PHYB	AT1G32640.1	AT2G18790.1	0.463
MYC2	STZ	AT1G32640.1	AT1G27730.1	0.635
MYC2	TGA6	AT1G32640.1	AT3G12250.4	0.522
PHYA	CCC1	AT1G09570.1	AT1G30450.2	0.41
PHYA	GA3OX1	AT1G09570.1	AT1G15550.1	0.523
PHYA	LHY	AT1G09570.1	AT1G01060.2	0.782
PHYA	PHYB	AT1G09570.1	AT2G18790.1	0.781
PHYA	PIF3	AT1G09570.1	AT1G09530.2	0.998
PHYA	TOC1	AT1G09570.1	AT5G61380.1	0.795
PHYB	GA3OX1	AT2G18790.1	AT1G15550.1	0.624
PHYB	LHY	AT2G18790.1	AT1G01060.2	0.834
PHYB	MYC2	AT2G18790.1	AT1G32640.1	0.463
PHYB	PHYA	AT2G18790.1	AT1G09570.1	0.781
PHYB	PIF3	AT2G18790.1	AT1G09530.2	0.999
PHYB	TOC1	AT2G18790.1	AT5G61380.1	0.832
PIF3	LHY	AT1G09530.2	AT1G01060.2	0.665
PIF3	PHYA	AT1G09530.2	AT1G09570.1	0.998
PIF3	PHYB	AT1G09530.2	AT2G18790.1	0.999
PIF3	TOC1	AT1G09530.2	AT5G61380.1	0.91
STZ	APX1	AT1G27730.1	AT1G07890.8	0.58
STZ	JAZ1	AT1G27730.1	AT1G19180.1	0.791
STZ	JAZ5	AT1G27730.1	AT1G17380.1	0.41
STZ	MPK11	AT1G27730.1	AT1G01560.2	0.417
STZ	MYC2	AT1G27730.1	AT1G32640.1	0.635
TGA6	COI1	AT3G12250.4	AT2G39940.1	0.461
TGA6	MYC2	AT3G12250.4	AT1G32640.1	0.522
TOC1	LHY	AT5G61380.1	AT1G01060.2	0.995
TOC1	PHYA	AT5G61380.1	AT1G09570.1	0.795
TOC1	PHYB	AT5G61380.1	AT2G18790.1	0.832
TOC1	PIF3	AT5G61380.1	AT1G09530.2	0.91
Total				**65**

Table S2: Artemisia annua L. EST contigs involved in different Biological process

Source	Term_ name	Term_ id	Adjusted_p_value	Negative_log10_of_ adjusted_p_value	Term_ size	Query_ size	Intersection_ size	Effective_ domain_size	Intersections
GO:MF	DNA-binding transcription factor activity	GO:0003700	0.048997503	1.309826048	1634	23	8	20815	"AT1G49720.1,AT5G03790.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT3G12250.1,AT1G59750.1,AT1G04370.1"
GO:BP	response to hormone	GO:0009725	1.16E-07	6.936801322	1858	23	14	22991	"AT1G49720.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT1G07890.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to endogenous stimulus	GO:0009719	1.50E-07	6.822830639	1895	23	14	22991	"AT1G49720.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT1G07890.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to acid chemical	GO:0001101	6.28E-07	6.201992171	1338	23	12	22991	"AT1G49720.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT1G01470.1,AT2G17290.1,AT1G15550.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to chemical	GO:0042221	1.05759E-06	5.97568246	3203	23	16	22991	"AT1G49720.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT3G12250.1,AT1G07890.1,AT1G01470.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to organic substance	GO:0010033	1.10979E-06	5.954758443	2204	23	14	22991	"AT1G49720.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT1G07890.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to oxygen-containing compound	GO:1901700	1.14646E-06	5.94064165	1785	23	13	22991	"AT1G49720.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT1G07890.1,AT1G01470.1,AT2G17290.1,AT1G15550.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to lipid	GO:0033993	4.20496E-05	4.376238029	862	23	9	22991	"AT1G49720.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT2G17290.1,AT1G15550.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to stimulus	GO:0050896	5.85787E-05	4.232260367	6521	23	19	22991	"AT1G49720.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT3G12250.1,AT1G07890.1,AT1G01470.1,AT1G01120.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G02400.1,AT1G48270.1,AT1G01560.2"
GO:BP	hormone-mediated signaling pathway	GO:0009755	7.60765E-05	4.118749685	924	23	9	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1"
GO:BP	cellular response to hormone stimulus	GO:0032870	0.000224864	3.648079812	1050	23	9	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1"
GO:BP	response to light stimulus	GO:0009416	0.000240227	3.61937837	754	23	8	22991	"AT1G09530.1,AT1G01060.1,AT1G27730.1,AT1G01470.1,AT1G01120.1,AT1G15550.1,AT1G02400.1,AT1G48270.1"
GO:BP	response to radiation	GO:0009314	0.000301691	3.520438356	777	23	8	22991	"AT1G09530.1,AT1G01060.1,AT1G27730.1,AT1G01470.1,AT1G01120.1,AT1G15550.1,AT1G02400.1,AT1G48270.1"
GO:BP	cellular response to endogenous stimulus	GO:0071495	0.000305754	3.514627674	1089	23	9	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1"
GO:BP	cellular response to chemical stimulus	GO:0070887	0.000464428	3.333081169	1518	23	10	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT1G07890.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1"
GO:BP	cellular response to organic substance	GO:0071310	0.000647609	3.18868723	1191	23	9	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT1G48270.1"
GO:BP	signal transduction	GO:0007165	0.000777896	3.109078332	2045	23	11	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G48270.1,AT1G01560.2"
GO:BP	cellular response to acid chemical	GO:0071229	0.000798042	3.097974508	600	23	7	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G15550.1,AT1G48270.1"
GO:BP	signaling	GO:0023052	0.000947917	3.023229885	2086	23	11	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to abscisic acid	GO:0009737	0.00123854	2.907089943	641	23	7	22991	"AT1G49720.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT2G17290.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to alcohol	GO:0097305	0.001290722	2.889167267	645	23	7	22991	"AT1G49720.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT2G17290.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to abiotic stimulus	GO:0009628	0.001352113	2.868986876	2162	23	11	22991	"AT1G09530.1,AT1G01060.1,AT1G27730.1,AT2G05710.1,AT1G19180.1,AT1G07890.1,AT1G01470.1,AT1G01120.1,AT1G15550.1,AT1G02400.1,AT1G48270.1"
GO:BP	cell communication	GO:0007154	0.002886871	2.539572632	2335	23	11	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G48270.1,AT1G01560.2"
GO:BP	cellular response to oxygen-containing compound	GO:1901701	0.005729766	2.241863141	809	23	7	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT2G17290.1,AT1G15550.1,AT1G48270.1"
GO:BP	biological regulation	GO:0065007	0.007985836	2.097679619	6776	23	17	22991	"AT1G49720.1,AT5G03790.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT1G17380.1,AT1G19180.1,AT3G12250.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G02400.1,AT1G48270.1,AT1G01560.2,AT1G30450.1"
GO:BP	cellular response to stimulus	GO:0051716	0.009547462	2.020112062	3211	23	12	22991	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT1G07890.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to gibberellin	GO:0009739	0.010876545	1.963509037	161	23	4	22991	"AT1G09530.1,AT1G01060.1,AT1G15550.1,AT1G48270.1"
GO:BP	regulation of cellular process	GO:0050794	0.011024477	1.957641991	5289	23	15	22991	"AT1G49720.1,AT5G03790.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT1G17380.1,AT1G19180.1,AT3G12250.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G48270.1,AT1G01560.2"
GO:BP	response to wounding	GO:0009611	0.026945217	1.569518317	203	23	4	22991	"AT1G27730.1,AT1G17380.1,AT1G19180.1,AT1G01470.1"
GO:BP	gibberellic acid mediated signaling pathway	GO:0009740	0.040192429	1.395855751	80	23	3	22991	"AT1G09530.1,AT1G15550.1,AT1G48270.1"

GO:BP	gibberellin mediated signaling pathway	GO:0010476	0.041711076	1.37974861	81	23	3	22991	"AT1G09530.1,AT1G15550.1,AT1G48270.1"
GO:BP	cellular response to gibberellin stimulus	GO:0071370	0.046491195	1.332629287	84	23	3	22991	"AT1G09530.1,AT1G15550.1,AT1G48270.1"
GO:BP	regulation of biological process	GO:0050789	0.049742196	1.303275048	5957	23	15	22991	"AT1G49720.1,AT5G03790.1,AT1G09530.1,AT1G01060.1,AT1G27730.1,AT1G17380.1,AT1G19180.1,AT3G12250.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G48270.1,AT1G01560.2"
GO:CC	intracellular part	GO:0044424	0.019838978	1.702480706	13817	22	22	20312	"AT1G49720.1,AT5G03790.1,AT1G09530.1,AT1G01780.1,AT1G01060.1,AT1G27730.1,AT1G05000.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT3G12250.1,AT1G07890.1,AT1G01470.1,AT1G01120.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G48270.1,AT1G01560.2,AT1G30450.1"
GO:CC	intracellular	GO:0005622	0.026630365	1.574622875	14003	22	22	20312	"AT1G49720.1,AT5G03790.1,AT1G09530.1,AT1G01780.1,AT1G01060.1,AT1G27730.1,AT1G05000.1,AT2G05710.1,AT1G17380.1,AT1G19180.1,AT3G12250.1,AT1G07890.1,AT1G01470.1,AT1G01120.1,AT2G17290.1,AT1G59750.1,AT1G15550.1,AT1G04370.1,AT3G54840.1,AT1G48270.1,AT1G01560.2,AT1G30450.1"
KEGG	Plant hormone signal transduction	KEGG:04075	0.000820171	3.086095731	273	13	6	4851	"AT1G49720.1,AT1G09530.1,AT1G17380.1,AT1G19180.1,AT3G12250.1,AT1G59750.1"
KEGG	Diterpenoid biosynthesis	KEGG:00904	0.03302247	1.481190451	22	13	2	4851	"AT1G15550.1,AT1G02400.1"