COMPARATIVE OMICS ANALYSIS FOR SOME YIELD-RELATED TRAITS ASSOCIATED WITH SALT TOLERANCE IN SUNFLOWER (*Helianthus annuus* L.).

F. M. ABD EL TWAB[1], A. ABO DOMA[1], M. H. AMAR[2] AND M. S. RIZK[1,2]

1. *Genetics Department, Faculty of Agriculture, Ain Shams University (ASU), Cairo, Egypt.*
2. *Egyptian Deserts Gene Bank, Desert Research Center (DRC), Egypt.*

Sunflower (*Helianthus annuus* L.) remain one of the essential producers of edible oil over the world, and consider an excellent source of unsaturated fatty acids, necessary protein, fibers and essential nutrients like vitamins such as (A, D and E) (El-Wakil, 2014). The wild sunflowers genus *Helianthus* is a diploid species (2n = 2x = 34), are branched annuals and perennials of family *Asteraceae* originated in North America (Bukhsh *et al*., 2011; Aboki *et al*., 2012). Egypt is facing a considerable reduction in the production of edible oil and imports annually about 93% to supplement the needs for oil to meet the increasing consumption. Therefore, it is representing that the best choice for increasing local production of edible oils to overcome the shortage of its production (Kandil *et al*., 2016). Apparently, classification of sunflower genotypes based on morphological characteristics were associated with the environmental changes, making classifications based on morphological traits unreliable. Hence, with the aid of molecular identification of sunflower, it is becoming easier and faster to evaluation the genetic diversity and structure of plant populations using both morphological characters and molecular markers (Zhang *et al*., 2015). Therefore, selection by molecular markers is particularly a useful approach to improve the genetics of yield-related traits and to enhance the development of marker-assisted selection (MAS) methods. Amplified Fragment Length Polymorphism (AFLP) is a molecular technique with high resolution, efficiency, reproducibility, and no prior sequence information for the identification of genetic diversity of plants is needed (Mueller and Wolfenbarger, 1999). AFLP is bi-allelic, dominant and they typically cover the entire genome though they sometimes tend to cluster around centromeres (Manel *et al*., 2010).

Another system to make AFLP bands relatively clear was by limiting homoplasy, via *in-silico* AFLP analysis (Zhang *et al*., 2014). Ii is a computational tool to evaluate theoretical results of PCR amplicons amplified from sequenced genome or transcriptome (Boutros and Okey, 2004). *In-silico* PCR analysis was helpful to test the efficiency of PCR marker techniques and identify potential genes with the genome coverage (Alsamman *et al*., 2017). Amongst AFLP protocol, fluorescent dye labeling fluorescent (F-AFLP) protocol was the result of an optimization process using a capillary

sequencer to achieve high polymorphism, reproducibility, and sensitivity ease of genotyping together with a wide abroad of fragment resolution (Wolf *et al*., 2004; Yang *et al*., 2016a). F-AFLP also was known as a genome scan for identifying marker loci that are linked to selectively relevant target loci through 'genetic hitchhiking' (Luikart *et al*., 2003). Genome scans are widely used to detect signatures of local adaptation to environmental conditions among plant populations, especially for species whose genomes have not been sequenced (Yang *et al*., 2016a). It can potentially be discovered by a strong correlation between allele frequencies and environmental parameters (Coop *et al*., 2010). Recently, an increased emphasis has been placed on developing AFLP-based genome scans to revel outlier loci linked to of floral divergence (Herrera and Bazaga, 2008); host plants (Manel *et al*., 2010); insecticide resistance (Paris *et al*., 2010) ecotype divergence (Yang *et al*., 2016a); adaptation to altitude (Yang *et al*., 2016b); or selection during domestication (Rossi *et al*., 2009).

Salinity is one of the problems that restrict production of crops in arid and semi-arid areas. Egypt is one of the countries that have been profoundly affected by salinity, and a large area of agricultural lands is now rated as unproductive as a result of salinity (Athar and Ashraf, 2009). Quantitative real-time polymerase chain reaction (qRT-PCR) is a necessary technology used to analyze gene-expression levels and become a high-throughput method for accurate gene-expression profiling (Zhang *et al*., 2016). qRT-PCR can be used to calculate low gene copy number differences in abundances of a template and low gene expression levels (Lu *et al*., 2012). Intracellular signaling via calcium-dependent protein kinases (*CDPKs*) and pyrroline-5-carboxylate synthetase (*P5CS*) were as two significant pathways that are involved in the signaling of abiotic and biotic stresses in plant cells and widely used to react to a changing environment (Mehlmer *et al*., 2010). As is well known *CDPKs* plays essential roles in cell signaling and plant growth, development, and response to environmental cues and crucial sensors of Ca^{2+} flux resulting from salt stress, as well as, signal transduction of calcium within and out of cell vacuole (Cai *et al*., 2015). However, proline accumulation is an essential mechanism for osmotic regulation under salt stress, *P5CS* and glutamate dehydrogenase activities, the key enzymes in the glutamate pathway of proline biosynthesis, and up-regulated *P5CS* expression, but it decreased the activity of proline dehydrogenase (Pro DH), a key enzyme of proline degradation (Yang *et al*., 2016a). In the current investigation, we sought to determine a new basis for the ongoing discussion about; first, to assess the yield-related traits with an AFLP genome scan to identify candidate loci by selection along an population structure, secondly, discovering genome coverage using *in-silico* PCR analysis and finally, screening the sunflower genotypes under investigation for their salt tolerance and select the most promising genotypes in salt tolerance

and evaluation of gene expression using qRT-PCR.

MATERIALS AND METHODS

Sampling collection

Fourteen sunflower genotypes were kindly obtained from various sources, Oil Crops Research Department, Field Crops Res. Institute (FCRI), Agricultural Research Center (ARC), Giza, Egypt, and The National Gene Bank of Egypt (NGB), their names and sources are presented in (Table 1).

Morphological traits

A field experiment was conducted to evaluate fourteen sunflower genotypes for their morphological, yield and yield components traits. The studied yield related traits are; plant height (cm), number of leaves/plan, stem diameter (cm), head diameter (cm), head weight (g), seeds weight (g), number of seeds/head and 1000 seeds weight (g).

DNA extraction

Genomic DNA was extracted using Gene JET™ Plant Genomic DNA Purification Kit (#K0791, Thermo Scientific, Lithuania). The quality and concentration of the DNA samples were checked in a Quawell Q5000 UV-Vis spectrophotometer (V2.1.4, USA). A portion of the DNA was diluted to 50 ng/µl for use in F-AFLP analysis. Both the stock and diluted portions were kept at -20°C.

Testing DNA contamination

Testing DNA for possible Eukaryotic DNA contamination was performed by amplifying the internal transcribed spacer ITS region according to (White *et al*., 1990). Samples showed an only a single band at the expected length (~600 bp) were considered for further analysis.

F-AFLP protocol

The original protocol of Vos *et al*. (1995) was applied, the primers were fluorescently rather than radioactively labeled as a modification (Table 2). Six different selective PCR combinations were successfully applied to 1٤ specimens. PCR amplification was performed using the original protocol without modifications. Private Service was contracted to visualize the amplified products using ABI3730 DNA analyzer (Applied Biosystems, USA) with a size standard GS500-LIZ (Macrogen Fragment Analysis Service, Korea).

In silico PCR analysis

The overall genome length of anchored scaffolds in the merged assembly was 2.45 Gb about 68% of the 3.6 Gb sunflower genome, was used as template for the *in-silico* PCR analysis against all studied of six AFLP primer combinations, Table (2) to reveal the possible PCR amplimers. In home PERL scripts were used for performing the *in-silico* PCR analysis and detect adjoin genes. For the primer genome coverage statics, the overlap layout consensus algorithm was used

to report each primer-covering area. This algorithm ensures that if two PCR amplimers share same genomic areas, only the collective area will be reported, and redundant area will be removed. The same algorithm was used for genome sequence assembly. For calculating the total coverage for each PCR marker type, all *in-silico* PCR amplimers belonging to each marker type were processed collectively. The same procedure was conducted in calculating the total primer set coverage. Circos software package (Circos 0.66) was used for visualizing output results in a circular layout.

Greenhouse salt tolerance screening

The rationale for this greenhouse screening was to determine salt tolerance sunflower genotypes under controlled conditions through vegetative stages. Seven sunflower genotypes were screened during three replications. At the seventh to tenth leaf stages, approximately two weeks after planting three salt treatments were used: 0 ppm (control), 4000 ppm and 6000 ppm NaCl the salt treatment was initiated. Data recorded were number of leaves, plant height, leaf length, root length and fresh and dry plant weight.

RNA isolation

Total RNA was extracted using plant RNA extraction kit (Zomanbio, Beijing, China) according to the manufacturer's instructions, and treated with RNase-free DNase I (Takara, Dalian, China). Equivalent amount of total RNA was quantified using Quawell Q5000 UV-Vis spectrophotometer (V2.1.4, USA). The reverse transcription reactions were conducted by almost 2 microgram total RNA as template and with the kit from Takara, which synthesizes first- and second-strand cDNAs using random hexamer and oligo-dT primers.

Real time PCR

Real-time PCR was performed with Fast SYBR Green® LO- Rox master mix (Bio Line, UK). Reactions were run on a Strata gene Mx3000p (Agilent Technologies, Santa Clara, CA). Conditions were: 10 min at 95°C, 40 cycles of 5s at 95°C, 45s at 60°C, 20s at 72°C. PCR products were examined by melt curve analysis. The reaction was achieved with the primers (Table 3), each reaction was repeated three times for each cDNA sample. Actin gene was used as a reference gene for data normalization. The Ct values that were used for further analysis were obtained using the MxPro Mx3000P v3.00 software (Agilent Technologies, Santa Clara, CA).

Data analysis

The collected data were statistically analyzed using CoStat program and correlations between eight yield related-traits were estimated. The differences among means were compared using Duncan's new multiple ranged test (Duncan, 1955). The results were considered significant at $P<0.01$. While, Regression analysis was performed to define the dependency of 1000 seeds weight on other correlated variables.

Molecular data

Loci scoring

Automated fluorescent AFLP scoring was performed using two programs Peak Scanner™ (Applied Biosystems, USA) for peak calling and RawGeno V2 for automated scoring, according to the software's manuals. Loci with rare frequency were cut-off (i.e., a frequency below 10% and/or above 90%).

Population and genetic structure

Based on the F-AFLP, a graphic demonstration of phylogenetic tree was constructed using Kinship relationships (KSR) according to VanRaden methods. To gain further perspectives on the population genetic structure among the sunflower germplasm, Principal Coordinates Analysis (PCoA) was carried out to display the multidimensional genetic relationship and its partition among varieties across the binary matrix using TotalLab120 software (Schlueter and Harris, 2006), taxonomic group was stated prior.

Real-time PCR data analysais

The most commonly used method for relative quantification is the $2^{-\Delta\Delta Ct}$ method. The relative difference in gene expression using the $2^{-\Delta\Delta Ct}$ method was calculated as follows: Relative fold change in gene expression = $2^{-\Delta\Delta Ct}$, where: $\Delta\Delta_{Ct} = \Delta_{Ct}$ control – Δ_{Ct} treatment, and Δ_{Ct} = (Ct genes – Ct actin gene) as shown in Table (3).

RESULTS AND DISCUSSION

Correlation among yields–related traits

Correlations between eight yields–related traits were evaluated among 14 genotypes of sunflower (Table 4). The results indicated positive correlations between head diameter and the four traits including; stem diameter, head weight, seeds weight and number of seeds per head, which recorded 0.6147, 0.7247, 0.6025 and 0.5453, respectively. A significant correlation was observed between head weight and both of seeds weight (0.8177) and 1000 seeds weight (0.6816). However, seed weight/plant recorded highly significant and positive correlation with 1000 seeds weight (0.7640).

Molecular characterization

Amplification scoring and readability

PCR amplification was successful for six pairs of primers combination as exposed in the Table (7). Peak analyses and automated band scoring were successful, and quality tests showed adequate quality. A total of 1007 amplicons were scored with number of amplicons for each primer pair ranged between 50-659 bp. Out of 1007 amplicons, 981 were polymorphic amplicons (PA) accounting for 97.42% with the minimum and maximum values of polymorphism percentage per combination was 119 and 208, respectively.

Markers associated with yield related traits

Eight traits across the six combination of F-AFLP were used to detect the marker-traits associations in sunflower genotypes. In details, only 73 MAS were generated within the studied traits (plant height, number of leaves, stem diameter, head diameter, head weight, seeds weight, number of seeds per head and 1000 seeds weight) ranged from 52 to 682 bp in size, while the P-value for the association between F-AFLP markers and the aforementioned sunflower traits ranged from 2 to 3.66 under ($p \leq 0.01$), (data not shown). The highest number of MAS was observed with the following four traits, plant height, 1000 seeds weight, head diameter and number of leaves/Plant accounting for 19, 16, 11 and 11 molecular markers, respectively. While, the lowest numbers of MAS were shown with the following three traits, head weight, number of seeds per head and seed weight accounting for 2, 4 and 4, molecular markers, respectively.

In silico PCR analysis as revealed by the F-AFLP markers

In the present study six primer combinations of F-AFLP were used to cover the genome scan of *Helianthus annuus* across the 17 chromosomes. The results indicated that *in- silico* analysis created a total of 11671 amplicons with coverage of 1378 genes. Out of 2.45 Gb, only 1865440 bp was identified across the six *in-silico* primer combinations with the percentage of the total chromosomal being 0.0512% as shown in Table (6) and Fig. (3). The highest number of band count was observed in C1 primer combination, while the lowest number was scored in C4 with (2882 and 601), respectively. Furthermore, the highest and the lowest number of genome coverage were presented in C6 and C4 primers combination being 473315 and 94821 bp, respectively. On behalf to the gene count, C5 and C4 primer combinations scored the highest and the lowest number of gene count being 322 and 100 genes, respectively. Chr13 was the highest covered (0.055%), while, Chr7 was the lowest (0.044%) using all AFLP combinations. On the other hand, C4 combination gave the lowest genome coverage (0.0026%), while C6 gave the highest genome coverage with 0.0129%. However, Fig. (4) illustrate the percentage of individual chromosome coverage across the six combinations. Chromosomes number 1, 5, 16 and 9 showed the highest percentages of chromosomes coverage with the percentages of 10%, 8%, 8% and 8%, respectively. Whereas, chromosomes number 11 and 15 revealed the lowest coverage percentages of 1%.

Detection of positive selection signatures (outlier's)

The AFLP genome scan analysis revealed 73 loci under selection among a total of 1007 loci scored in this study. The 73 outlier's loci were under directional selection due to genetic variation assumed to be related to seed production. Combination C3 showed the maximum number of

outlier's loci (24 loci), followed by C6 (22 loci), C2 (18 loci), C1 (5 loci), and finally the pair of primer combinations C4 and C5 exhibited the minority values of outlier loci (2 loci). In conclusion, the difference between the actual covered genomic areas and the total area covered by the combined data was about 75 kb. This may suggested that these markers are located in different genomic areas compared to the total genome scan.

Assessment of genetic variation using F-AFLP markers

Based on the F-AFLP, a graphic demonstration of phylogenetic tree was constructed using Kinship relationships (KSR) according to VanRaden methods (Fig. 6). Indeed, some differences in the positioning among the fourteen *Helianthus annuus*. L genotypes were observed in the consensus trees. In general, the results revealed that three strongly clades, supported that resolved the species into distinct branches among the fourteen *Helianthus annuus* L. genotypes. Out of the fourteen genotypes, twelve were sorted together into three clades, while the two genotypes Luxor13 and Luxor14 were clustered jointly as out-group in the basal position of the tree. In details, the first clade comprised jointly two sub-clads, the first one included New Valley4 is close to New Valley5. While, AlBuhayrah3 is close to Qena2 in the second sub-clad, however, Sakha53* is a common sister of all other *Helianthus annuus* L. genotypes in the first clade. In the second clade Sohag10 and Luxor11 were close together as a sister to Luxor12. Meanwhile, the third clade represented Aswan is closely related to Bani Sweif, at the same time being sister to Sohag9 and Qena6.

To further determination of the genetic relationships among the fourteen genotypes of *Helianthus annuus* L. pomegranate, a graphic demonstration of principal component analysis (PCoA) was presented (Fig. 7). The results indicated that it could be divided into three major categories. Category I comprised Sakha53*, Qena2, Al Buhayrah, New Vally4 and New Vally5 close together in the first group. Category II assembled Sohag10, Luxor11 and Luxor12. However, Qena6, Sohag9, Bani Sweif, and Aswan were closely related in Category III. Meanwhile, Luxor13 and Luxor14 were clustered alone in separate group and distantly related to the three main groups of the two axes of PCoA. Interestly, Sakha53* was isolated separately outside all of *Helianthus annuus* L. genotypes. In conclusion, KSR and PCoA results were highly identical, with slight differences in the position among the fourteen genotypes of *Helianthus annuus* L.

Assessment of salt stress in sunflower genotypes

Salt tolerance in the seven sunflower genotypes was assessed using the following yield related traits (number of leaves, plant height, root length and whole plant dry weight). Table (9) presented the overall reduction ratios across the four traits. The results indicated that the differences among the seven sunflower geno-

types were significant with the salt concentrations 4000 and 6000 ppm of NaCl, in number of leaves, plant height, root length and whole plant dry weight. S3 genotype recorded the lowest reduction ratio with a rank of 1.5, while, S10 genotype was the highest reduction ratio with a rank of 7. Therefore, it possible to conclude that Al-Buhyrah genotype was selectively the most salt tolerant genotype while, Sohag genotype was the most sensitive one.

Gene expression profile of CDPKs and P5CS in response to salt stress

Quantitative real time PCR was carried out to validate and confirm the mechanism of salt tolerance among the most tolerant and sensitive genotypes (Al Buhayrah) and (Sohag), respectively. The results of expression patterns in response to salt tolerance revealed significant increase in the expression level of *CDPKs* and *P5CS*. The expression of *CDPKs* in the salt tolerant genotype (Al Buhayrah) was up regulated in response to salt stress with 2.464 folds over the control. On the other hand, in the salt sensitive genotype (Sohag) the gene expression was down regulated from 1.39 in control to 0.38 in treated plants with 0.273 folds below the control. While, the results of expression level of *P5CS* indicated that in the salt tolerant genotype (Al Buhayrah) the expression was up regulated in response to salt stress with 2.715 folds over the control. On the other hand, in salt sensitive genotype (Sohag) the gene expression was down regulated from 1.78 in the control to 0.29 in treated plants with 0.163 folds of the control as shown in Fig. (8 a,b) and Table (8).

Despite the vast number of existing sunflower cultivars used for seeds and oil production, there is continuing need to develop a new variety with resistant tolerances to further expand into subtropical climates change in Egypt; increased seed quality and appearance are being targeted by breeders worldwide. Seed yield and its related components and adaptability have priority in breeding programs. Apparently, classification of sunflower genotypes based on morphological characteristics were associated with the environmental changes, making classifications based on morphological traits unreliable. Hence, with the aid of molecular identification of sunflower, it is becoming easier and faster to characterize germplasm and identify genotypes with desirable traits in breeding programs. To meet these challenges, a mixture of morphological yield-related traits and molecular markers is considered as the optimal to find the informative markers-assisted selection (MAS) and is particularly a useful approach in sunflower breeding strategy (Zhang *et al*., 2015). Hence, the statistical analysis that estimates both environmental and genetic components of variability need to elucidate the genetics of these traits fully. F-AFLP markers is the powerful model of choice and are accessible for performing MAS for several species, due to quality control of the AFLP procedures (Manel *et al*., 2010). Quantitative trait loci which were detected could be used as MAS for

the investigated traits. The markers with highest significance may have the highest contribution and therefore can be used for developing markers for MAS. The detected accessions included in this study have shown to harbor significant levels of morphological and genetic variation. Our results provided a reliable and discriminant set of markers for the cost-effective molecular characterization of sunflower collection under investigation and offer the guidelines for the delineation of sampling strategies for yield components, thus aiding the efficient management and exploitation of these valuable genetic resources (Dong *et al*., 2007).

Another way to make F-AFLP bands relatively clear *via in-silico* AFLP analysis in model species (Zhang *et al*., 2014). *In-silico* PCR analysis can assist in the selection of newly designed primers and avoid potential problems before primer synthesis. The regions matching the two primers should be localized on the selected sequence in a way allowing PCR amplification, which forces the relative orientation of the matches and the distance between them. To simulate real PCR conditions, the *in-silico* PCR algorithm should allow some mismatches between the primers and the target sequences (Ficetola *et al*., 2010). In subsequent studies, Paris *et al*. (2010) evaluated *in-silico* AFLP analysis in model species, it was concluded that homoplasy sharply increases with the number of peaks per AFLP outline and that the number of co-migrating fragments in a single peak depends on the repetitiveness of the genome. This technique has advantage that could improve primer selectivity needed to target genes and the average of genome coverage. Certainly, the comparison between the results of *in-silico* and in vitro PCR highlights the usefulness of the *in-silico* PCR for assessing the genetic variability, and MAS in crop improvement **(**Mokhtar *et al*., 2016). In the interpretation of Alsamman *et al*. (2017), *in-silico* PCR is the optimal approach to compare between the efficiency of different molecular marker techniques and helped to improve the analysis of F-AFLP and provide more informative and reliable results.

Accordingly, discover the locus-specific under divergent choice is a serious step to considerate the evolutionary process for the population genetic variations that affect fitness in diverse environments, which can possibly be detected by a high correlation between allele frequencies and ecological parameters (Beaumont and Balding, 2004; Roesti *et al*., 2012). In the present investigation, it has demonstrated that the 73 outlier locus can be distinguished explicitly in the genes under selection and also in neutral flanking regions due to hitchhiking effects (Wang *et al*., (2012). According to Feng *et al*. (2015) and Yang *et al*. (2016b), such a result refers to outlier AFLP loci are likely not the target of natural selection, but the neighboring genes of these loci might involve in significant economic trials that are difficult to select by phenotype early in the plant life cycle (Byrne *et al*., 2012). It is worth mentioning that, the outlier loci must be confirmed through further analy-

sis, including sequencing and molecular functional analysis of neighboring genes.

To identify the population's structure, we used Kinship analysis and PCoA to resolve patterns of genetic variation and assess genetic distances among groups of populations of sunflower genotypes (Fig. 8). Our findings demonstrated that PCoA and KSR analysis produced propensity to four sub-population and KSR analysis appeared a strong effect in identifying migrants in clustered samples by geographical region. Currently, a comprehensive analysis was conducted for the behavior of F-AFLP using (KSR) and (PCoA) to maximize the accuracy of the results. These findings agreed with a previous study by Jie *et al*. (2003), Dong *et al*. (2007) and Jannatdoust *et al*. (2016).

Salt stress causes reduction in yield components of sunflower varieties with multiple thresholds at salt stress, some are salt-sensitive glycophytes, meanwhile, others are moderately or highly salt tolerant halophytes (Sadak *et al*., 2012). Therefore, gene expression analysis is an outstanding approach, providing insight into the genetic and developmental mechanisms in biological research. Calcium-dependent protein kinases (*CDPKs*) are important sensors of Ca^{++} flux in plants, plays essential roles in plant development and adaptation to various abiotic stresses. In subsequence study, Roche *et al*. (2009), reported that expression levels of some protein kinase, and phosphatase genes along with other transcription factors were up-regulated in leaves of tolerant genotypes of sunflower compared to the non-tolerant genotype. A subset of *CDPK* genes in this family exhibits inducible expression patterns and have been previously demonstrated to play pivotal roles in plant response to abiotic stresses such as salinity (Vivek *et al*., 2013; Campo *et al*., 2014). While that, Bagdi *et al*. (2015) reported that increasing of NaCl concentration from 85 mM to 425 mM progressively increased the proline content in rice, while, the maximum increase in proline content was recorded at 425 mM NaCl concentration as compared to control. This accumulation in the proline content was coupled to significant induction of the gene encoding *P5CS* approximately, 1.35 folds under salt stress conditions. These findings are in partial agreement with our results indicated that expression profiling of the two salt tolerances related genes (*CDPKS* and *P5CS*) were induced in the salt tolerance genotype (Al-Buhyrah) but not in the sensitive genotype (Sohag). Interestingly, the mechanism of *CDPKS* and *P5CS*, beneficial effects on agronomic traits and abiotic stresses, possibly through transcriptional regulation of the key genes in salt tolerance pathways. The cross-talk between agronomic traits and gene expression analysis indicated that the improved salt tolerance in the distinguished genotypes may be attributed to the gene induction of related pathways. Plants are subjected to various stress conditions including salt during their life cycles that may adversely affect their productivity, and adaptation to environmental changes is an important

fitness trait for crop development (Ewas *et al*., 2017 a,b). In the interpretation of Skz *et al*. (2018) the KEGG pathway enrichment analysis identified significantly enriched metabolic pathways or signal transduction pathways associated with differentially expressed genes compared with the whole genome background. Finally, the data from our study can be utilized to better understand the gene expression networks involved in sunflower and other oil plant under salt stress and could provide tools to maximize the benefits for crop production.

SUMMARY

In the present investigation, we sought to assess the genetic diversity among fourteen Egyptian genotypes of *Helianthus annuus* using the fluorescence-based AFLP (F-AFLP) and evaluation of yield-related traits. The genome coverage was assessed using *in-silico* PCR analysis and the real-time PCR technique (qPCR) was used to determine gene expression of some salt-related genes. The analysis of variance for yield-related traits indicated that there were significant differences among genotypes for all studied traits ($p \leq 0.01$). A total of 1007 amplicons were scored across six F-AFLPs combinations with 97.42% polymorphic percentage. The results of *in-silico* analysis revealed that F-AFLPs exhibited the highest chromosomal coverage (0.055%) and targeted 1378 genes with the average area (75kb) covered by the F-AFLPs. Out of 1007 F-AFLP loci, only 73 outlier loci were identified as a putatively positive outlier to seeds weight loci. PCoA and KSR analysis produced propensity to four sub-population and KSR analysis appeared a strong effect in identifying migrants in clustered samples by geographical region. In the context, seven genotypes out of the fourteen were tested under greenhouse conditions on different salt stress concentrations at 4000 and 6000 ppm NaCl. We can point out that, Sohag genotype was considered as the most sensitive genotype; while Al-Buhyrah genotype was considered as the most salt tolerant one based on the combined results of the investigated yield related traits. To explorer the genetic factors behind the improved salt tolerance in selected sunflower genotypes, gene expressions of *CDPKs* and *P5CS* genes were estimated using real time qRT-PCR and showed substantial differentiation between the two aforementioned contrasting genotypes regarding salt tolerance.

REFERENCES

Aboki, M. A., M. Mohammed, S. H. Musa, B. S. Zuru, H. M. Aliyu, M. Gero, I. M. Alibe and B. Inuwa (2012). Physicochemical and antimicrobial properties of sunflower (*Helianthus annuus* L.) seed oil. Int. J. Sci. Tech., 2: 151-154.

Alsamman, A. M., S. S. Adawy, S. D. Ibrahim, B. A. Hussein and E. H. A. Hussein (2017). Selective Amplification of Start codon Polymor-

phic Loci (SASPL): a new PCR-based molecular marker in olive. Plant Omics, 10: 64-77.

Athar, H. R. and M. Ashraf (2009). Strategies for crop improvement against salinity and drought stress: An overview. Salinity and Water Stress, p: 1-16.

Bagdi, D. L., B. P. Shaw, B. B. Sahu, G. K. Purohit (2015). Real time PCR expression analysis of gene encoding P5CS enzyme and proline metabolism under NaCl salinity in rice. J. Environ Biol., 36: 955-961.

Beaumont, M. A. and D. J. Balding (2004). Identifying adaptive genetic divergence among populations from genome scans. Molecular Ecology, 13: 969-980.

Boutros, P. C. and A. B. Okey (2004). PUNS: transcriptomic-and genomic *in-silico* PCR for enhanced primer design. Bioinformatics, 20: 2399-2400.

Bukhsh, M. A., J. Iqbal, S. Kaleem, A. Wasaya and M. Ishaque (2011). Qualitative analysis of spring planted sunflower hybrids as influenced by varying nutritional area. Pakistan Journal of Nutrition, 103: 291-295.

Byrne, D. H., M. B. Raseira, D. Bassi, M. C. Piagnani, K. Gasic, G. L. Reighard, M. A. Moreno and S. Pérez (2012). Peach. In Fruit Breeding, p: 505-569.

Cai, H., J. Cheng, Y. Yan, Z. Xiao, J. Li, S. Mou, A. Qiu, Y. Lai and D. S. H. Guan (2015). Genome-wide identification and expression analysis of calcium-dependent protein kinase and its closely related kinase genes in *Capsicum annuum*. Front. Plant Sci., 6: 737.

Campo, S., P. Baldrich, J. Messeguer, E. Lalanne, M. Coca and B. SanSegundo (2014). Over expression of a calcium-dependent protein kinase confers salt and drought tolerance in rice by preventing membrane lipid peroxidation. Plant Physiol., 165: 688-704.

Coop, G., D. Witonsky, A. Di Rienzo and J. K. Pritchard (2010). Using environmental correlations to identify loci underlying local adaptation. Genetics, 185: 1411-1423.

Dong, G., G. Liu and K. Li (2007). Studying Genetic Diversity in the Core Germplasm of Confectionary sunflower (*Helianthus annuus* L.) in China Based on AFLP and Morphological Analysis. Russian Journal of Genetics, 43: 627-635.

Duncan, D. B. (1955). Multiple ranges and multiple F test. Biometrics, 11: 1.

El-Wakil, D. A. (2014). Seed-Borne fungi of sunflower (*Helianthus annuus*

L.) and their impact on oil quality. IOSR Journal of Agriculture and Veterinary Science, 6: 38-44.

Ewas, M., E. Khames, K. Ziaf, R. Shahzad, E. Nishawy, F. Ali and J. Luo (2017a). The Tomato DOF Daily Fluctuaions 1, TDDF1 acts as flowering accelerator and protector against various stresses. Nature Sci. Rep., 7: 1-16.

Ewas, M., Y. Gao, F. Ali, E. M. Nishawy, R. Shahzad, H. Subthain, M. H. Amar, C. Martin and J. Luo (2017b). RNA-seq reveals mechanisms of SlMX1 for enhanced carotenoids and terpenoids accumulation along with stress resistance in tomato. Science Bull., 62: 476-485.

Feng, X. J., G. F. Jiang and Z. Fan (2015). Identification of outliers in a genomic scan for selection along environmental gradients in the bamboo locust, Ceracris kiangsu. Scientific Reports, 5: 13758.

Ficetola, G. F., E. Coissac, S. Zundel, T. Riaz, W. Shehzad, J. Bessière and F. Pompanon (2010). An *in silico* approach for the evaluation of DNA barcodes. BMC genomics, 11: 434.

Herrera, C. M. and P. Bazaga (2008). Population-genomic approach reveals adaptive floral divergence in discrete populations of a hawk moth-pollinated violet. Molecular Ecology, 17: 5378-5390.

Jannatdoust, M., R. Darvishzadeh, R. Ziaeifard, M. A. Ebrahimi, H. H. Maleki, E. Gholinezhad and A. A. Hatamnia (2016). Analysis of genetic diversity and population structure of confectionery sunflower (*Helianthus annuus* L.) native to Iran Journal of Crop Science and Biotechnology, 19: 37-44.

Jie, L., L. Gong-she and C. C. Jan (2003). Comparison of genetic diversity of the germplasm resources of confectionary sunflower *Helianthus annuus* in China based on RAPDs and AFLPs. Acta Botanica Sinica., 45: 352-358.

Kandil, A. A., A. E. Sharief and A. M. Amira (2016). Seedling parameters of some sunflower cultivars as affected by seed priming and salinity stress. Int. J. Agro. Agri. Res., 9: 117-136.

Lu Y., L. Xie and J. Chen (2012). A novel procedure for absolute real- time quantification of gene expression patterns. Plant Methods, 8: 1-11.

Luikart, G., P. R. England, D. Tallmon, S. Jordan and P. Taberlet (2003). The power and promise of population genomics: from genotyping to genome typing. Nature Reviews Genetics, 4: 981-994.

Manel, S., S. Joost, B. K. Epperson, R. Holderegger, A. Storfer, M. S. Rosenberg, K. T. Scribner, A. Bonin and M. J. Fortin (2010). Per-

spectives on the use of landscape genetics to detect genetic adaptive variation in the field. Molecular Ecology, 19: 3760-3772.

Mehlmer, N., B. Wurzinger, S. Stael, D. Hofmann-Rodrigues, E. saszar, B. Pfister and M. Teige (2010). The Ca^{2+} dependent protein kinase CPK3 is required for MAPK independent salt-stress acclimation in Arabidopsis. The Plant Journal, 63: 484-498.

Mokhtar, M. M., S. S. Adawy, S. S. El-Assal and E. H. A. Hussein (2016). Genetic, intergenic SSR database generation, SNPs determination and pathway annotations, in date palm (*Phoenix dactylifera* L.). PLOS ONE, 11: 1-19.

Mueller, U. G. and L. L. Wolfenbarger (1999). AFLP genotyping and finger printing. Trends in Ecology and Evolution, 14: 389-394.

Paris, M., S. Boyer, A. Bonin, A. Collado, J. P. David and L. Despres (2010). Genome scan in the mosquito *Aedes rusticus*: population structure and detection of positive selection after insecticide treatment. Molecular Ecology, 19: 325-337.

Roche, J., T. Hewezi, A. Bouniols and L. Gentzbittel (2009). Real- time PCR monitoring of signal transduction related genes involved in water stress tolerance mechanism of sunflower. Plant Physiology and Biochemistry, 47: 139-145.

Roesti, M., W. Salzburger and D. Berner (2012). Uninformative polymorphisms bias genome scans for signatures of selection. BMC Evolutionary Biology, 12: 94.

Rossi, M., E. Bitocchi, E. Bellucci, L. Nanni, D. Rau, G. Attene and R. Papa (2009). Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. Evolutionary Applications, 2: 504-522.

Sadak, M. S., A. A. A. El-Monem, H. M. S. El-Bassiouny and N. M. Badr (2012). Physiological response of sunflower (*Helianthus annuus* L.) to exogenous arginine and putrescine treatments under salinity stress. Journal of Applied Sciences Research, 8: 4943-4957.

Schlueter, P. M. and S. A. Harris (2006). Analysis of multilocus fingerprinting data sets containing missing data. Molecular Ecology Notes, 6: 569-572.

Skz, A., S. Vardharajula and S. S. K. P. Vurukonda (2018). Transcriptomic profiling of maize (*Zea mays* L.) seedlings in response to Pseudomonas putida stain FBKV2 inoculation under drought stress. Annals of Microbiology, 68: 331-349.

Vivek, P. J., N. Tuteja and E. V. Soniya (2013). CDPK1 from ginger pro-

motes salinity and drought stress tolerance without yield penalty by improving growth and photosynthesis in *Nicotiana tabacum*. PLoS One, 8: 1-15.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. V. Lee, M. Hornes, A. Friters, J. Pot, J. Paleman, M. Kuiper and M. Zabeau (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research, 23: 4407-4414.

Wang, T., G. Chen, Q. Zan, C. Wang ad Y. Su (2012). AFLP genome scan to detect genetic structure and candidate loci under selection for local adaptation of the invasive weed *Mikania micrantha*. PLoS One, 7: e41310.

White, T. J., T. Bruns, S. Lee and J. W. Taylor (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In PCR Protocols: A guide to methods and applications edited by: Innis, M. A., Gelfand, D. H., Sninsky J. J., White T. J. New York: Academic Press Inc., 315-322.

Wolf, P., B. Doche, L. Gielly and P. Taberlet (2004). Genetic structure of Rhododendron ferrugineum at a wide range of spatial scales. J. Hered., 95: 301-308.

Yang, S., N. Guo and H. Ge (2016a). Morphological and AFLP-based genetic diversity in *Rosa platyacantha* population in eastern *Tianshan mountains* of northwestern China. Horticultural Plant Journal, 2: 55-60.

Yang, A. H., N. Wei, P. W. Fritsch and X. H. Yao (2016b). AFLP genome scanning reveals divergent selection in natural populations of Liriodendron Chinense (*magnoliaceae*) along a latitudinal transect. Frontiers in Plant Science, 7: 698.

Zhang, Z., F. R. D. van Parijs and B. Xiao (2014).The status of AFLP in the genomics era and a pipeline. Mol. Breeding, 34: 1245-1260.

Zhang, M. F., Q. Liu and G. X. Jia (2016). Reference gene selection for gene expression studies in lily using quantitative real-time PCR. Genetics and Molecular Research: GMR, 15: 1-14.

Zhang, Q., R. Jia, C. Meng, C. Ti and Y. Wang (2015). Diversity and population structure of a dominant deciduous tree based on morphological and genetic data. Aob Plants. Research Article, 1-13.

Table (1): Names of the fourteen sunflower genotypes and their sources.

Code No.	Common Name	Source
S1	Sakha53*	OCRD –FCRI-ARC
S2	Qena	(NGB)- ARC
S3	Al Buhayrah	(NGB)- ARC
S4	The new Valley	(NGB)- ARC
S5	The new Valley	(NGB)- ARC
S6	Qena	(NGB)- ARC
S7	Bani Sweif	(NGB)- ARC
S8	Aswan	(NGB)- ARC
S9	Sohag	(NGB)- ARC
S10	Sohag	(NGB)- ARC
S11	Luxor	(NGB)- ARC
S12	Luxor	(NGB)- ARC
S13	Luxor	(NGB)- ARC
S14	Luxor	(NGB)- ARC

Sakha*53 Local variety, obtained from Oil Crops Research Department, Field Crops Res., Inst., Agricultural Res. Center, Giza, Egypt, while genotypes from S2 to S14 are accessions from the National Gene Bank, Agricultural Research Center, Giza, Egypt.

Table (2): List of oligonucleotides (primer and/or adaptors) used for F-AFLP amplification.

Primer/ Adaptor	5´sequence 3´
EcoRI - A1	CTCGTAGACTGCGTACC
EcoRI - A2	AATTGGTACGCAGTC
Mse I - A1	GACGATGAGTCCTGAG
Mse I - A2	TACTCAGGACTCAT
Eco + A	GACTGCGTACCAATTC**A**
Mse + C	GATGAGTCCTGAGTAA**C**
Eco + ACA	FAM-GACTGCGTACCAATTC**ACA**
Eco + AGG	HEX-GACTGCGTACCAATTC**AGG**
Eco + ATA	CY3-GACTGCGTACCAATTC**ATA**
Mse + CAA	GATGAGTCCTGAGTAA**CAA**
Mse + CTC	GATGAGTCCTGAGTAA**CTC**

Table (3): List of qRTPCR primer sequences, gene name and Actin primer sequence.

Gene Name	F	R
$CDPK_s$	5´ ATG ATG ACA AAA GAG CAG CC´3	5´ CTT GAT ACC CGT TCC ATT TG´3
P_5Cs	5´ GGC AAT GGA ACT GAA GGC´3	5´ CCT CTT CCC ATC CTT GAC TTA´3
Actin	5´GCC GTG CTT TCT CTT TAT GC´3	5´ CTC TCT GGA GGA GCA ACC AC´3

Table (4): Correlation between the eight yields–related traits of sunflower.

Traits	Plant height (cm)	No. of leaves	Head diameter (cm)	Stem diameter (cm)	Head weight (g.)	1000 Seeds weight (g.)	No. of seeds per head
No. of Leaves	-0.1042						
Head diameter (cm)	0.3487	-0.2833					
Stem diameter (cm)	0.2508	0.2672	0.6147*				
Head weight (g.)	0.2513	-0.2191	0.7247*	0.2161			
Seeds weight (g.)	0.0933	0.1443	0.6025*	0.3263	0.8177*		
No. of seeds per head	-0.0495	0.1912	0.5453*	0.5084	0.4029	0.3505	
1000 seeds weight (g.)	0.0802	-0.1322	0.3409	-0.0254	0.6816*	0.7640*	-0.2477

*= Correlation significant test at ($p \leq 0.05$).

Table (5): Summary of the polymorphic, monomorphic amplicons, % of polymorphic and positive unique marker monomorphic amplicons by six F-AFLP combinations.

Combination name	Monomorphic amplicons (MA)	Polymorphic amplicons (PA)	Total number of amplicons	Percentage of monomorphic amplicons	Percentage of polymorphism (PA%)	Total of positive unique marker
C6-yellow	1	207	208	0.48%	99.52%	43
C3-yellow	14	208	222	6.31%	93.69%	31
C2-Green	8	155	163	4.91%	95.09%	19
C4-blue	2	166	168	1.19%	98.81%	27
C1-Blue	0	126	126	0.00%	100.00%	34
C5-green	1	119	120	0.83%	99.17%	24
Total	26	981	1007	2.58%	97.42%	178

Table (6): Illustrated the band count, genome coverage and genes count as revealed by six F-AFLP combinations.

Combination name	Number of Band count	Genome coverage (bp)	Number of genes count
C5	2801	337965	322
C3	2039	391858	260
C1	2882	424924	320
C4	601	94821	100
C6	1945	473315	208
C2	1403	218439	168

Table (7): The overall ranks of the seven genotypes of sunflower across the four yield-related traits based on reduction ratio.

Traits Genotypes	L/P	P.H	R.L	D.W	Overall ranks
S_1	4	5	5	5	5.0
S_2	5	5	3	5	4.5
S_3	2	1	1	2	1.5
S_4	4	4	4	2	3.5
S_5	3	2	5	2	3.0
S_{10}	7	7	7	7	7.0
S_{14}	4	4	4	6	4.5

Table (8): Transcript levels of the CDPKs and P°Cs gene in response to the expression of two sunflower genotypes using real-time PCR analysis (R T PCR).

Gene name	Treatments	Cultivars	
		S3 Al Buhayrah (Tolerant)	S10 (Sohag) (Sensitive)
CDPKs	Control	2.760	1.390
	Treated	6.800	0.380
	Folds	2.464	0.273
P5Cs	Control	3.050	1.780
	Treated	8.280	0.290
	Folds	2.715	0.163

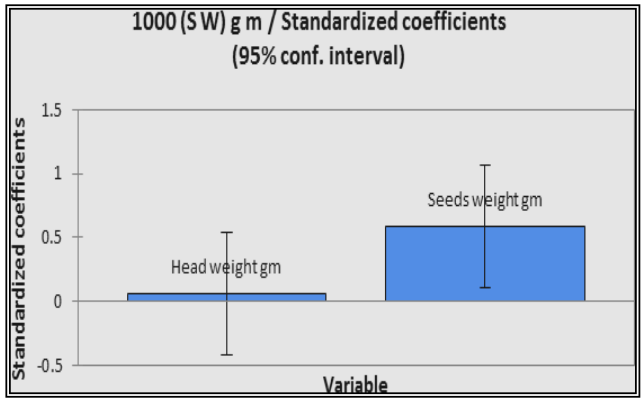


Fig. (1): Effect of independent variable seeds weight and head weight on 1000 seeds weight.

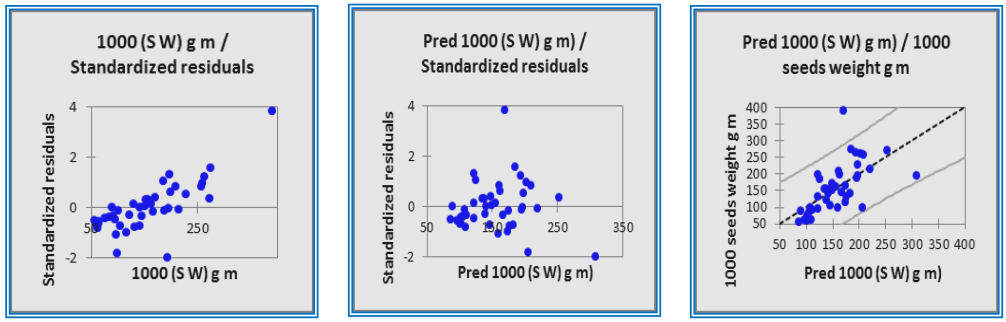
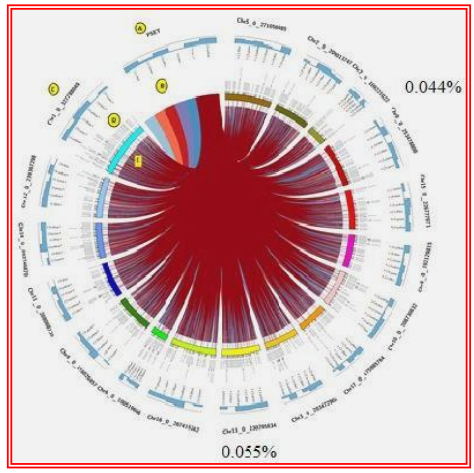


Fig. (2): Observed and predicted records based on the linear regression model.

Fig. (3): The schematic represented *in-silico* PCR analysis of six F-AFLP combinations while, (A) against sunflower genome, (B) revealing the possible adjoined genes, (D) primer total genome coverage percentages statics (A), each primer total chromosome coverage percentages statics (C) and the position of possible PCR amplimers with genes (extended lines) or without (short lines) (E).



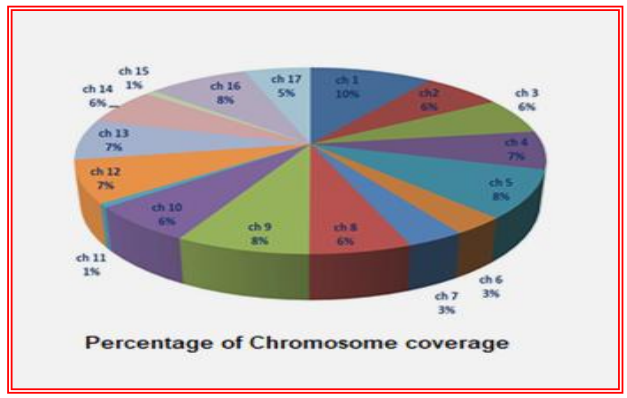


Fig. (4): Percentage of set chromosomal coverage on the sunflower genome through *in-silico* PCR analysis by F-AFLP combination primers.

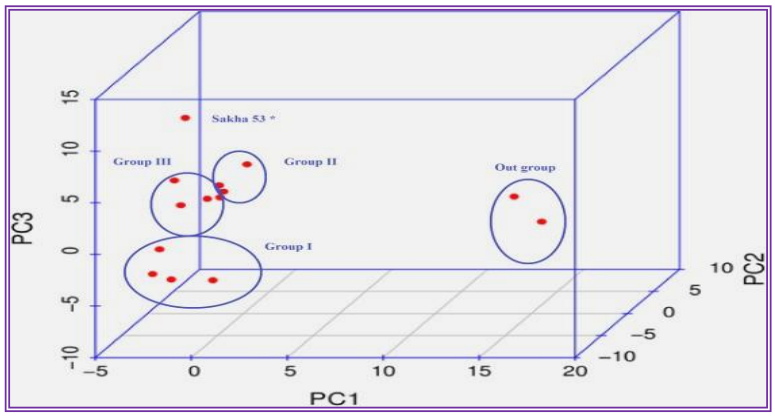


Fig. (5): Kinship relationships among 14 genotypes of *Helianthus annuus* L. according to Van Raden method (2008).

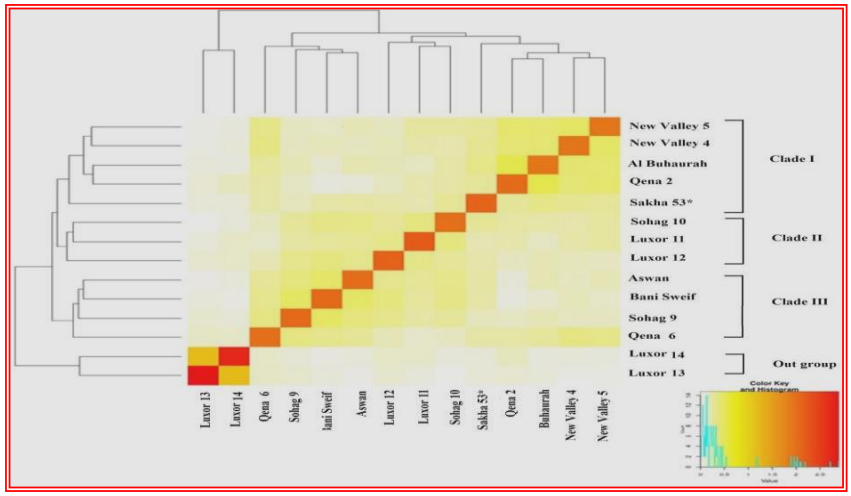


Fig. (6): Principal component analysis (PCA) among 14 genotypes of *Helianthus annuus* L.

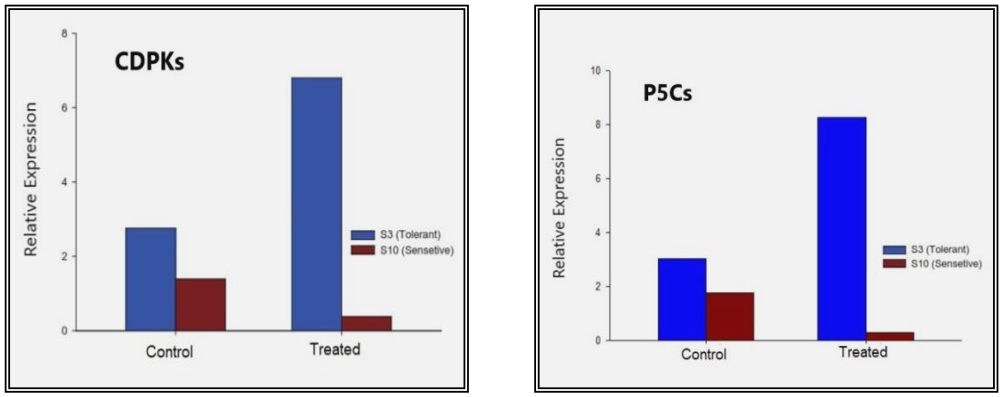


Fig. (7): Diagram of transcript levels of the *CDPK2* and *P5Cs* genes in response to salt tolerance in the two contrasting genotypes of sunflower using real-time Panalysis.

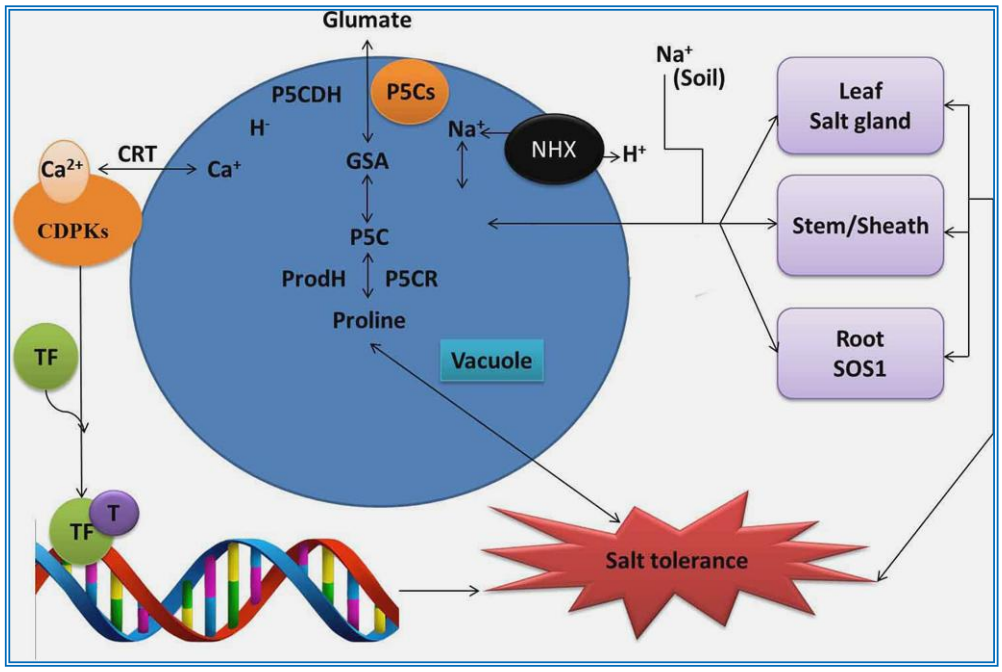


Fig (8): Schematic KEGG pathway analysis representation of *CDPK2* and *P5Cs*, mechanisms for improved salt tolerance in sunflower crop, according to Ewas *et al*. (2017a).