Bulletin of Biological and Allied Sciences Research
ISSN: **2521-0092**
www.bbasr.org
DOI: https://doi.org/10.64013/bbasr.v2026i1.110
*Bull. Biol. All. Sci. Res., Volume, **11**: 110*
Original Research Article ***Open Access***

GENOME-WIDE CHARACTERIZATION AND EXPRESSION ANALYSIS OF THE BCCP GENE FAMILY IN SOYBEAN: IMPLICATIONS FOR FATTY ACID BIOSYNTHESIS UNDER SALT STRESS AND MELATONIN TREATMENT

RASHEED MU[1(†)], MALIK A[1(†)], TUFAIL MT, SAMI A[2(†)], HAIDER MZ[1], ALI Q[1], JAVED MA[1], ALI D[3]

[1]Department of Plant Breeding and Genetics, University of the Punjab, P.O BOX. 54590, Lahore, Pakistan
[2]Agricultural Biotechnology Research Center, National Cheng Hsing University, Academia Sinica, Taipei, Taiwan
[3]Department of Zoology, College of Science, King Saud University, PO Box 2455, Riyadh 11451, Saudi Arabia
(†) authors have equal contribution
**Correspondence Author Email Address: ur36402@gmail.com*

(Received, 24th May 2024, Accepted 20th December 2025, Published 1st January 2026)

Abstract *Soybean oil is a complex mixture of triacylglycerols rich in unsaturated fatty acids which are major sources of lipids for human and animal consumption. The biosynthesis of soybean oil, an important agricultural commodity, involves a complex network of enzymes and regulatory pathways. One of the major pathways, BCCP domain of Acetyl-CoA Carboxylase (ACC) is considered as a vital key role player for this process but its role in soybeans remained poorly understood. In this study, the identified and characterized 11 BCCP genes in the soybean genome (GmBCCP) revealed their diverse roles in fatty acid biosynthesis and stress response. One of the 4 subunits of Acetyl-CoA Carboxyl (ACC), BCCP (Biotin Carboxyl Carrier Protein) domain have great importance for fatty acid biosynthesis in plants, impacting the oil content and its composition. To evaluate the composition, function, and structure of this domain in soybean, this study was conducted to explore their gene structures, conserved motifs, chromosomal distribution, physiochemical properties, evolutionary relationships, and expression patterns under melatonin and salt stress. Phylogenetic analysis revealed the evolutionary conservation of GmBCCP genes across diverse oilseed crops, particularly with Arabidopsis. Subcellular localization highlighted their diverse roles in key metabolic compartments of cells. Cis-regulatory element analysis showed the potential functions in stress responses, growth, development, and hormone signaling. The distinction in intron-exon structures indicates potential regulatory complexity through alternative splicing. Gene duplication analysis identified segmental and tandem duplications contributing to the expansion and functional diversification of the GmBCCP gene family. RNA-seq data demonstrated that melatonin treatment upregulated GmBCCP1 and GmBCCP10 expression, potentially stimulating fatty acid biosynthesis. Additionally, melatonin lessened the salt-induced downregulation of GmBCCP2, suggesting its role in stress tolerance. These findings provide a comprehensive groundwork for understanding the functional roles of GmBCCP genes in soybean, with implications for crop improvement strategies for enhancing oil production and stress resilient soybean cultivars. These insights provide a foundation for targeted breeding strategies aimed at enhancing oil yield and developing stress-tolerant soybean cultivars.*

[Citation: Rasheed, M.U., Malik, A., Tufail, M.T., Sami, A., Haider, M.Z., Ali, Q., Javed, M.A., Ali, D. (2026). Genome-wide characterization and expression analysis of the bccp gene family in soybean: implications for fatty acid biosynthesis under salt stress and melatonin treatment. *Bull. Biol. All. Sci. Res.* **11***: 110. doi:* https://doi.org/10.64013/bbasr.v2026i1.110]
Keywords: *BCCP genes; Soybean (Glycine max); Fatty acid biosynthesis; Melatonin; Gene expression*

Introduction

Soybean (*Glycine Max L.*), being a leguminous crop, is a major and most economic source of plant protein and oil. Due to its great dietary and nutritionary value, and its various products like soy milk, soya sauce, soya oil and soy beverages etc., Soybean has a great impact on the world's economy. About 5000 years ago, from wild species of soybean in China *G. soja seib. & Zuccc*., the cultivated soybean is domesticated by initial natural selection and then artificial selection of desirable traits of higher yield, larger inflorescence structures, larger seeds, greater seed weight and other positive heterotic agronomic traits (Liu et al., 2020). Currently, An EST database, full-length cDNAs and cDNA microarrays, and a haplotype map (GmHapMap) are among the additional datasets that have been created. These genomic resources provide up a world of possibilities for studying gene function using different forward and reverse genetic techniques, improving soybeans through marker-assisted breeding, transgenic and genome editing, and other methods (Xu et al., 2022).

Pakistan's reliance on imported soybean oil has increased significantly over the last decade, which is highlighting the domestic production gap. (FAO_Stat,

2020-2022). Need of soyabean oil in Pakistan is increasing day by day but climate change, limited availability of high oil yielding cultivars and inadequate pest and disease management has adversely affected the yield from 2020-2022 (Siddiqui et al., 2012). Oil accumulation is highly affected by salt quantity in soil. In Pakistan, 7 million hectares of agricultural land is saline which is also a major cause of such decline in soya oil production (FAO.org, 2017).

Acetyl-CoA Carboxylase (ACC) is a multi-subunit enzyme that catalyzes the first committed step in de-novo fatty acid biosynthesis. It is composed of several functional domains, including the Biotin Carboxyl Carrier Protein (BCCP) subunit, which is responsible for transferring carboxyl groups and is indispensable for the enzyme's overall activity. Biotin Carboxyl Carrier Protein (*BCCP*) is one of the important parts of Acetyl-CoA Carboxyl (ACC). Biotin dependent, Acetyl-CoA Carboxylase (ACCase) is an enzyme which plays an important role in biosynthesis of fatty acids in plants by carboxylation of Acetyl-CoA into malonyl-CoA (Salie et al., 2016). Biotin Carboxylase (BCase) is one of other three functional domains of ACCase, responsible for binding to ATP, activation of CO_2 by converting into carboxyphosphate and transfers the activated carboxyl group from carboxyphosphate to the biotin molecule attached to the Biotin Carboxyl Carrier Protein (*BCCP*) domain of ACCase (Cui et al., 2017; Reverdatto et al., 1999). This reaction yields Malonyl Co-A formation, the first committed step of fatty acids synthesis in plants as it serves as substrate for Fatty Acid Synthase (FAS) which results in formation of long chained fatty acids, primarily Palmitic Acid (16 Carbons) and Stearic Acid (18 Carbons). (Baud and Lepiniec, 2009) These lipo-compounds further undergoes the esterification with glycerol to form Triacylglycerols (TAGs) (Li-Beisson et al., 2013). These TAGs accumulate in large amounts in mature seeds which ultimately increases the oil content of the seeds (Cui et al., 2017).

Overall reaction: Acetyl-CoA + HCO3- + ATP → Malonyl-CoA + ADP + Pi

Melatonin is a plant growth regulator and powerful antioxidant protecting the cells from environmental stress. . As an antioxidant, this hormone has role in enhancement of stress tolerance of soybean and accumulation of fatty acids (Wei et al., 2015). Under normal and abiotic stress conditions, melatonin upregulates the genes involved in glycolysis. During the glycolysis, glucose molecules are converted to pyruvate, which is further converted into Acetyl Co-A, responsible for fatty acid biosynthesis (Wei et al., 2015; Zhang et al., 2016). Melatonin has emerged as a pleiotropic signaling molecule in plants which enhances the tolerance to various abiotic stresses, including salinity (Wei et al., 2015). It is known to regulate the expression of numerous stress-responsive genes. However, whether melatonin specifically targets the BCCP gene family to modulate fatty acid biosynthesis as part of its protective mechanism in soybean remains an open and important question

While the BCCP gene family has been studied in species like Arabidopsis but the research on the BCCP gene family in soybean (Glycine max) has been limited. Abiotic stresses, particularly high salinity, are major constraints on soybean productivity and are known for its impact metabolic pathways which include lipid biosynthesis. However, a comprehensive analysis of the *GmBCCP* gene family and its transcriptional regulation under salt stress has not yet been performed. This represents a significant gap in our understanding of how soybeans regulate fatty acid production under adverse environmental conditions. Therefore, this study was undertaken to: (i) perform a genome-wide identification and comprehensive bioinformatic characterization of the BCCP gene family in Glycine max; (ii) analyze the expression patterns of *GmBCCP* genes under salt stress and melatonin treatment using available RNA-seq data; and (iii) elucidate the potential role of the *GmBCCP* genes in regulating fatty acid biosynthesis in response to these conditions. This research provides a foundation for understanding this crucial gene family and for future efforts in soybean crop improvement.

Material and Method

BCCP Sequence retrieval from databases and BLAST in G. Max

National Center for Biotechnology Information (NCBI) Protein database https://www.ncbi.nlm.nih.gov/ was used to retrieve the peptide sequence of BCCP domain from Arabidopsis. Gene Bank ID of BCCP Domain is AAM63178.1. Retrieved sequence was validated by using motif finder database https://www.genome.jp/tools/motif/. Query sequence for Biotin Carboxylase C-terminal Domain (Pfam ID: PF02785) from motif finder was utilized to BLAST on Phytozome v13 (https://phytozome-next.jgi.doe.gov/) by selecting genome of *Glycine Max Lee* v1.1 with Expect (E) threshold value at 1e-5 on 02 July, 2024 at 17:47 Pakistan time. All Sequences from Phytozome were further analyzed by using online database, NCBI CDD (Conserved Domain Database) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). for the identification of presence of Biotin Carboxylase C-terminal Domain in all genes.

Phylogeny Analysis

For the analysis of evolutionary relation of *GmBCCP* genes with other crops, Phylogenetic tree was constructed. MEGA 11, an offline software, was utilized for the alignment of different peptide sequences of *BCCP* from different crops and *G. max L.* Other crops were *Brassica rapa*, *Arabidopsis thaliana*, *Helianthus annuus*, and *Gossypium hirsutum L.* MUSCLE alignment program was used to align peptide sequences. Phylogenetic trees were constructed by using neighborhood joining algorithms in MEGA 11 with bootstrap by number of 1000

replications. Newick file from MEGA 11 was used to represent the phylogenetic tree on an online tool, iTOL (https://itol.embl.de/upload.cgi).

Physiochemical Characteristics and Subcellular Localization

An online platform Expasy Protparam (https://web.expasy.org/protparam/) was used to identify multiple parameters like, number of amino acids, protein length, protein molecular weight, protein GRAVY value, Isoelectric point of protein, Instability index, and Aliphatic index of retrieved protein. Subcellular localization of protein in cell was identified by using the online tool WoLF PSORT (https://wolfpsort.hgc.jp/) to find the presence of protein and its functional area. Details of original gene name IDs and Chromosomal location of *GmBCCP* were found by using Phytozome v13 (https://phytozome-next.jgi.doe.gov/).

Gene Structure and Cis Regulatory Elements (CREs) Identification

A gene includes promotor, ORF and terminator region. CREs are part of the Promotor region. For CREs, Plant CARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html) database was used to locate the position and presence of specific CREs in the promotor region of each gene. Promotor sequences were retrieved from 1500 upstream of genomic sequence in Phytozome v13. ORFs include motifs, and Intron-Exons of that gene. Conserved motifs in the BCCP protein sequences were identified using an online database, MEME suite (http://meme.sdsc.edu/meme/website/intro.html) with the following parameters as number of motifs = 10, minimum width = 6, maximum width = 50. Meme.xml file was downloaded and visualized by using TBtool. For the evaluation of ORF, Intron-Exon identification was conducted by using Gene Structure Display Server v2.0 (GSDS) (http://gsds.cbi.pku.edu.cn/). Genomic and CDS sequences, retrieved from Phytozome v13, were used to construct gene structure, introns and exons.

Chromosomal Mapping and Evolutionary Analysis

The molecular evolutionary rates and duplication of genes were determined by using simple Ka_Ks calculation in TBtool. For this analysis, CDS sequence derived from Phytozome and Gene pair were used. These estimates were further calculated by using lambda (λ) value of *Capsicum annuum* for the calculation time of evolution in gene (million years ago MYA). The formula used was (T=Ks/2 λ). Where $\lambda=6.9\times10^{-9}$ (Pan et al., 2021).

Chromosomal Mapping was visualized by using the function, Advance Gene Location, in TBtool. This analysis required 4 files which were created by using Phytozome database. Data required for files was Chromosomal length, location of gene on chromosome, start and end position of gene. Evolution of *BCCP* genes was also examined by synteny analysis. Identification of linkage of genes present on different chromosomes was made by using Advance Circos feature present in TBtool. This analysis describes the interaction of orthologous genes to perform the specific function in organism.

Gene Expression Analysis

Gene Expression Profiling of Soybean Cultivar Response to Melatonin and Salt Stress

To investigate the response of soybean cultivar to plant growth regulator and abiotic stress, melatonin and salt (NaCl) treatments were used. Data of expression profiling by high throughput sequencing was used from NCBI Geo, an online database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE57960). Three-week old seedlings were treated with water, 100 µM melatonin, 1% NaCl and 100 µM melatonin plus 1% NaCl. Then expressions of different genes recorded and targeted the expression of 11 *BCCP* genes. It is important to note that the transcriptome data used in this study was not validated with qRT-PCR. Future experimental validation is needed to confirm the precise expression levels of these *GmBCCP* genes.

Gene Ontology and P-P Interaction

The biological, molecular and cellular functions of *BCCP* gene were further analyzed by using an online database ShinyGo v. 0.745 (http://bioinformatics.sdstate.edu/go/) by GO (Gene Ontology) annotation. For protein-protein interaction (PPI) analysis, STRING v11.0 (https://string-db.org/) was used with a high confidence score of 0.7. The PPI network was constructed by integrating active interactions with an interaction score exceeding 0.4, sourced from text mining, studies, gene fusion data, databases, and co-expression analysis (Sami et al., 2024).

Results

BCCP genes in G. Max and its Physiochemical Properties

BCCP (Biotin Carboxyl Carrier Protein) gene, retrieved from sequence of *A. thaliana* from NCBI database (Gene Bank = AAM63178.1) was used to investigate its presence in the genome of *glycine max* Wm82.a2. v1 by BLASTP program of phytozome. 11 BCCP hits were found in the genome of *glycine max* and named as *GmBCCP1*, *GmBCCP2*, *GmBCCP3*, *GmBCCP4*, *GmBCCP5*, *GmBCCP6*, *GmBCCP7*, *GmBCCP8*, *GmBCCP9*, *GmBCCP10*, *GmBCCP11* as peptide sequences were retrieved from database. The results of physiochemical properties of all these genes showed their length in this species varied between 334-2260 amino acids, and their predicted molecular weight (MW) and isoelectric point varied between 36.960 to 252.37045 KDa and 4.9-9.12, respectively. Gravy values of all genes were found negative except one, *GmBCCP7*. Instability index (II) and Aliphatic index (AI) values also ranged from 28.7-46.06 and 72.19-94.88; Supplementary File; Table 1.

Phylogeny Analysis

To assess the evolution of *BCCP* homologs, Phylogenetic trees were constructed among 31 BCCP

peptide sequences from different oilseed crops like *B. rapa, H. annus, A. thaliana, G. hirsutum*, and 11 BCCP peptide sequences of *G. max* by using neighbor-joining (NJ) in MEGA 11. Phylogenetic tree (Figure 1) was classified into 5 different clades based on the presence of 4 *BCCP* genes from *Arabidopsis thaliana* in clade. 3 clades out of 5 revealed high sequence similarity and a close orthologous relationship between *GmBCCP2* and *GmBCCP4* with *AtBCCP3* and *AtBCCP2*, *GmBCCP3* and *GmBCCP9* with *AtBCCP1 and GmBCCP1* and *GmBCCP10* with *AtBCCP4.* The remaining *GmBCCP*(s) didn't show any specified distinguished genetic functionality with *Arabidopsis thaliana.*

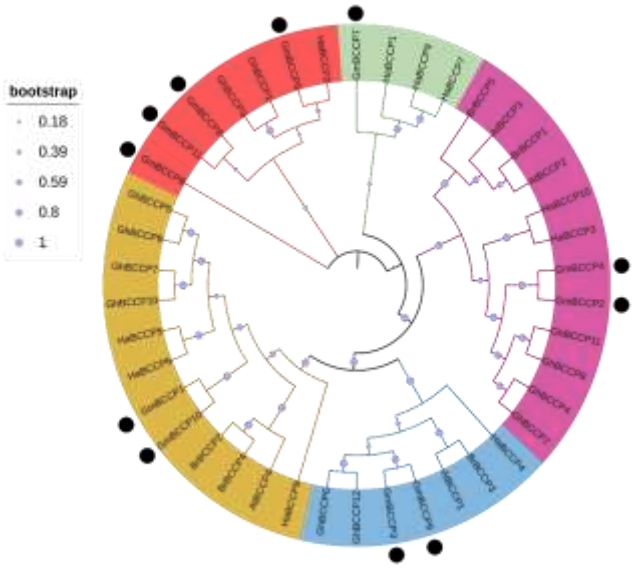

Figure 1 Phylogenetic analysis and gene structure of the BCCP gene family in soybean. (A) The unrooted phylogenetic tree of 11 GmBCCP proteins. The proteins were clustered into five clades (I-V), indicated by different background colours. (B) Exon-intron structure of the GmBCCP genes. Green boxes represent exons, and black lines represent introns. The scale at the bottom indicates gene length in kilobases (kb).

Subcellular Localization

Protein subcellular localization is important for proper functioning and pathway analysis. According to Wolf PSORT assessment, the result of the single peptide prediction showed the concentration of the presence of genes in cellular organelles as in (Figure 2). Proteins like *GmBCCP1, GmBCCP2, GmBCCP4, GmBCCP7,* and *GmBCCP10* were found in chloroplast. *GmBCCP3 and GmBCCP9* were found abundantly in mitochondria. *GmBCCP5* and *GmBCCP11* were found in nucleus of the cell. Some amounts of *GmBCCP6* and a little of *GmBCCP8 and GmBCCP11* were also found in cytoplasm.

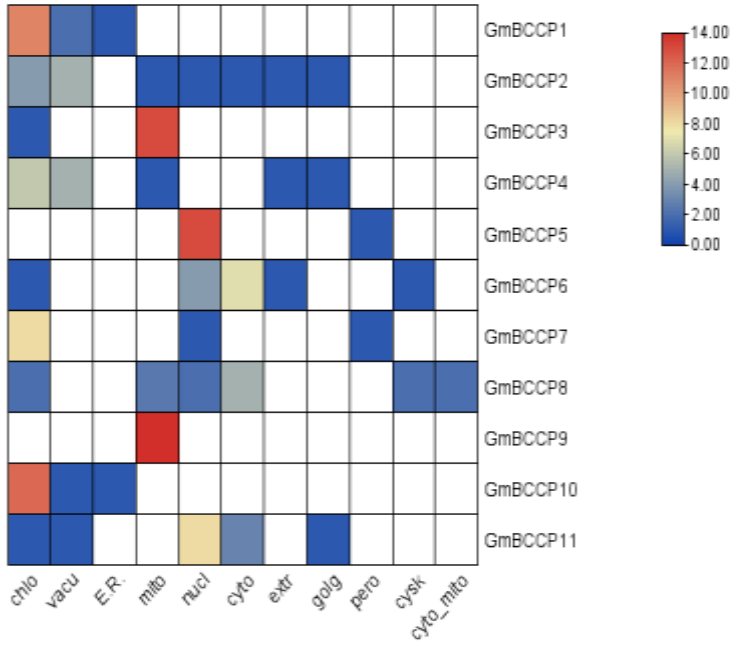

Figure 2 ***Subcellular Localization Prediction of GmBCCP Proteins****: This figure displays the predicted subcellular locations for each of the 11 GmBCCP proteins, including chloroplast, mitochondria, nucleus, cytoplasm, and vacuole, suggesting their diverse functional roles within the cell.*

Cis Regulatory Elements (CREs) Identification

Cis-Regulatory Elements (CREs) are very important for the specific function and behavior of a gene. CREs analyzed from PlantCARE databse (Figure 3) revealed multiple binding sites for transcription factors on genes which are key role player for upregulation or downregulation of gene for that specific function. All genes except *GmBCCP6* have high numbers of TATA-Box in their promotor region as it serves as a recognition site for the binding of the TATA-binding protein (TBP) and other general transcription factors. Respectively, CAAT-Box was found abundantly in all genes which significantly boosts the efficiency of transcription. Many other CREs, e.g., ABRE (ABA-responsive element), ARE (auxin-responsive element), AT~TATA-box, CGTCA-motif, ERE, G-Box, GT1-motif, MYB (for drought-inducibility), MYC, STRE, and TCT-motif were also found in apparently counts and provided in Supplementary file Table 2

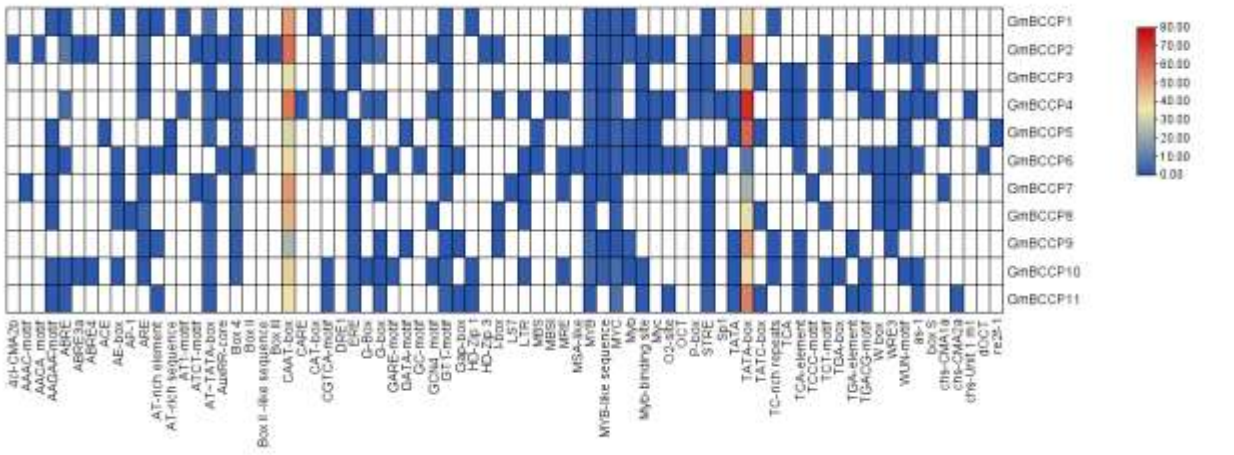

Figure 3 Heat map of statistics of Cis Regulatory Elements (CREs) for their number of presence which are associated with different functions with plant developmental process

Conserved Domain and Motif Analysis

ORF is the main part of a gene which has a secret code for a specific function. As the basic rule of transcription and translation, a gene produces multiple polypeptide chains as the domains of whole protein. Motif analysis of the *BCCP* gene family provided further evidence of conserved domains and motifs within the Glycine max genomes. The analysis showed that *BCCP* genes harbor characteristic domains such as Biotin Carboxylase (both N and C terminal), CPSase (L_D2 and L_D2_superfamily), and Biotin_lipoyl, alongside additional features like ACC_central and Carboxyl_trans. Notably, Motifs 3, 4, and 5 were consistently identified in all *BCCP* genes, highlighting their functional significance. Furthermore, a cluster of 10 motifs (Motifs 2, 6, 7, 9, 1, 10, 8, CPSase_L_D2, Biotin_carb_C, and Biotin_carb_N) were conserved across various domains, including Biotin_lipoyl_domains, CPSase_L_D2_superfamily, Biotin_carb_C_superfamily, and ACC_central, predominantly observed in *GmBCCP1*, *GmBCCP10*, *GmBCCP3*, *GmBCCP9*, and *GmBCCP2*.

In contrast, a distinct set of 14 motifs (Motifs 4, 3, 5, 1, 10, PK_Tyr_Ser-Thr_superfamily, RING_Ubox_superfamily, SMC_N_superfamily, DUF1665, PPR_long_superfamily, PPR_2_superfamily, PPR_3_superfamily, MutS_V, and MAS_V) were uniquely conserved within domains (PK_Tyr_Ser-Thr_superfamily, RING_Ubox_superfamily, SMC_N_superfamily, DUF1665, PPR_long_superfamily, PPR_2_superfamily, PPR_3_superfamily, MutS_V, MAS_V, MIS, WD40, and ANAPC4_WD40_superfamily) identified in *GmBCCP4*, *GmBCCP8*, *GmBCCP7*, *GmBCCP5*, and *GmBCCP6*. Comprehensive details concerning motif sequences and their potential functions within the *BCCP* gene family can be accessed in (Figure 4).

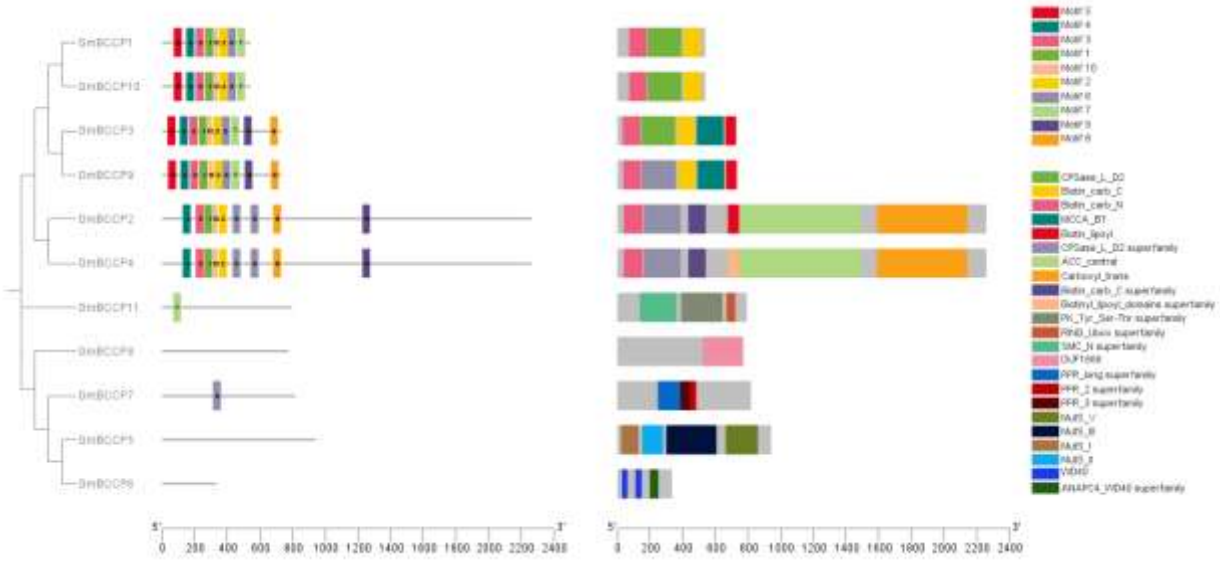

Figure 4 A colour-labelled bar graph shows the results of a motif distribution analysis of soybean BCCP protein by using MEME-Suite. The bar graph is connected to a phylogenetic tree to show the relationship between BCCP proteins and motif distribution.

Intron-Exon Identification:

Analysis of the *GmBCCP* gene structures (Figure 5) reveals a range of exon-intron. One gene, *GmBCCP7*, lacks introns entirely. Two genes (*GmBCCP6* and *GmBCCP8*) possess a relatively simple structure with fewer than 5 introns. The majority of *GmBCCP* genes (7 out of 11) exhibit more complex structures with 10 or more introns, suggesting potential for alternative splicing and greater regulatory complexity. *GmBCCP2* and *GmBCCP4* stand out with the highest number of introns at 30, suggesting a particularly complex gene architecture.

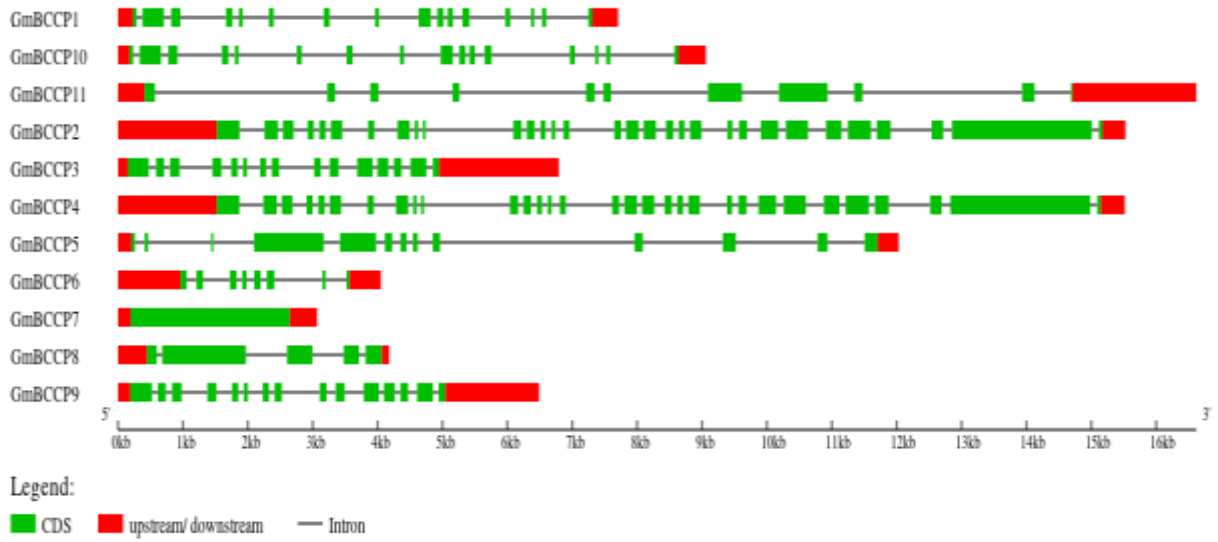

Figure 5 Different colours indicate the Intron-Exon, green shapes represent exons(CDS) sequences, red shapes represent UTR regions and Black lines represent Introns

Chromosomal Mapping and Gene Duplication

Chromosomal mapping allows to anticipate the spread of genes on different chromosomes in the genome. *BCCP* genes were found on chromosome numbers 2, 4, 5, 6, 8, 10, 14, 17, 19 and 20. No *BCCP* genes were found on 1, 3, 7, 9, 11, 12, 13, 15, 16, and 18. *GmBCCP6* and *GmBCCP7* were found on chromosome number 19. The remaining genes were present as one gene on one chromosome as shown in (Figure 6).

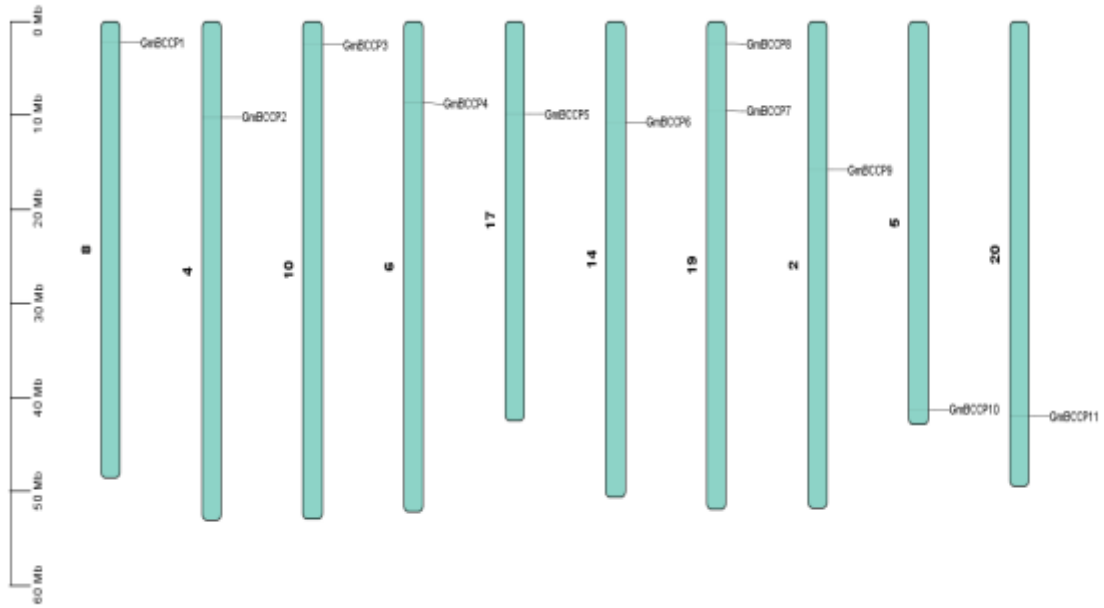

Figure 6 Chromosomal distribution of BCCP genes in Glycine Max. This provides the insight of spatial arrangement and potential interactions among BCCP genes.

For the gene pair *GmBCCP1_GmBCCP10*, Ka_Ks value was relatively low as 0.065659589 indicating that gene have experienced relatively few substitutions. The low value of MYA value of gene pair *GmBCCP2_GmBCCP4* (0.671494559) suggests that the gene divergence occurred relatively recently. In the same respect, the gene pair *GmBCCP6_GmBCCP7* indicates relatively high Ka_Ks value as 0.739349885, suggesting that genes have experienced many substitutions over the time.

The high MYA value of gene pair *GmBCCP7_GmBCCP8* (34.28541418) suggests that these genes diverged from a common ancestor long ago. The two gene pairs have undergone different levels of evolutionary change, with *GmBCCP6_GmBCCP7* experiencing more divergence over time than *GmBCCP1_GmBCCP10* (Figure 7).

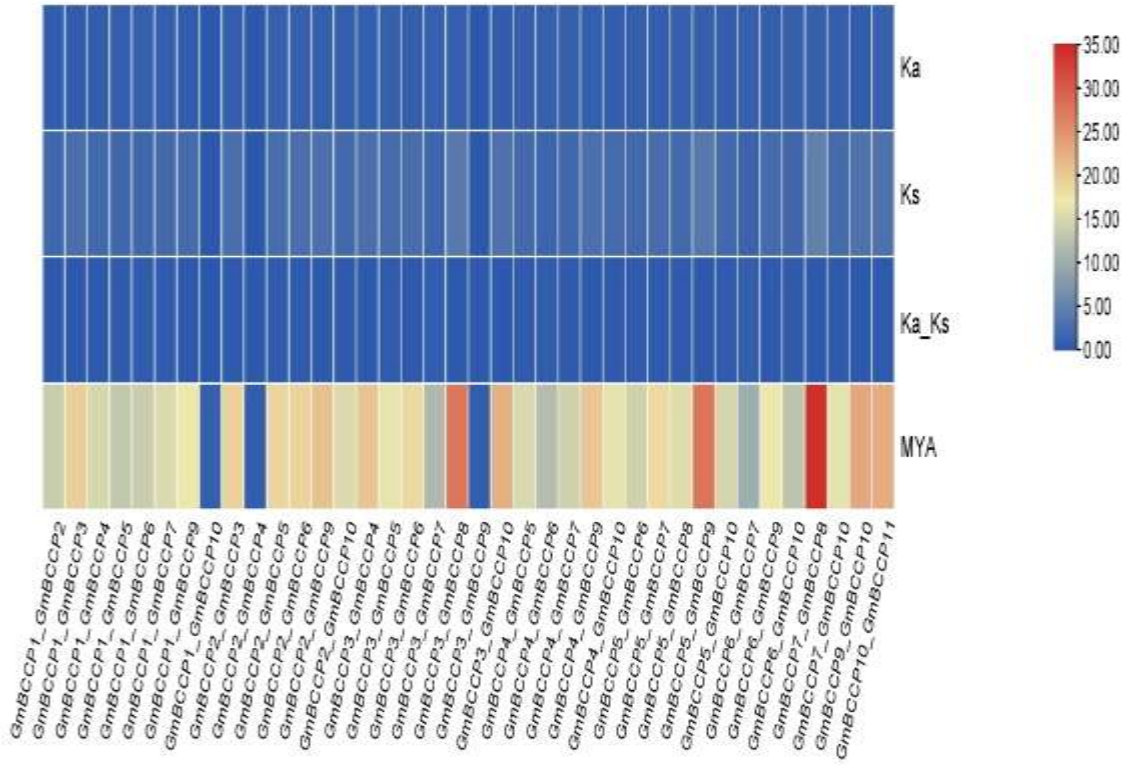

Figure 7 The Ks (synonymous substitution rate) and Ka (nonsynonymous substitution rate) were estimated by using TBtools. The Clock-like rate (λ) for soybean was 6.9×10^{-9}. The date of duplication was found by using formula (T=Ks/2 λ).

Synteny Analysis

Gene duplication is a fundamental process in evolution, serving as a crucial mechanism for generating functional divergence and innovation in organisms. To study locus relationship among the *BCCP* gene in glycine max, we identified the paralogous gene pairs in the genome of the *Glycine max L.* This analysis revealed that all our genes are present on different chromosomes as in (Figure 6) and *GmBCCP7* and *GmBCCP8* on chromosome 19 showed tandem duplication and other shows segmental duplication on their respective chromosomes. In (Figure 8) the chromosome number is indicated at the top of each chromosome, highlighting that segmental duplication of genes is more prevalent than tandem duplication within the *GmBCCP* gene family (Sami et al., 2024)

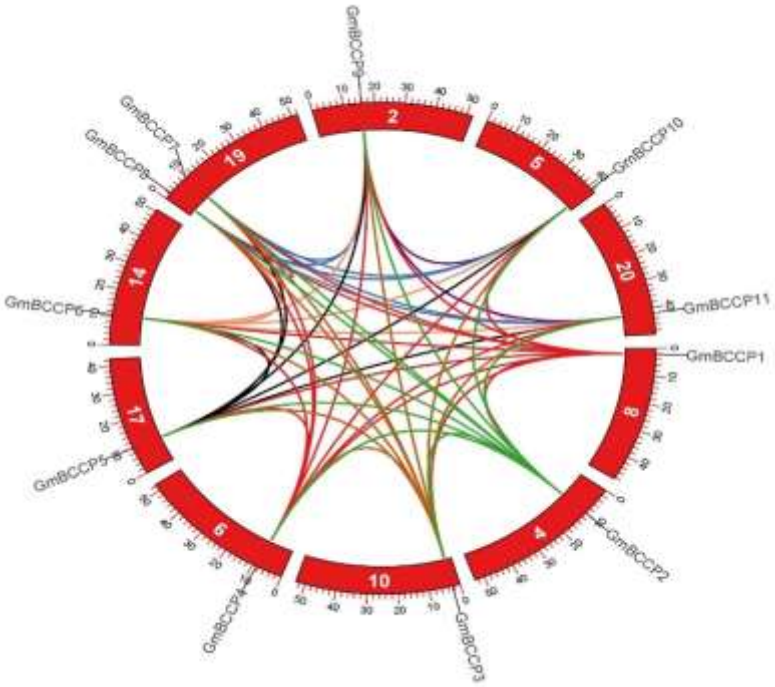

Figure 8 Synteny Analysis of GmBCCP across the genome of Glycine Max indicating the potential duplication on different chromosomes.

Gene Expression Analysis

This experiment compared gene expressions in soybeans treated with water, salt, melatonin, or both and helped to identify *GmBCCP* genes affected by melatonin and salt, individually and together. *GmBCCP1* and *GmBCCP10* showed significant upregulation under melatonin treatment compared to water controls, suggesting their potential involvement in melatonin-mediated growth promotion. In this respect, *GmBCCP2* was downregulated in the presence of salt, but this effect was mitigated when melatonin was also present, implying a possible role for melatonin in alleviating salt stress. *GmBCCP3* and *GmBCCP5* showed downregulation under salt stress, which was further enhanced by the addition of melatonin. Melatonin acted as fatty acid production enhancer and stress reducer.

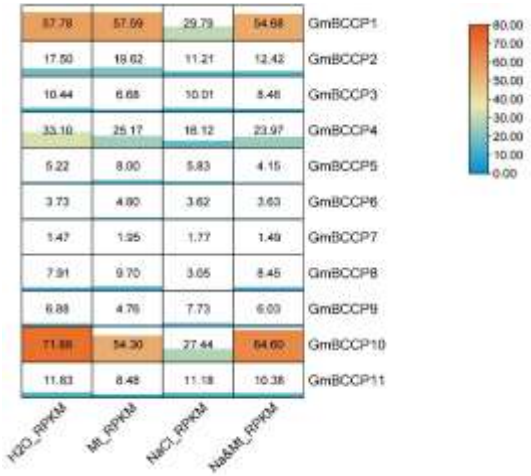

Figure 9 Heat map of the transcriptomic expression profile of Soybean's BCCP genes under water, melatonin, salt and salt + Melatonin conditions

Gene Ontology (GO) and P-P Interaction (PPI)

The GO enrichment analysis was carried out for the evaluation of the function of *BCCP* gene in G*lycine max L.* As a result, in (Figure 10), *BCCP* genes have a great role in Propanoate metabolism, Fatty acid biosynthesis, and other metabolic processes, majorly controlling the oil composition and oil content in soybean.

The network comprises 11 nodes (Proteins) and 10 edges (Interactions)indicating an average of 1.82 interaction per protein. The avg. local clustering coefficient is 0.303 suggesting a moderate level of clustering in the network Protein-Protein Interaction (PPI) in (Figure 11) revealed the interaction between 6 proteins out of 11 proteins (*GmBCCP1*, *GmBCCP2*, *GmBCCP3*, *GmBCCP4*, *GmBCCP9*, and *GmBCCP10*). Other Proteins (*GmBCCP5*, *GmBCCP6*, *GmBCCP7*, *GmBCCP8*, and *GmBCCP11*) didn't show any interaction.

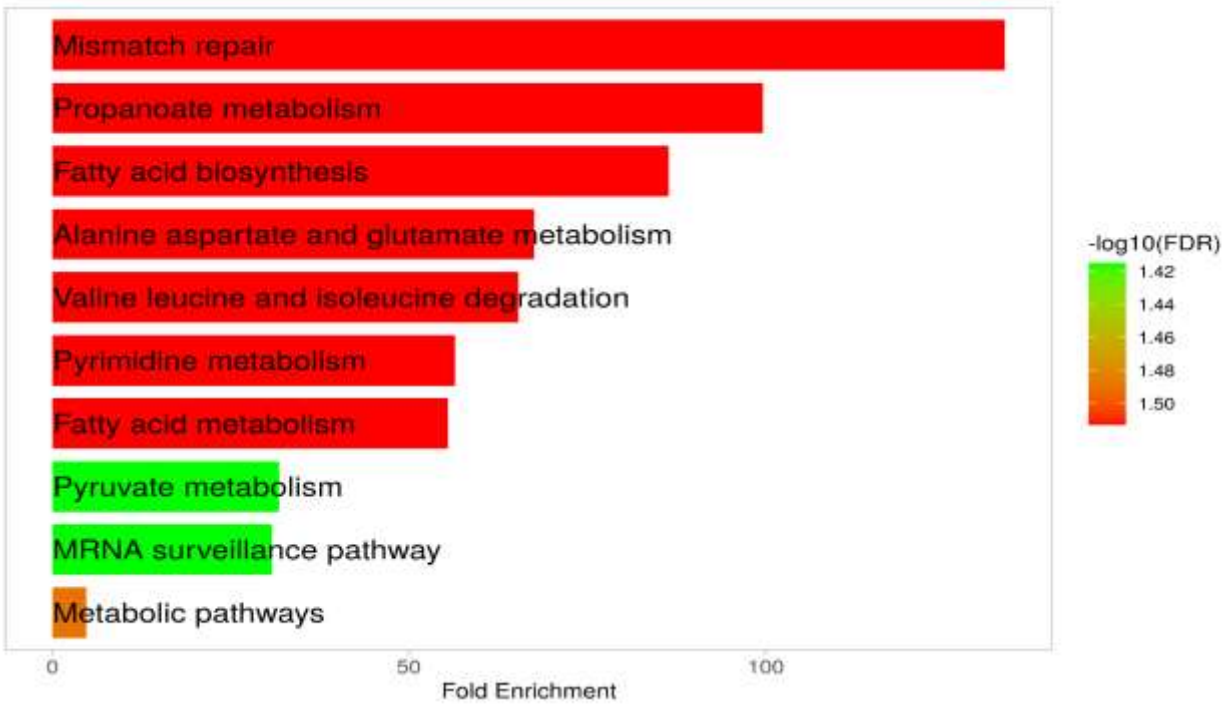

Figure 10 Fold Enrichment chart analysis representing the overlapping GmBCCP genes functions majorly involved in fatty acid biosynthesis.

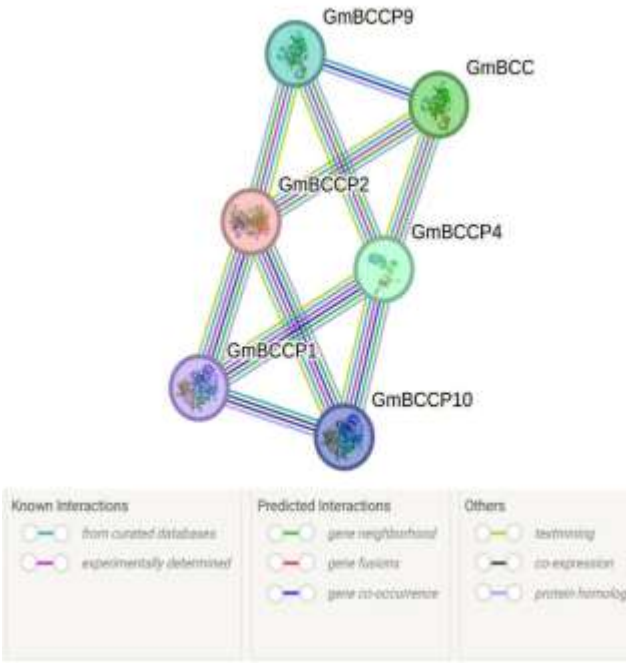

Figure 11 Protein-Protein Interaction of BCCP proteins in Glycine Max. Each colour represents a specific highly significant GO enrichment and coloured lines representing several types of interactions.

Discussion

Soybean is a significant oilseed crop, soybean seeds are major source of protein (13%) and oil (19%) (Nair et al., 2023). Biosynthesis of fatty acids in plants is done by multiple pathways. BCCP protein is one of the 4 different protein molecules of Acetyl CoA Carboxylase (ACCase) (Choi-Rhee and Cronan, 2003) and Acetyl-CoA Carboxylase (ACCase) helps in carboxylation of Acetyl-CoA into Malonyl-CoA to produce fatty acids by an irreversible reaction in different oilseed species like *Gossypium hirsutum L* and *Brassica sp*. etc. (Fukuda et al., 2013; Thelen et al., 2000). However, research on the *GmBCCP* genes family in soybean has been relatively limited compared to other species, leaving a gap in our knowledge about *BCCP* in Glycine max. Edible oil content studied in Arabidopsis was highly affected by mutation or modification in *BCCP* protein (Thelen et al., 2000). This study successfully identified 11 members of the BCCP gene family in soybean and provided the first comprehensive analysis of their characteristics and expression under salt stress, offering novel insights into their potential regulatory roles.

Physicochemical properties of *BCCP* proteins were studied to see how they differed from each other in the same clade of proteins. All identified *BCCP* proteins except one, were hydrophilic based on negative GRAVY value showing a significant interaction with water with net electric charge at different pH levels (Priya et al., 2019). BCCP proteins are naturally hydrophilic due to surface exposure as it is carrier protein for biotin, biotin binding due to binding site for biotin on BCCP is likely to be composed of hydrophilic amino acids, structural stability by hydrogen bonds and other favorable interactions, and aqueous environments as a functional requirements (Cronan, 2001). The Instability Index of 5 proteins was <40 indicating the stable protein molecules in In-vitro conditions and 6 proteins >40 showing the stable protein molecules (Ji et al., 2023). Subcellar localization analysis revealed the presence of proteins in different organelles such as chloroplast, mitochondria, nucleus, cytoplasm, and vacuole. About 50% of proteins have presence in chloroplast which indicates the biosynthesis of fatty acids by the pyruvate and acetyl-CoA as described in (Heredia-Martínez et al., 2018; Sasaki and Nagano, 2004). These proteins were also found in cytoplasm (where initial steps of fatty acid synthesis occur), mitochondria and nucleus (Sami et al., 2024).

Phylogeny analysis among 11 *BCCP* genes of *G. max*, 4 *BCCP* genes of *A. thaliana*, 10 *BCCP* genes of *H. annus*, 12 *BCCP* genes of *G. hirsutum* and 5 *BCCP* genes of *B. rapa*. Phylogenetic tree was classified based on presence of *AtBCCP(s)* as it was consistent to previously studies in *Brassicaceae* (Cui et al., 2017; Megha et al., 2022). The expansion and evolution of gene families through duplication events are key drivers of functional diversification in plants. Our analysis revealed that the 11 *GmBCCP* genes are structured into five distinct clades, with some members showing clear orthologous relationships to Arabidopsis counterparts which suggest functional conservation since their divergence. (Yao et al., 2022). These structural features provide a framework for understanding the potential functional specialization among *GmBCCP* genes

It indicates the similarities between those genes with Arabidopsis suggesting the same evolutionary time. Such comparisons of genomes between different species provide a valuable information about the evolutionary changes in genes throughout the time (Yao et al., 2022). During this evolutionary study, we identified 39 pairs of paralogous genes in genome which evolute from gene duplication events in the past. Such events in past provides the important information about the type of gene duplication either tandem or segmental duplication (Wei et al., 2022).

Previous studies have shown that the arrangement of exons and introns within groups of related genes have same role in their evolutionary development (Megha et al., 2022; Sami et al., 2024). This current research supports this statement, as analysis of gene structure and motifs revealed that genes within the same population or clade shared similar numbers and positions of exons, introns, and motifs (Wei et al., 2022). Exons and introns are the most important functional units of a gene. Under the current studies, all *BCCP* genes have introns and exons. It was found that motifs of the *BCCP* gene family were discovered to be more conserved than those of

the *CCD* subfamily (Megha et al., 2022). Relationship of Transcription Factors (TFs) and Cis-Regulatory Elements (CREs) like a lock and key mechanism to perform a specific function. Under this study, various CREs were found to have specific functions.

The identified Cis-Regulatory Elements (CREs) points towards a diverse array of regulatory functions in the plant. Several CREs are associated with stress and defense responses: ABRE and "ARE" are involved in abscisic acid and anaerobic responses, respectively; the TCA-element is linked to salicylic acid responsiveness; and the W-box is crucial for pathogen defense. Additionally, multiple CREs are implicated in growth and development: the G-Box and Box 4 are involved in light responsiveness (Naeem et al., 2024); the Myb family of transcription factors play a role in drought and low-oxygen responses (Ahmad et al., 2024b); and the GT1-motif is potentially associated with light-regulated gene expression (Wei et al., 2022). Furthermore, the CAAT-box and TATA-box are fundamental components of core promoters, while the ERE potentially regulates hormone signaling. The presence of the STRE suggests a role in general stress responses (Megha et al., 2022). Finally, several CREs like CGTCA-motif and TGACG-motif are implicated in the MeJA (methyl jasmonate) responsiveness pathway, which plays a role in plant defense and development (Cui et al., 2017; Sami et al., 2024).

RNA-seq data for current studies revealed the effects of melatonin and salt stress, individually and in combination, on the expression of *GmBCCP* genes in soybeans. These genes are involved in fatty acid metabolism(Rasheed and Malik, 2022). Our investigation highlight on the complex interaction between melatonin, salt stress, and *GmBCCP* gene expression and highlighting melatonin's potential roles in promoting growth and minimizing salt stress (Sami et al., 2024) (Ahmad et al., 2024a). The upregulation of *GmBCCP1* and *GmBCCP10* under melatonin treatment, as compared to water controls, suggests their potential involvement in melatonin-mediated growth promotion. It is possible that these genes contribute to increased fatty acid synthesis, providing the energy and building blocks necessary for growth (Cui et al., 2017).

The downregulation of *GmBCCP2* under salt stress, which effect was decreased by the addition of melatonin, suggesting the positive role of melatonin in the decrease of salt stress (Cronan, 2001; Mushtaq *et al*., 2024). This suggests that melatonin may counteract the negative impact of salt on *GmBCCP2* expression, potentially helping to maintain fatty acid metabolism and overall cellular function under stress conditions (Choi-Rhee and Cronan, 2003; Raza *et al*., 2025) and same as under abiotic stress in cucumber (Amjad et al., 2024). Further research is needed to determine the exact mechanisms by which melatonin effect *GmBCCP2* expression and to explore its broader role in salt stress tolerance. *GmBCCP3* and *GmBCCP5* showed downregulation under salt stress, which was enhanced by the addition of melatonin (Cui et al., 2017; Megha et al., 2022). This experiment highlights the interactions between melatonin, salt stress, and *GmBCCP* gene expression. It is possible that melatonin's effects on these genes are context-dependent or that they interact with other signaling pathways involved in the stress response. Furthermore, gene ontology shows the significant role of these genes in soybean.

Our results highlight specific *GmBCCP* genes that are highly responsive to stress and melatonin, marking them as prime candidates for future functional characterization. Validating these results with qRT-PCR and using techniques like gene editing (e.g., CRISPR/Cas9) to study knockout or overexpressing lines will be essential to definitively establish their roles in fatty acid synthesis and stress tolerance. Such research could lead to the development of new soybean varieties with improved oil content and resilience to environmental stress.

Conclusion

In this study, 11 *BCCP* genes were found in the whole genome of *G.max.* Based on structural analysis of promotor of these genes, cis-regulatory elements related to ABA-responsive element, auxin-responsive element, developmental and hormone responsiveness, recognition sites to enhance efficiency of transcription and for drought-inducibility in *Glycine max*. Structural analysis of gene, number of introns and exons ranges from one to thirty-one. As identified by an RNA-Seq data analysis, *GmBCCP1* and *GmBCCP10* under melatonin treatment in combination of salt, showed upregulation in fatty acid biosynthesis. Salt stress decreases the fatty acid biosynthesis which effect can be minimized by application of melatonin. However, additional research, including gene cloning and functional analysis, is required to confirm the significance of these genes in various physiological and biological processes.

References

Ahmad, A., Han, S., Sami, A., Haider, M. Z., Ali, Q., Shafiq, M., Ali, D., Iqbal, J., Manzoor, M. A., and Sabir, I. A. (2024a). Genome-wide identification of Cytochrome P450 gene in Fall Armyworm (Spodoptera frugiperda) in response to insecticide resistance. *Plant Stress* **14**, 100579. https://doi.org/10.1016/j.stress.2024.100579

Ahmad, A., Sami, A., Habib, U., Ali, M., Shafiq, M., Haider, M. Z., Ahmad, S., Ali, B., Harakeh, S., and Makki, R. M. (2024b). Genome wide analysis of carotenoid cleavage oxygenases (CCO) gene family in Arachis hypogaea (peanut) under biotic stress. *BMC genomics* **25**, 469. doi: 10.1007/s11103-024-01433-0

Amjad, M., Wang, Y., Han, S., Haider, M. Z., Sami, A., Batool, A., Shafiq, M., Ali, Q., Dong, J., and

Sabir, I. A. (2024). Genome wide identification of phenylalanine ammonia-lyase (PAL) gene family in Cucumis sativus (cucumber) against abiotic stress. *BMC Genomic Data* **25**, 76. https://doi.org/10.1186/s12863-024-01259-1

Baud, S., and Lepiniec, L. (2009). Regulation of de novo fatty acid synthesis in maturing oilseeds of Arabidopsis. *Plant Physiology and Biochemistry* **47**, 448-455. https://doi.org/10.1016/j.plaphy.2008.12.006

Choi-Rhee, E., and Cronan, J. E. (2003). The biotin carboxylase-biotin carboxyl carrier protein complex of Escherichia coli acetyl-CoA carboxylase. *Journal of Biological Chemistry* **278**, 30806-30812. https://doi.org/10.1074/jbc.m302507200

Cronan, J. E. (2001). The biotinyl domain of Escherichia coli acetyl-CoA carboxylase: evidence that the "thumb" structure is essential and that the domain functions as a dimer. *Journal of Biological Chemistry* **276**, 37355-37364. https://doi.org/10.1074/jbc.M106353200

Cui, Y., Zhao, Y., Wang, Y., Liu, Z., Ijaz, B., Huang, Y., and Hua, J. (2017). Genome-Wide Identification and Expression Analysis of the Biotin Carboxyl Carrier Subunits of Heteromeric Acetyl-CoA Carboxylase in Gossypium. *Frontiers in Plant Science* **8**. https://doi.org/10.3389/fpls.2017.00624

FAO.org (2017). National Policy Dialogue on Salt-affected Soils.

FAO_Stat (2020-2022). Import and Export of Pakistan *In* "2020-2022". https://www.fao.org/faostat/.

Fukuda, N., Ikawa, Y., Aoyagi, T., and Kozaki, A. (2013). Expression of the genes coding for plastidic acetyl-CoA carboxylase subunits is regulated by a location-sensitive transcription factor binding site. *Plant molecular biology* **82**, 473-483. doi: 10.1007/s11103-013-0075-7

Heredia-Martínez, L. G., Andrés-Garrido, A., Martínez-Force, E., Pérez-Pérez, M. E., and Crespo, J. L. (2018). Chloroplast damage induced by the inhibition of fatty acid synthesis triggers autophagy in Chlamydomonas. *Plant Physiology* **178**, 1112-1129. https://doi.org/10.1104/pp.18.00630

Ji, F., Wu, J., and Zhang, Z. (2023). Identification and Characterization of CCD Gene Family in Rose (Rosa chinensis Jacq.'Old Blush') and Gene Co-Expression Network in Biosynthesis of Flower Scent. *Horticulturae* **9**, 115. https://doi.org/10.3390/horticulturae9010115

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M. X., Arondel, V., Bates, P. D., Baud, S., Bird, D., DeBono, A., and Durrett, T. P. (2013). Acyl-lipid metabolism. *The Arabidopsis book/American Society of Plant Biologists* **11**. Doi:10.1199/tab.0161

Liu, S., Zhang, M., Feng, F., and Tian, Z. (2020). Toward a "green revolution" for soybean. *Molecular plant* **13**, 688-697. https://doi.org/10.1016/j.molp.2020.03.002

Megha, S., Wang, Z., Kav, N. N., and Rahman, H. (2022). Genome-wide identification of biotin carboxyl carrier subunits of acetyl-CoA carboxylase in Brassica and their role in stress tolerance in oilseed Brassica napus. *BMC genomics* **23**, 707. https://doi.org/10.1186/s13038-022-02194-8

Mushtaq, F., Akram, M., Usman, M., Mohsin, M., & Nawaz, M. (2024). Global climate change and its influence on crop production. *Journal of Life and Social Sciences*, **2024**(1), 27. https://doi.org/10.64013/bbasrjlifess.v2024i1.27

Naeem, S., Wang, Y., Han, S., Haider, M. Z., Sami, A., Shafiq, M., Ali, Q., Bhatti, M. H. T., Ahmad, A., and Sabir, I. A. (2024). Genome-wide analysis and identification of Carotenoid Cleavage Oxygenase (CCO) gene family in coffee (coffee arabica) under abiotic stress. *BMC Genomic Data* **25**, 71. DOI: 10.1186/s12863-024-01248-4

Nair, R. M., Boddepalli, V. N., Yan, M.-R., Kumar, V., Gill, B., Pan, R. S., Wang, C., Hartman, G. L., Silva e Souza, R., and Somta, P. (2023). Global status of vegetable soybean. *Plants* **12**, 609. doi:10.3390/plants12030609

Pan, X., Hu, M., Wang, Z., Guan, L., Jiang, X., Bai, W., Wu, H., and Lei, K. (2021). Identification, systematic evolution and expression analyses of the AAAP gene family in Capsicum annuum. *BMC genomics* **22**, 463. https://doi.org/10.1186/s12864-021-07765-1

Priya, R., Sneha, P., Dass, J. F. P., Manickavasagam, M., and Siva, R. (2019). Exploring the codon patterns between CCD and NCED genes among different plant species. *Computers in biology and medicine* **114**, 103449. doi: 10.1016/j.compbiomed.2019.103449

Rasheed, M., and Malik, A. (2022). Mechanism of drought stress tolerance in wheat. *Bulletin of Biological and Allied Sciences Research* **2022**, 23-23. https://doi.org/10.54112/bbasr.v2022i1.23

Raza, A., Ayub, M., & Abbas, A. (2025). Genome-wide identification and characterization of pbs3 plant-specific transcription factor gene family in carrot species (*Daucus carota* L.). *Journal of Physical, Biomedical and Biological Sciences*, **2025**(1), 39. https://doi.org/10.64013/jpbab.v2025i1.39

Reverdatto, S., Beilinson, V., and Nielsen, N. C. (1999). A multisubunit acetyl coenzyme A carboxylase from soybean. *Plant physiology* **119**, 961-978. https://doi.org/10.1104/pp.119.3.961

Salie, M. J., Zhang, N., Lancikova, V., Xu, D., and Thelen, J. J. (2016). A family of negative regulators targets the committed step of de novo fatty acid biosynthesis. *The Plant Cell* **28**, 2312-2325. https://doi.org/10.1105/tpc.16.00317

Sami, A., Haider, M. Z., Shafiq, M., Sadiq, S., and Ahmad, F. (2024). Genome-wide identification and in-silico expression analysis of CCO gene family in sunflower (Helianthus annnus) against abiotic stress. *Plant Molecular Biology* **114**, 34. https://doi.org/10.1007/s11103-024-01433-0

Sasaki, Y., and Nagano, Y. (2004). Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. *Bioscience, biotechnology, and biochemistry* **68**, 1175-1184. https://doi.org/10.1271/bbb.68.1175

Siddiqui, R., Samad, G., Nasir, M., and Jalil, H. H. (2012). The impact of climate change on major agricultural crops: evidence from Punjab, Pakistan. *The Pakistan Development Review*, 261-274. https://doi.org/10.30541/v48i1pp.104-107.

Thelen, J., Mekhedov, S., and Ohlrogge, J. (2000). Biotin carboxyl carrier protein isoforms in Brassicaceae oilseeds. Portland Press Ltd. https://doi.org/10.1042/bst0280595.

Wei, H., Liu, G., Wang, Y., Chen, J., Chen, Y., Lian, B., Zhong, F., Yu, C., and Zhang, J. (2022). Genome-Wide Identification and Expression Analysis of Carotenoid Cleavage Oxygenase Genes in Crape Myrtle. *Available at SSRN 4294416*.

Wei, W., Li, Q.-T., Chu, Y.-N., Reiter, R. J., Yu, X.-M., Zhu, D.-H., Zhang, W.-K., Ma, B., Lin, Q., and Zhang, J.-S. (2015). Melatonin enhances plant growth and abiotic stress tolerance in soybean plants. *Journal of experimental botany* **66**, 695-707. https://doi.org/10.1093/jxb/eru392

Xu, H., Guo, Y., Qiu, L., and Ran, Y. (2022). Progress in soybean genetic transformation over the last decade. *Frontiers in Plant Science* **13**, 900318. https://doi.org/10.3389/fpls.2022.900318

Yao, Y., Jia, L., Cheng, Y., Ruan, M., Ye, Q., Wang, R., Yao, Z., Zhou, G., Liu, J., and Yu, J. (2022). Evolutionary origin of the carotenoid cleavage oxygenase family in plants and expression of pepper genes in response to abiotic stresses. *Frontiers in Plant Science* **12**, 792832. https://doi.org/10.3389/fpls.2021.792832

Zhang, N., Sun, Q., Li, H., Li, X., Cao, Y., Zhang, H., Li, S., Zhang, L., Qi, Y., and Ren, S. (2016). Melatonin improved anthocyanin accumulation by regulating gene expressions and resulted in high reactive oxygen species scavenging capacity in cabbage. *Frontiers in Plant Science* **7**, 197. https://doi.org/10.3389/fpls.2016.00197

Statements and Declarations

Data Availability Statement

All data are fully available and can be found within the manuscript and its supplementary file.

Acknowledgments

The authors would like to extend their sincere appreciation to the Ongoing Research Project Program (ORP-2025-165), King Saud University, Riyadh, Saudi Arabia.

Conflicts of Interest

The authors declare no conflict of interest.

Ethics Approval Statement

Not applicable.

Ethics Approval Statement

Not applicable.

Consent to Participate

Not applicable.

Consent to Publish

Not applicable.

Author Contributions

Conceptualization by MUR, AM and AS; methodology by MZH. and AS; software by MAJ; validation by DA., MZH and QA.; resources by AM and MZH; data curation by DA., AS and QA; writing—original draft preparation was prepared by MUR., writing—review and editing by QA., MZH., AS., and MAJ; visualization by QA and DA; supervision, MAJ and QA; project administration, DA and QA.; All authors have read and agreed to the published version of the manuscript."

Funding

Ongoing Research Funding Program, (ORF-2025-165), King Saud University, Riyadh, Saudi Arabia.