J Appl Biotechnol Rep. 2025 September;12(3):1712-1722

Journal of
Applied Biotechnology
Reports

doi 10.30491/jabr.2025.479140.1783

Original Article

Regulation of the Tocopherol Biosynthetic Pathway by AREB5 in *Cannabis sativa*

Sara Kazemzadeh [1], Naser Farrokhi [1*], Asadollah Ahmadikhah [1*]

[1] Department of Cell and Molecular Biology, Faculty of Life Sciences & Biotechnology, Shahid Beheshti University, Tehran, Iran

Corresponding Authors: Naser Farrokhi, PhD, Associate Professor, Tel: +989124405840, E-mail: n_farrokhi@sbu.ac.ir; Asadollah Ahmadikhah, PhD, Assistant Professor, Tel: +989112734072, E-mail: ahmadikhaha@gmail.com; Department of Cell and Molecular Biology, Faculty of Life Sciences & Biotechnology, Shahid Beheshti University, Tehran, Iran.

Received September 22, 2024; **Accepted** January 7, 2025; **Online Published** September 30, 2025

Abstract

Introduction: The oil extracted from hemp (*Cannabis sativa*) seeds can be used in human nutrition, and several components of hemp oil such as tocopherols benefit human health. Increasing the content of natural antioxidants like tocopherols in the oil seed is achieved through the metabolic engineering of key enzymes or regulatory transcription factors involved in tocopherol biosynthesis. Although there is significant information regarding the genes involved in the tocopherol biosynthesis pathway, the molecular mechanisms regulating its transcription remain unclear.

Materials and Methods: In this study, using fluorescent HPLC, the contents of vitamin E isomers (α, δ, and γ) were measured at four developmental stages (germination, milky, doughy, and mature) of cannabis seed. Based on previous study, candidate TFs involved in tocopherol biosynthesis were identified. RT-qPCR was used to evaluate the expression of candidate transcription factors in different seed stages. Statistical analyses were performed using two-way ANOVA and Pearson's correlation coefficients.

Results: High-performance liquid chromatography (HPLC) identified three tocopherol isomers (α, γ, δ), with γ-tocopherol being predominant. The highest tocopherol concentration (12.5 μg/g) was observed at the seed maturation stage. Expression analysis of transcription factors (*AREB5*, *NAC072*, and *PIF3*) revealed significant upregulation of *AREB5* during seed maturation, correlating strongly with tocopherol levels (*r* : 0.86–0.94). These findings suggest *AREB5* pivotal role in tocopherol biosynthesis and provide insights for future metabolic engineering in hemp.

Conclusions: This study offers insights for ongoing biofortification efforts to breed and/or engineer vitamin E in cannabis and other oilseeds.

Keywords: Hemp, HPLC, Seed Development, Transcription Factor, Vitamin E

Citation: Kazemzadeh S, Farrokhi N, Ahmadikhah A. Regulation of the Tocopherol Biosynthetic Pathway by AREB5 in *Cannabis sativa.* J Appl Biotechnol Rep. 2025;12(3):1712-1722. doi:10.30491/jabr.2025.479140.1783

Introduction

Cannabis sativa L., commonly known as hemp, a member of the Cannabaceae family, is one of the world's oldest domesticated crops.[1] It is believed to have originated in Central Asia, from where its cultivation rapidly spread throughout Asia and Europe. Cannabis has a diploid genome ($2n = 20$) consisting of nine pairs of autosomes and a pair of sex (X and Y) chromosomes.[2] It is predominantly dioecious, meaning a plant is either male or female, with estimated haploid genome sizes of 843 Mb and 818 Mb for male and female plants, respectively.[3]

Hemp seed oil has been praised for its nutritional properties and health benefits due to of its relatively high levels of polyunsaturated fatty acids (PUFAs) and natural antioxidants. The oil contains linoleic acid (55.41–59.64%), α-linoleic acid (16.51–20.40%), oleic acid (11.40–15.88%), palmitic acid (6.08–6.82%), and stearic acid (2.34–2.67%)[4]. Additionally, 25–35% of the seeds are proteins, 10–15% are fibers, and 20–30% are carbohydrates[5]. Certain ecotypes may contain up to 4% of γ-linoleic acid.[6]

Tocopherols are widely known bioactive compounds present in hemp seed oil. Four different isomers of tocopherols are found in the oil, namely α-, β-, γ- and δ, of which γ-tocopherol exhibits the highest antioxidant activity[7] and prevents unsaturated fatty acid oxidation,[9] which may reduce the risk of cardiovascular diseases, cancer and age-related macular degeneration.[8] The latest research shows that tocopherols can act as chemically protective agents against hepatotoxicity.[11] The recommended daily allowance of vitamin E is 15 mg per day for adults.[12] The total tocopherol content of hemp seed oil is higher than canola oil (60–70 mg/100 g), sunflower oil (~90 mg/100 g), and soybean oil (~115 mg/100 g).[10]

However, tocopherols are vulnerable when exposed to heat, light and alkali as they easily form oxidation products during processing and storage. The oxidation of tocopherol would cause the loss of the antioxidant properties in the oil.[13]

Vitamin E is synthesized by photosynthetic organisms[14] and its biosynthetic pathway is highly conserved in the plant kingdom.[15] Vitamin E biosynthesis occurs in the inner envelope of plastids via a combination of two main pathways (Figure 1a). The shikimate (SK) pathway gives rise to the chromanol ring from homogentisate (HGA). Meanwhile, the methylerythritol phosphate (MEP) pathway provides the prenyl tail from geranylgeranyl diphosphate (GGDP) and phytyl diphosphate (phytyl-DP) for the synthesis of tocotrienol and tocopherol, respectively.[16]

Transcription factors (TFs) are *trans*-acting proteins that bind to specific *cis*-elements in the promoter regions of downstream genes and regulate gene expression. TFs are typically modular, consisting of a DNA-binding domain (DBD)

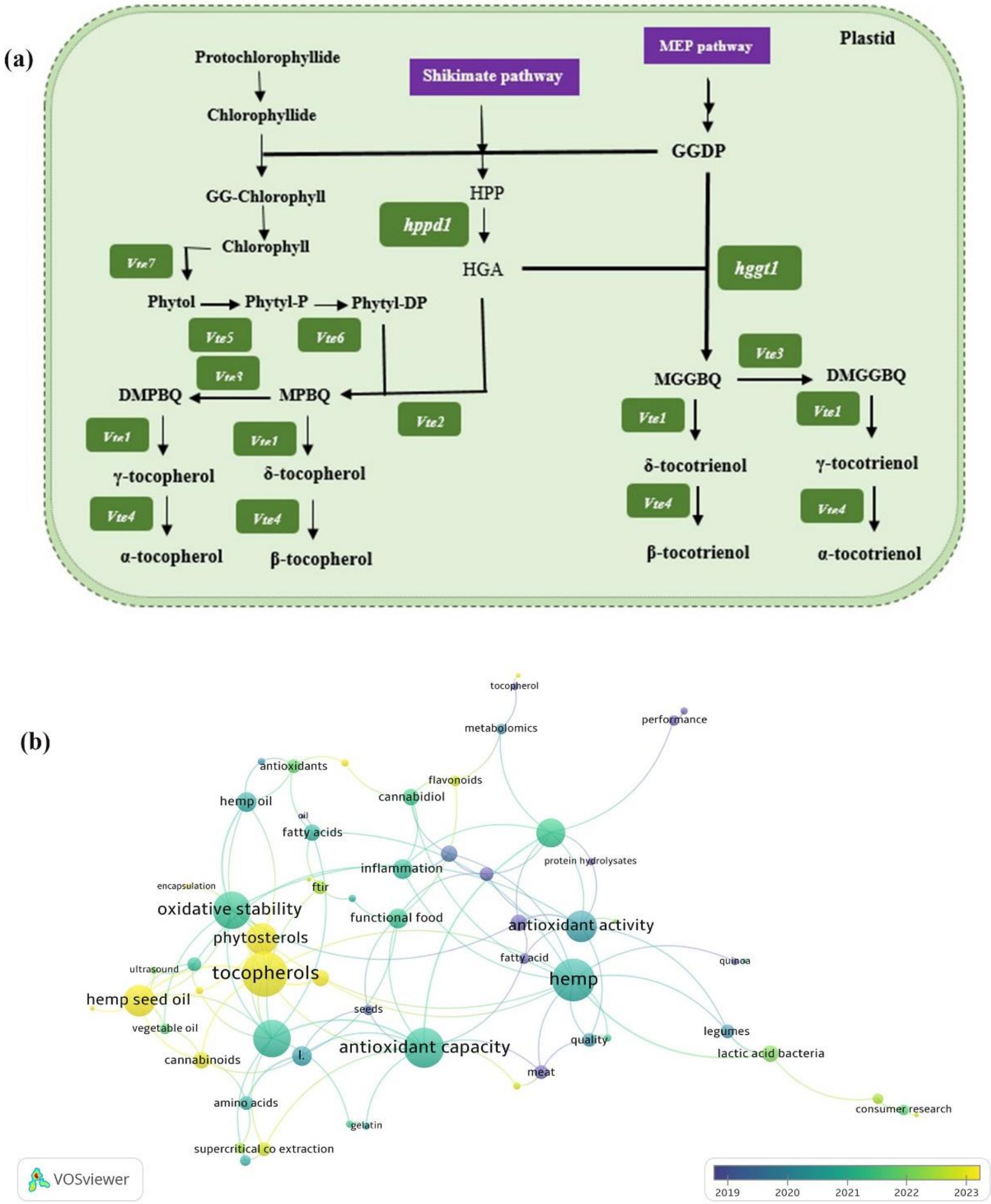


Figure 1. a) Simplified model of the tocopherol biosynthesis pathway. The shikimate pathway provides HGA as the precursor for tocopherol synthesis. PDP as the lipid precursor is derived from free phytol, released during chlorophyll degradation. Prenylation of HGA with PDP results in MPBQ production. This is the key step for tocopherol biosynthesis. Methylation of MPBQ yields DMPBQ. *Vte1* converts MPBQ and DMPBQ to δ- and γ- tocopherol, respectively. *Vte4* methylates δ- and γ-tocopherol to β- and α-tocopherol, respectively. DMGGBQ, 2,3-dimethyl-5-geranylgeranyl-1,4-benzoquinol; DMPBQ, 2,3-dimethyl-6-phytyl-1,4-benzoquinol; GGDP, geranylgeranyl diphosphate; GG-Chlorophyll, geranylgeranyl-chlorophyll a; HGA, homogentisic acid; HPP, hydroxyphenylpyruvate; MEP, methylerythritol 4-phosphate; MGGBQ, 2-methyl-6-geranylgeranyl-1,4-benzoquinol; MPBQ, 2-methyl-6-phytyl-1,4-benzoquinol; Phytyl-DP (PDP), phytyl diphosphate; Phytyl-P, phytyl monophosphate; *Hppd1*, p-hydroxyphenylpyruvate dioxygenase; *hggt1*, homogentisate geranylgeranyltransferase; *vte1*, tocopherol cyclase; *vte2*, homogentisate phytyltransferase; *vte3*, MPBQ/MGGBQ methyltransferase; *vte4*, γ-tocopherol methyltransferase; *vte5*, phytol kinase; *vte6*, phytyl phosphate kinase, *vte7* α/β-hydrolase family protein (created at BioRender.com). **b)** Bibliometric map of the keywords used in this article from 2014 to 2024. Data were collected from the ScienceDirect database using the keyword search: "cannabis, tocopherol". A shift toward tocopherol and hemp seed oil is evident.

and an activation domain (AD) that interacts with other regulatory proteins to activate or repress transcription. With the recent development of genomic tools, it has become clear that overexpressing pathway regulators does more than just induce genes that encode pathway enzymes. Because of their ability to control multiple pathway steps and cellular processes, necessary for metabolite accumulation, TFs offer much promise to manipulate metabolic pathways.[17] Synthetic zinc finger transcription factors (ZFP-TFs) were designed to upregulate the expression of the endogenous *Arabidopsis* γ-tocopherol methyltransferase gene. This gene encodes the enzyme responsible for the conversion of γ-tocopherol to α-tocopherol. Overexpression of ZFP-TF under the control of the seed-specific promoter caused a 20-fold increase in α-tocopherol compared to the control in *Arabidopsis* seeds.[18]

In tomato, the overexpression of a MYB, *ODORANT1*, led to a significant increase in phenylpropanoid content concomitantly with transcriptional upregulation of several SK pathway genes[19]. Recently, in a Genome-wide association study (GWAS) conducted by Kazemzadeh et al (2024), associations between genetic variations (SNPs) and traits related to tocopherol biosynthesis were analyzed, leading to the identification of the transcription factors *OsMADS59* and *WRKY39* in rice.[20]

To get a panoramic view of studies conducted in the last 10 years (2014-2024) on tocopherol biosynthesis in cannabis, we performed a co-occurrence analysis of high-frequency keywords to construct a knowledge map and visualize biblio metric networks[21] (Figure 1b). A shift toward tocopherol and hemp seed oil indicates a growing interest in these areas.

Here, a bioinformatic analysis was carried out, resulting in the identification of three candidate TFs (*AREB5*, *NAC072* and *PIF3*) possibly involved in vitamin E biosynthesis in *C. sativa*. Among these, *AREB5* showed changes in the expression profile during seed development stages in qRT-PCR.

Materials and Methods

Plant Materials

The field experiments were carried out on 29 April 2022 at the research greenhouse of Shahid Beheshti University (SBU, latitude 51° 29′, longitude 51◦39′), Tehran, Iran. *Cannabis sativa* ecotype Hamedan was grown without fertilization or chemical plant protection. Plot size was 1 × 2 m^2. Seeds were harvested in the middle of September 2022 every three weeks in 3 repetitions at milky, doughy, and mature developing stages. The germination stage was considered as the control. The seeds were wrapped in foil, placed in liquid nitrogen and stored at -80 ºC for further use.

Cannabis Oil Extraction

Oil was extracted from hemp seeds using n-hexane as the solvent following a standardized procedure.[22] Initially, 200 mg of hemp seeds were crushed in a crusher for 1 minute. The resulting powder was then mixed with n-hexane at a 2:1 (v/v) ratio. The mixture was vortexed for 1 minute and transferred to Falcon tubes (50 ml) wrapped in aluminum foil to protect from light exposure. The tubes were placed in a shaker incubator at 220 rpm and 26 °C for 16 hours. After shaking, the tubes were removed, and the supernatant was transferred to a new tube for solvent evaporation at room temperature (22 °C). The remaining oil (500 µl) was filtered through a PTFE filter (0.45 µm) and mixed with acetonitrile in a 1:1 ratio.

Measurement of Tocopherol Content

For tocopherol analysis, the method described by the Iranian National Standardization Organization (INSO.7211) was used with minor modifications. Tocopherol contents in the seed oils were determined by high-performance liquid chromatography (LC pump K-1001, KNAUER, Germany). The apparatus was coupled with a fluorescence detector set at an excitation wavelength of 295 nm and an emission wavelength of 330 nm. The mobile phase was 100% acetonitrile at a flow rate of 1 ml/min· The column temperature was set to 40 ºC and the injection volume was 20 ml. A C18 reverse phase column (GALAK, 150 mm length, 5 µm particle size, pH 2-11, 100 A˚, 330 m^2/g) was used to separate the tocopherols. The external standard method was used for the quantification of tocopherols in cannabis seeds under the same HPLC conditions. Each sample was analyzed in three repetitions to ensure accuracy and reproducibility. Amounts of tocopherol were calculated based on the established standard curve. The total tocopherol was a simple sum of all detected four individual major components. The unit of each isoform and the total content of tocopherol were expressed as micrograms of tocopherol per gram of seed (µg/g). This solvent system did not effectively separate the β-tocopherol content.[23, 24]

RNA Extraction and cDNA Synthesis

Cannabis seeds at different developmental stages (stage 1: germination, stage 2: milky, stage 3: doughy, stage 4: maturity) were used to isolate RNA. Total RNA at each sampling time was extracted using TRIzol (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. After treatment with *DNaseI* (Fermentas, MA, USA) to remove genomic DNA, the quantity and quality of the isolated RNA were determined using a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, USA) and agarose gel electrophoresis. First-strand cDNA was synthesized from total RNA using the Thermo Scientific RevertAid RT Kit (#K1691, Thermo Fisher Scientific). Synthesis of cDNAs was confirmed by PCR, followed by 1.5% agarose gel electrophoresis in TBE buffer. cDNA was kept at -20 ºC until used.

Identification of Candidate Genes

According to Mastour et al., 2019 in sunflower,[25] orthologous genes were identified (Table 2). Briefly, the sequences of key genes involved in the tocopherol biosynthesis pathway were extracted from the NCBI database for Sunflower. The 1500 nucleotides upstream of the start codon of each gene were considered as the promoter region, retrieved using the online Plantpan3 tool (http://plantpan.itps.ncku.edu.tw/) and transcription factors were identified using PlantTFDB5.0 (http://planttfdb.gao-lab.org/), a comprehensive plant transcription factor database. The protein-protein interaction network of the vitamin E biosynthesis pathway enzymes and transcription factors was visualized using STRING (Search Tool for the Retrieval of Interacting Genes - https://string-db.org/). Chromosome location and collinearity analysis of candidate TFs between *C. sativa*, *A. thaliana,* and *H. annuus* were obtained using the One-step MCScanX function of TBtools. Due to its simplicity and high efficiency in collinear block detection, MCScanX as a powerful tool for conducting diverse synteny and evolutionary analyses was used. These analyses offer valuable insights into the evolutionary relationships and functional conservation of TFs across different plant genomes.

Tissue-specific Expression of Candidate Genes

The expression profiles of the candidate genes in various tissues, including dry seed, leaf and shoot were obtained from the *Arabidopsis* eFP Browser database (https://bar.utoronto.ca/). A heatmap was generated to visualize the gene expression patterns across the different tissues.[26]

Analysis of Gene Expression by RT-qPCR

Primers were designed using Primer3Plus (https://www.primer3plus.com) in the last exon and 3´-UTR regions and synthesized by Pishgam (Tehran, Iran). The *PP2a* gene was used as the reference[27] (Table 1). The RT-qPCR was performed using a Magnetic Induction Cycler qPCR (Mic qPCR Cycler, Bio Molecular Systems, Queensland, Australia) and read using the Mic qPCR Cycler Software v.2.10.5 (Bio Molecular Systems, Queensland, Australia). For RT-qPCR analysis two technical replicates were used and the reaction mixture had 5 µl of SYBR Green II master mix (Applied Biosystems SYBR Green, Carlsbad, CA, USA), 0.7 µl each of forward and reverse primers (10 pmol each), and 240 ng cDNA (2.5 µl) to make the final reaction volume of 10 µl. The cycling program consisted of 94 ºC for 10 min and 40 cycles of [94 ºC: 30 s; 56 ºC: 30 s; 72 ºC: 10 s]. Relative gene expression was quantified using the $2^{-\Delta\Delta CT}$ method as described in Livak and Schmittgen (2001).[28] Two-way ANOVA calculated the statistical significance of Ct between different seed developmental stages with Tukey's post-hoc test comparisons using the GraphPad Prism 9.5.1 software. Pearson’s correlation coefficients (r) were used to assess the relationship between tocopherols and the relative expression of the candidate TF in Excel.

Table 1. Primers Used for Quantitative Real-time PCR.

Gene name	Primer sequence (5´→3´)	Tm (˚C)	Amplicon size (bp)
PP2A	F-AGCAACGTTCAGCCCGTTAA R-GACACTTCCCTCCAATTCGAAA	59	100
AREB5	F-CAGAGCCAAAGAACCAGC R-CCAAATACAAGTCCTGACCG	56	271
NAC072	F-ATCCCTCCTCTCTGCCAC R-GCACTCCCTACTTTGATTG	61	296
PIF3	F-CACCTATTCCTGGGCACTC R-CCTGAGTAGATGTCTAATTCGT	64	297

Results

Tocopherol Content

The seeds of the Hamedan ecotype (Figure 2a) were used to study tocopherol content. One month after flowering, the cannabis seeds were harvested every three weeks. Germination was considered the first step in measuring the amount of vitamin E (Figure 2b). The amounts of tocopherol at different stages of seed development (Figure 2c) were quantified using HPLC with a fluorescent detector. The retention time was between 10 15 min. Three tocopherol (T) isomers were detected, namely αT, γT, and δT (Figure 3a1-4). The predominant antioxidant in hemp seeds was γ-tocopherol (Figure 3-b). It was followed by α and δ tocopherol (γT: 7.10, αT: 3.98, δT: 1.41 µg/g). The highest amount of tocopherol was found in the mature seeds (stage 4) at the rate of 12.5 µg/g for total tocopherol (Figure 3-c).

Expression Analysis of Candidate TFs

The results of PPI network analysis showed that eight genes (Table 2) are co-expressed with the main genes in the tocopherol biosynthesis pathway, and they were selected as the final candidates. BLASTx retrieved orthologous TF genes in *C. sativa* and *Arabidopsis thaliana* using sequence data from *Helianthus annuus* (Table 2). The comparison of synteny analysis between *C. sativa*, *H. annuus,* and *A. thaliana* suggests a strong evolutionary connection (Figure 4).

To further verify the potential impact of the candidate genes on the regulation of tocopherol biosynthesis, the expression patterns of the candidate genes in various tissues (dry seeds, leaves and shoots) based on data from the *Arabidopsis* eFP Browser database was analyzed (Figure 5a). The highest expression of *NAC072, AREB5,* and *PIF3*

Figure 2. a) Cannabis Plant (*Hamedan ecotype*). **b)** Different Stages (1-germination, 2-milky, 3-doughy and 4-maturity) of seed development. **c)** Oil was isolated from hemp seeds by *n*-hexane. Color differences and droplet concentration between cannabis oil in four stages are clear. The germination stage (1) was considered as the control.

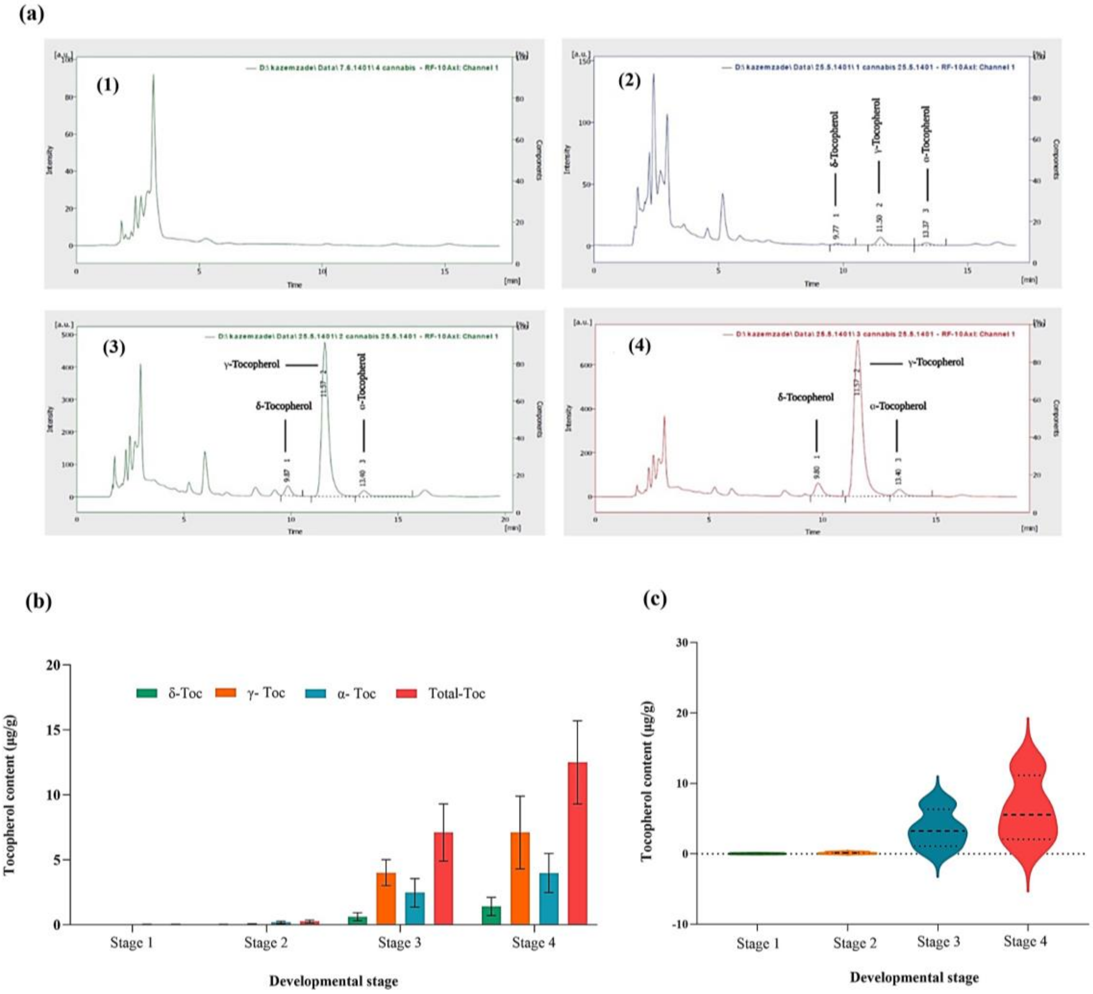


Figure 3. a. Chromatograms of the vitamin E homologs in different stages of cannabis seed development. Tocopherol contents in seed oils were analyzed using high-performance liquid chromatography (HPLC) with a fluorescence detector (excitation: 295 nm, emission: 330 nm). The mobile phase was 100% acetonitrile at a 1 ml/min flow rate. A C18 reverse-phase column (GALAK, 150 mm × 5 µm, pH 2–11, 100 E, 330 m^2/g) was used at 40 ºC with a 20 µl injection volume. **1)** stage 1 (germination), **2)** stage 2 (milky), **3)** stage 3 (doughy), **4)** stage 4 (maturity). **b.** Average of each isoform α, δ, γ, and total tocopherol at different stages of seed development **c.** Tocopherol average content in four developmental stages which reached its highest level at the maturation stage.

Table 2. Candidate TF Genes

TF name	TF family	*H. annuus*	Ortholog in *A. thaliana*	Ortholog in *C. sativa*
NAC072	NAC	*XM_022137133.2*	*AT4G27410*	*XM_030633854.1*
MYB308	MYB	*XM_022133204.2*	*AT4G38620*	*XM_030631923.1*
bZIP16	bZIP	*XM_022173311.2*	*AT2G35530*	*XM_030645360.1*
AREB5	bZIP	*XM_022132555.2*	*AT3G56850*	*XM_030645877.1*
WRKY50	WRKY	*XM_022125987.2*	*AT5G26170*	*XM_030645831.1*
MADS4	MIKC_MADS	*XM_022141472.2*	*AT1G24260*	*XM_030640317.1*
MYC2	bHLH	*XM_022113658.2*	*AT1G32640*	*XM_030646017.1*
PIF3	bHLH	*XM_022160286.2*	*AT1G09530*	*XM_030648734.1*

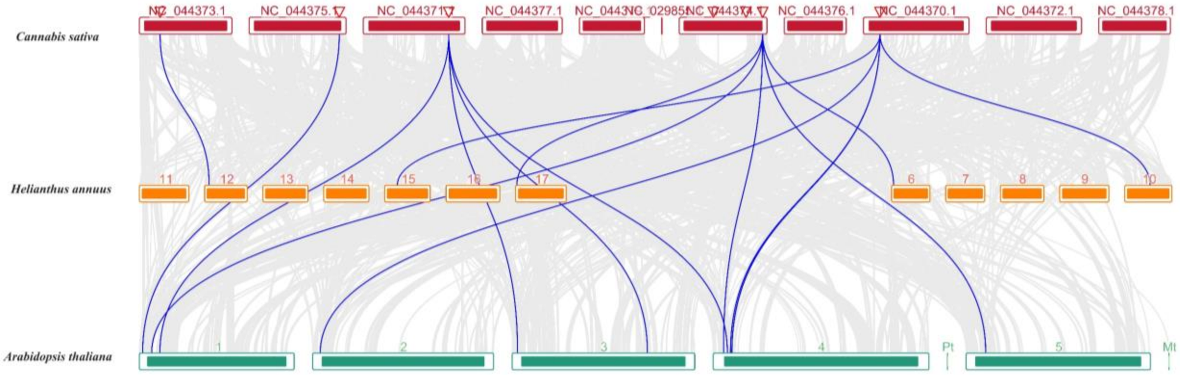


Figure 4. Collinearity analysis of the *NAC072, MYB308, bZIP16, AREB5, WRKY50, MADS4, MYC2,* and *PIF3* genes among *C. sativa, H. annuus* and *A. thaliana* species. Gray lines in the background indicate the collinear blocks of plant genomes. Blue-colored lines indicate the collinearity of genes.

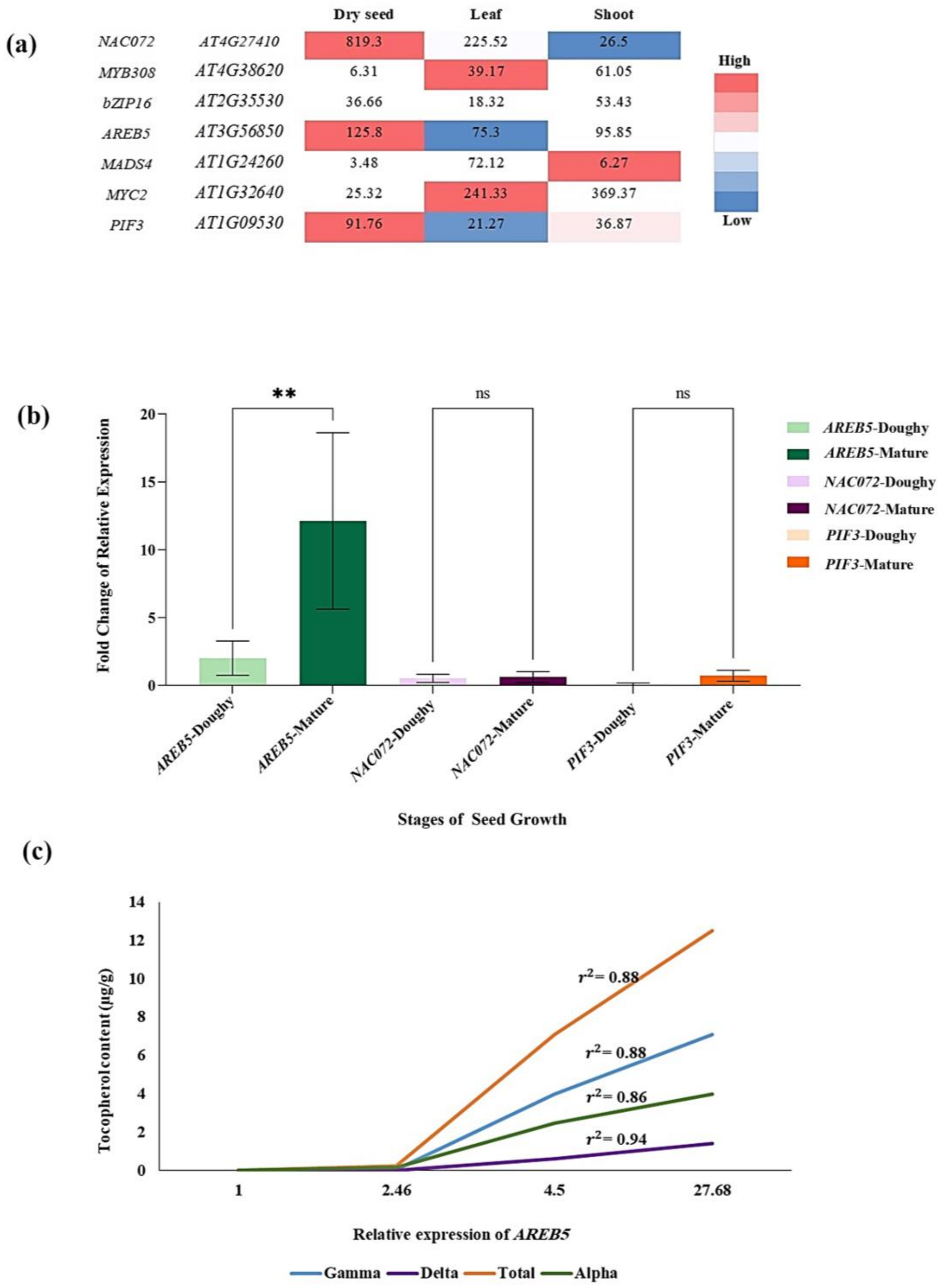
were in seeds. Based on these results, these genes were selected as the candidates for validation using real-time qRT-PCR.

The results obtained from the gene expression analyses (Figure 5b) showed that *AREB5* gene expression increased 27 times in the mature seeds (stage 4) compared to the germinating seed (stage 1- control stage). The increase in expression was 2 times in the green seed stage (stage 2) and 4.5 times in the third stage. The increase in expression was very significant in the maturation stage. In general, *AREB5* showed increased expression at the stage of seed maturity. The differential expressions of two other genes were not significant.

Additionally, a robust positive correlation between all isoforms of tocopherol and relative expression of *AREB5* was observed (α: $r^2 = 86$, γ and total: $r^2 = 0.88$, δ: $r^2 = 0.94$ δ - Figure 5c).

Discussion

Vitamin E is produced in the chloroplast envelope of plants and stored in the stroma. Its primary role is to quench reactive oxygen species in response to stress induced by salinity, drought, temperature, light, infection by pathogens, and the presence of heavy metals. It helps prevent lipid peroxidation by scavenging lipid peroxyl radicals in the thylakoid membranes of the chloroplast by donating a proton from its polar head. This protection from lipid peroxidation is crucial for maintaining photosynthetic membrane integrity.[29] Vitamin E also aids in stress tolerance, along with other antioxidants, such as ascorbate and glutathione. Tocopherols are lipid-soluble, whereas ascorbate and glutathione are water-soluble antioxidants. This antioxidant triad scavenges oxygen radicals in a coordinated manner in various phases within the cell.[30] Apart from being an essential dietary supplement, tocopherols have several physiological roles in treating cancer, cardiovascular, and ocular diseases.[31-33] Vitamin E has been reported to have a positive effect when taken as a supplement by specific populations, such as patients on hemodialysis or diabetic individuals with cardiovascular diseases[34] and aged people with respiratory diseases.[35] With the growing demand for natural cosmetic formulations, it is also one of the key ingredients in skin creams and hair formulations.[36] Furthermore, it is used as an additive in animal feeds because of its nutritive value.[37] With the growing need to combat lifestyle diseases and the inclination toward natural products, the market size of vitamin E is now on the rise. The global Vitamin E market size is predicted to grow to Approximately 3.66 billion USD by 2032 with a compound annual growth rate (CAGR) of around 5.3% between 2024 and 2032 (www.zionmarketresearch.com; Figure 6a). Photosynthetic organisms and chemical synthesis are the two common ways to obtain vitamin E to meet its rising demand. Genetic engineering, homologous/heterologous expression of functional or regulatory genes, and optimizing cultivation conditions have been shown to increase vitamin E content significantly. Some successful attempts can be enumerated.



(a)

		Dry seed	Leaf	Shoot
NAC072	*AT4G27410*	819.3	225.52	26.5
MYB308	*AT4G38620*	6.31	39.17	61.05
bZIP16	*AT2G35530*	36.66	18.32	53.43
AREB5	*AT3G56850*	125.8	75.3	95.85
MADS4	*AT1G24260*	3.48	72.12	6.27
MYC2	*AT1G32640*	25.32	241.33	369.37
PIF3	*AT1G09530*	91.76	21.27	36.87

Figure 5. a) Expression analysis of candidate TFs was conducted using publicly available RNA-Seq data for dry seeds, leaves, and shoots. Three genes, *NAC072*, *AREB5*, and *PIF3*, were selected as candidates for validation using real-time qRT-PCR. **b)** The expression analysis of *AREB5*, *NAC072*, and *PIF3* genes was performed using real-time qRT-PCR. Amplifications of cDNA from different developmental stages of seeds were carried out. A housekeeping gene, *PP2a*, was used as the control. Data were analyzed using a two-way ANOVA, followed by Tukey's multiple comparison test with a 95% confidence interval. Significantly different comparisons between doughy and maturity are displayed on graphs using asterisks (**, $p<0.01$; ns, not significant). Error bars represent mean ± SD (n = 2 biological replicates). **c)** The correlation between α, γ, δ, total tocopherol, and the relative expression of *AREB5* was examined. All r^2 values range from 0.86 to 0.94, indicating a strong correlation between the expression of the *AREB5* gene and tocopherol content in *C. sativa*.

For instance, overexpression of tocopherol cyclase caused an 18% increase in vitamin E content in maize and a 28% increase in *Arabidopsis*.[38] The overexpression of ZFP-TF under the control of the seed-specific promoter resulted in a 20-fold increase in α-tocopherol compared to the *Arabidopsis* control seeds.[18] Satish et al. (2018),[39] expressed HPT and TC genes from *Arabidopsis* in *Nicotiana benthamiana*. The levels of α-tocopherol in leaves expressing HPT, TC, and both genes increased by 4.2, 5.9, and 11.3 times, respectively, compared to the control. In tomato, the overexpression of a MYB, *ODORANT1*, led to a significant increase in phenyl propanoid content concomitantly with transcriptional up regulation of several SK pathway genes.[19] Downregulation of *DET1*, a suppressor of light responses in the absence of light in tomato fruit, resulted in increased contents of carotenoids, flavonoids, and tocopherol.[40] Also, in our recent study on the identification of regulatory elements by GWAS, we identified two TFs: *WRKY39* (*Os02g0265200*) and *MADS59* (*Os06 g0347700*) with possible regulatory roles in rice major tocopherol biosynthesis.[20] Although research on other plants has been conducted in previous years, studies on vitamin E in Cannabis are relatively recent. Figure 6b showed rather scarce plant metabolic engineering attempts to improve plant vitamin E content. This may be due to changes in research priorities and advances in the discovery of new medicinal and therapeutic properties.

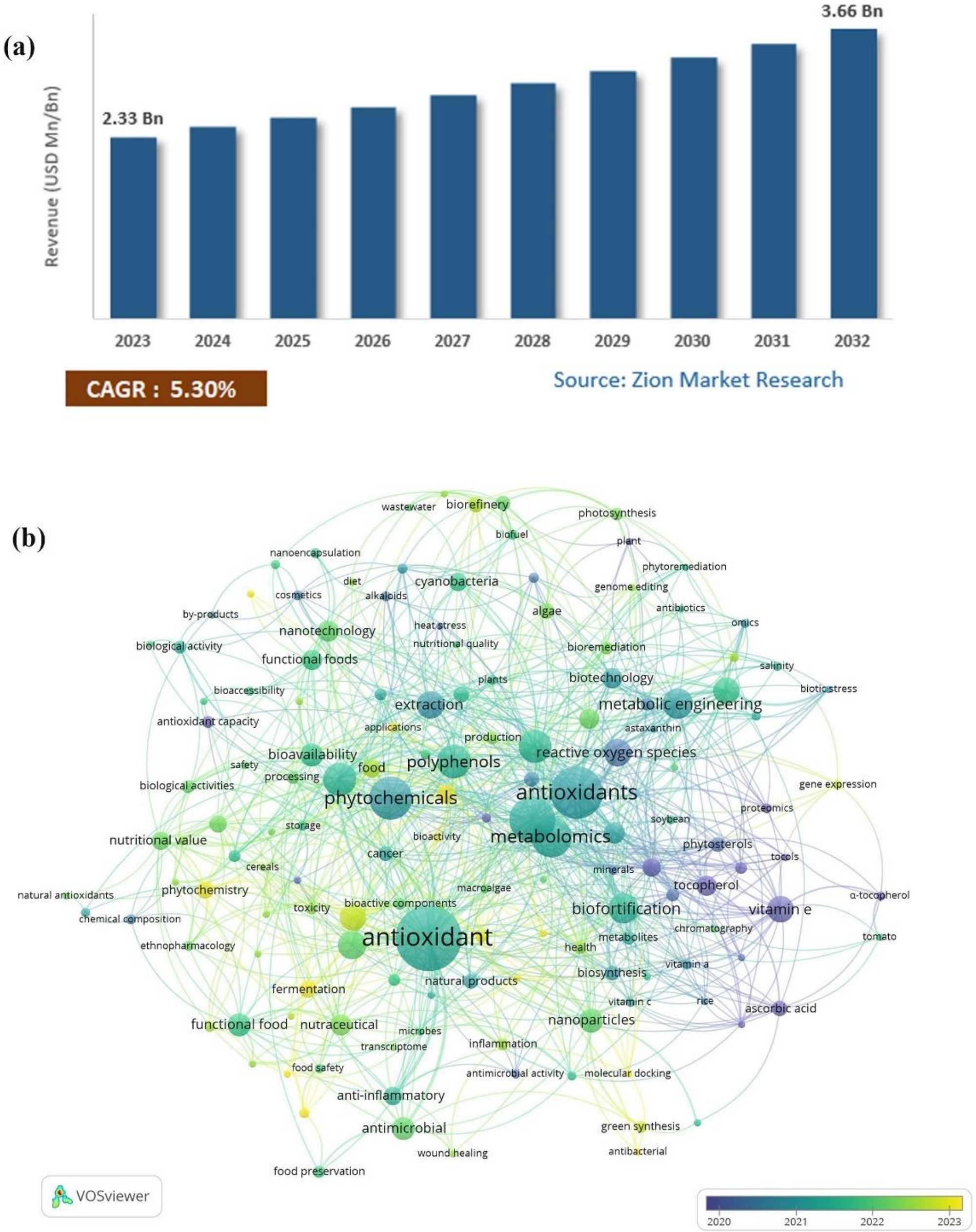


Figure 6. a) Global Vitamin E market from 2024 to 2032. The global vitamin E market size was approximately USD 2.33 billion in 2023 and is projected to reach around USD 3.66 billion by 2032, with a compound annual growth rate (CAGR) of approximately 5.3% between 2024 and 2032. **b)** Bibliometric map of the keywords used in this article from 2014 to 2024. Data were collected from the ScienceDirect database using the keyword search: "plant, tocopherol, vitamin E, metabolite engineering."

In this study, the total amount of tocopherols in hemp seed oil, a good source of vitamin E, was 12.5 µg/g, with γ-tocopherol as the predominant isomer accounting for 57% of the total amount of tocopherols at full seed maturity stage.

The other isomers were present at much lower levels: 31% for α and 11% for δ-tocopherol. The content of tocopherols (α-, β-, γ- and δ-) in hempseed oil ranges from 800 to 1500 ppm with a ratio of 0.5: 0.2: 9: 3, respectively[41, 42]. This is higher than those in common vegetable oils such as soybean, sunflower and canola at 600–1150 ppm.[43] The content of tocopherols may vary depending on the varieties, agronomic conditions, extraction methods and storage conditions[42]. Here, eight genes *Nac072* (NAC family), *MYB308* (MYB family), *bZIP16* (bZIP family), *AREB5* (bZIP), *WRKY50* (WRKY family), *MADS4* (MIKC_MADS family), *MYC2,* and *PIF3* (bHLH family) were selected as primary candidate TFs. Based on the expression analysis of candidate genes**,** three genes (*PIF3, AREB5,* and *NAC072*) were selected as the final candidate for RT-qPCR. Among these, *AREB5* of the AREB/ABF subfamily of group-A *bZIP* TFs[44] showed substantial changes during seed development. The *AREB5* ortholog in *Arabidopsis* is predominantly expressed during early silique development and is rarely expressed in vegetative tissues,[45,46] even under stress conditions.[47] A significant correlation between different isoforms of tocopherol and the relative expression of *AREB5* was noted in the present study. The highest correlation was observed for δT (r^2 = 0.94). Three candidate genes *PIF3, AREB5,* and *NAC072* were identified and checked by qRT-PCR analysis. *AREB5* showed substantial changes during seed development. Collectively, these findings provide genetic information for increasing/modulating tocopherol content in Cannabis through genetic engineering strategies. Further functional studies in the future, such as overexpression or knockout of *AREB5*, will help us better understand the molecular mechanisms of tocopherol biosynthesis in hemp.

Conclusion

Considering recent discoveries, *C. sativa* seeds have captured the attention of researchers and the food industry, evolving from being a byproduct of the industry to a primary source of nutrients and bioactive compounds. The health benefits and nutritional values of hemp seed oil are attributed to its unique fatty acid composition and bioactive minor components. These minor components, such as tocopherols, act as antioxidants, which can prevent the oil from deteriorating, and contribute to the health benefits for humans. Here, using fluorescent HPLC, the contents of vitamin E isomers (α, δ, and γ) were measured at four developmental stages of Cannabis seeds and a bioinformatic analysis was carried out that resulted in finding eight candidate TFs involved in vitamin E biosynthesis in *C. sativa*. Among these, *AREB5* showed changes in the expression profile of seed development stages in qRT-PCR. The study of TFs and other regulatory elements has benefits for targeted breeding and genetic engineering as they are responsible for controlling more than one metabolic pathway. This study provides fundamental information for breeding new hemp varieties with high tocopherol content and high-quality oils. Further functional studies, such as overexpression or knockout of these genes, will help us better understand the molecular mechanisms of tocopherol biosynthesis in hemp and other crops.

Authors' Contributions

SK performed experiments and prepared the first draft of the manuscript; NF proposed the research idea, supervised the research, performed data interpretation, and edited the manuscript; AA co-supervised the research, conducted *in silico* analyses, and contributed to data interpretation. All authors read and approved the submitted version.

Funding

Some financial support was from research grants in Shahid Beheshti University (Tehran, Iran).

Data Availability Statement

All relevant data can be found within the manuscript and its supporting materials and further inquiries can be directed to the corresponding author(s).

Conflict of Interest Disclosures

The authors declare that they have no conflicts of interest.

References

1. Long T, Wagner M, Demske D, Leipe C, Tarasov PE. Cannabis in Eurasia: origin of human use and Bronze Age trans-continental connections. Veget Hist Archaeobot. 2017;26(2):245-58. doi:10.1007/s00334-016-0579-6
2. McKernan KJ, Helbert Y, Kane LT, Ebling H, Zhang L, Liu B, et al. Sequence and annotation of 42 cannabis genomes reveals extensive copy number variation in cannabinoid synthesis and pathogen resistance genes. BioRxiv. 2020:2020-01. doi:10.1101/2020.01.03.894428
3. Van Bakel H, Stout JM, Cote AG, Tallon CM, Sharpe AG, Hughes TR, Page JE. The draft genome and transcriptome of *Cannabis sativa*. Genome Biol. 2011;12(10):R102. doi:10.1186/gb-2011-12-10-r102
4. Kiralan M, Gьl V, Kara SM. Fatty acid composition of hempseed oils from different locations in Turkey. Span J Agric Res. 2010;8(2):385-90. doi:10.5424/sjar/2010082-1220
5. Dąbrowski G, Skrajda MN. Lipid and protein fraction of hemp seed (*C. sativa* L.) and its beneficial influence on human health. J Educ Healt Sport. 2016;6(9):357-66. doi:10.5281/zenodo.62002
6. Da Porto C, Decorti D, Tubaro F. Fatty acid composition and oxidation stability of hemp (*Cannabis sativa* L.) seed oil extracted by supercritical carbon dioxide. Ind Crops Prod. 2012;36(1):401-4. doi:10.1016/j.indcrop.2011.09.015
7. Liang J, Appukuttan Aachary A, Thiyam-Holl∆nder U. Hemp seed oil: Minor components and oil quality. Lipid Technol. 2015;27(10):231-3. doi:10.1002/lite.201500050
8. Leger CL. La vitamine E: état actuel des connaissances, rôle dans la prévention cardio-vasculaire, biodisponibilité.

Oléagineux, Corps Gras, Lipides. 2000;7(3):258-65. doi:10.1051/ocl.2000.0258

9. Niki E. Lipid oxidation that is, and is not, inhibited by vitamin E: Consideration about physiological functions of vitamin E. Free Radic Biol Med. 2021;176:1-5. doi:10.1016/j.freeradbiomed.2021.09.001
10. Liang J, Appukuttan Aachary A, Thiyam-Holländer U. Hemp seed oil: Minor components and oil quality. Lipid Technol. 2015;27(10):231-3. doi:10.1002/lite.201500050
11. Goloshvili T, Akhalkatsi M, Badridze G, Kikvidze Z. Tocopherol contents and antioxidant activity in grape pomace after fermentation and alcohol distillation. Cell Mol Biol. 2021;67(1):112-5. doi:10.14715/cmb/2021.67.1.17
12. Xiao Y, Yu Y, Li G, Xie L, Guo X, Li J, et al. Genome-wide association study of vitamin E in sweet corn kernels. Crop J. 2020;8(2):341-50. doi:10.1016/j.cj.2019.08.002
13. Ghazani SM, Marangoni AG. Minor components in canola oil and effects of refining on these constituents: A review. J Am Oil Chem Soc. 2013;90(7):923-32. doi:10.1007/s11746-013-2254-8
14. Clarke MW, Burnett JR, Croft KD. Vitamin E in human health and disease. Crit Rev Clin Lab Sci. 2008;45(5):417-50. doi:10.1080/10408360802118625
15. Rébeillé F, Douce R. Biosynthesis of Vitamins in Plants Part B: Vitamins B6, B8, B9, C, E, K. Academic Press; 2011.
16. Muñoz P, Munne-Bosch S. Vitamin E in plants: biosynthesis, transport, and function. Trends Plant Sci. 2019;24(11):1040-51. doi:10.1016/j.tplants.2019.08.006
17. Broun P. Transcription factors as tools for metabolic engineering in plants. Curr Opin Plant Biol. 2004;7(2): 202-9. doi:10.1016/j.pbi.2004.01.013
18. Van Eenennaam AL, Li G, Venkatramesh M, Levering C, Gong X, Jamieson AC, et al. Elevation of seed α-tocopherol levels using plant-based transcription factors targeted to an endogenous locus. Metab Eng. 2004;6(2):101-8. doi:10.1016/j.ymben.2003.11.001
19. Dal Cin V, Tieman DM, Tohge T, McQuinn R, de Vos RC, Osorio S, et al. Identification of genes in the phenylalanine metabolic pathway by ectopic expression of a MYB transcription factor in tomato fruit. Plant Cell. 2011;23(7):2738-53. doi:10.1105/tpc.111.086975
20. Kazemzadeh S, Farrokhi N, Ahmadikhah A, Tabar Heydar K, Gilani A, Askari H, et al. Genome-wide association study and genotypic variation for the major tocopherol content in rice grain. Front Plant Sci. 2024;15:1426321. doi:10.3389/fpls.2024.1426321
21. Van Eck NJ, Waltman L. Crossref as a source of open bibliographic metadata. MetaArXiv. 2022. doi:10.31222/osf.io/smxe5
22. Kostić MD, Joković NM, Stamenković OS, Rajković KM, Milić PS, Veljković VB. Optimization of hempseed oil extraction by n-hexane. Ind Crops Prod. 2013;48:133-43. doi:10.1016/j.indcrop.2013.04.028
23. Sui M, Jing Y, Li H, Zhan Y, Luo J, Teng W, et al. Identification of loci and candidate genes analyses for tocopherol concentration of soybean seed. Front Plant Sci. 2020;11:539460. doi:10.3389/fpls.2020.539460
24. Graebner RC, Wise M, Cuesta-Marcos A, Geniza M, Blake T, Blake VC, et al. Quantitative trait loci associated with the tocochromanol (vitamin E) pathway in barley. PLoS One. 2015;10(7):e0133767. doi:10.1371/journal.pone.0133767
25. Mastour TN, TALEEI A, Farrokhi N, Naghavi MR, Ghaffari M. Identification of transcription factors involved in the transcriptional regulation of gamma tocopherol methyl transferase gene using bioinformatics tools. Mod Genet J. 2020;15(2):183-9.
26. Sashankar P, Hegde SN, Sathyanarayana N. Gene Identification and Structure Annotation. InBioinformatics in Rice Research: Theories and Techniques. Singapore: Springer Singapore. 2021. pp. 163-177. doi:10.1007/978-981-16-3993-7_8
27. Guo R, Guo H, Zhang Q, Guo M, Xu Y, Zeng M, et al. Evaluation of reference genes for RT-qPCR analysis in wild and cultivated *Cannabis*. Biosci Biotechnol Biochem. 2018;82(11):1902-10. doi:10.1080/09168451.2018.1506253
28. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods. 2001;25(4):402-8. doi:10.1006/meth.2001.1262
29. Munné-Bosch S. α-Tocopherol: a multifaceted molecule in plants. Vitam Horm. 2007;76:375-92. doi:10.1016/S0083-6729(07)76014-4
30. Jin S, Daniell H. Expression of γ-tocopherol methyl transferase in chloroplasts results in massive proliferation of the inner envelope membrane and decreases susceptibility to salt and metal-induced oxidative stresses by reducing reactive oxygen species. Plant Biotechnol J. 2014;12(9):1274-85. doi:10.1111/ pbi.12224
31. Jiang Q. Natural forms of vitamin E as effective agents for cancer prevention and therapy. Adv Nutr. 2017;8(6):850-67. doi:10.3945/an.117.016329
32. Fernández-Araque A, Giaquinta-Aranda A, Laudo-Pardos C, Rojo-Aragüés AA. Los antioxidantes en el proceso de patologнas oculares. Nutr Hosp. 2017;34(2):469-78. doi:10.20960/nh.420
33. Garg A, Lee JC. Vitamin E: Where are we now in vascular diseases?. Life. 2022;12(2):310. doi:10.3390/life12020310
34. Vardi M, Levy NS, Levy AP. Vitamin E in the prevention of cardiovascular disease: the importance of proper patient selection. J Lipid Res. 2013;54(9):2307-14. doi:10.1194/jlr.R026641
35. Khadangi F, Azzi A. Vitamin E–the next 100 years. IUBMB Life. 2019;71(4):411-5. doi:10.1002/iub.1990
36. Liu Y, Ilić T, Pantelic I, Savić S, Lunter DJ. Topically applied lipid-containing emulsions based on PEGylated emulsifiers: Formulation, characterization, and evaluation of their impact on skin properties *ex vivo* and *in vivo*. Int J Pharm. 2022;626:122202. doi:10.1016/j.ijpharm.2022.122202
37. Idamokoro EM, Falowo AB, Oyeagu CE, Afolayan AJ. Multifunctional activity of vitamin E in animal and animal products: A review. Anim Sci J. 2020;91(1):e13352. doi:10.1111/asj.13352
38. Kumar R, Raclaru M, Schüßeler T, Gruber J, Sadre R, Lühs W, et al. Characterisation of plant tocopherol cyclases and their overexpression in transgenic *Brassica napus* seeds. FEBS Lett. 2005;579(6):1357-64. doi:10.1016/j.febslet.2005.01.030
39. Sathish S, Preethy KS, Venkatesh R, Sathishkumar R. Rapid enhancement of α-tocopherol content in *Nicotiana benthamiana* by transient expression of *Arabidopsis thaliana* Tocopherol cyclase and Homogentisate phytyl transferase genes. 3 Biotech. 2018;8(12):485. doi:10.1007/s13205-018-1496-4
40. Enfissi EM, Barneche F, Ahmed I, Lichtlé C, Gerrish C, McQuinn RP, et al. Integrative transcript and metabolite analysis of nutritionally enhanced DE-ETIOLATED1 downregulated tomato fruit. Plant Cell. 2010;22(4):1190-215. doi:10.1105/tpc.110.073866
41. Matthдus B, Brьhl L. Virgin hemp seed oil: An interesting niche product. Eur J Lipid Sci Technol. 2008;110(7):655-

61. doi:10.1002/ejlt.200700311

42. Liang J, Appukuttan Aachary A, Thiyam-HollΔnder U. Hemp seed oil: Minor components and oil quality. Lipid Technol. 2015;27(10):231-3. doi:10.1002/lite.201500050

43. Ghazani SM, Marangoni AG. Minor components in canola oil and effects of refining on these constituents: A review. J Am Oil Chem Soc. 2013;90(7):923-32. doi:10.1007/s11746-013-2254-8

44. Fujita Y, Yoshida T, Yamaguchi-Shinozaki K. Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. Physiol Plant. 2013;147(1):15-27. doi:10.1111/j.1399-3054.2012.01635.x

45. Bensmihen S, Giraudat J, Parcy F. Characterization of three homologous basic leucine zipper transcription factors (bZIP) of the ABI5 family during *Arabidopsis thaliana* embryo maturation. J Exp Bot. 2005;56(412): 597-603.

46. Kim HS, Delaney TP. Over-expression of TGA5, which encodes a bZIP transcription factor that interacts with NIM1/NPR1, confers SAR-independent resistance in *Arabidopsis thaliana* to *Peronospora parasitica*. Plant J. 2002;32(2):151-63. doi:10.1046/j.1365-313X.2001.01411.x

47. Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM, Seki M, et al. AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in *Arabidopsis*. Plant Cell. 2005;17(12): 3470-88. doi:10.1105/tpc.105.035659