

Article

QTL and Candidate Genes for Seed Tocopherol Content in 'Forrest' by 'Williams 82' Recombinant Inbred Line (RIL) Population of Soybean

Dounya Knizia ^{1,2}, Jiazheng Yuan ³, Naoufal Lakhssassi ¹ , Abdelhalim El Baze ¹, Mallory Cullen ¹, Tri Vuong ⁴, Hamid Mazouz ² , Henry T. Nguyen ⁴ , My Abdelmajid Kassem ³  and Khalid Meksem ^{1,*} 

¹ School of Agricultural Sciences, Southern Illinois University, Carbondale, IL 62901, USA; dounya.knizia@siu.edu (D.K.); naoufal.lakhssassi@siu.edu (N.L.); abdelhalim.elbaze@siu.edu (A.E.B.); cullenmallory@gmail.com (M.C.)

² Laboratoire de Biotechnologies & Valorisation des Bio-Ressources (BioVar), Department de Biologie, Faculté des Sciences, Université Moulay Ismail, Meknes 50000, Morocco; h.mazouz@fs-umi.ac.ma

³ Plant Genomics and Biotechnology Laboratory, Department of Biological and Forensic Sciences, Fayetteville State University, Fayetteville, NC 28301, USA; jyuan@uncfsu.edu (J.Y.); mkassem@uncfsu.edu (M.A.K.)

⁴ Division of Plant Science and Technology, University of Missouri, Columbia, MO 65211, USA; vuongt@missouri.edu (T.V.); nguyenhenry@missouri.edu (H.T.N.)

* Correspondence: meksem@siu.edu



Citation: Knizia, D.; Yuan, J.; Lakhssassi, N.; El Baze, A.; Cullen, M.; Vuong, T.; Mazouz, H.; T. Nguyen, H.; Kassem, M.A.; Meksem, K. QTL and Candidate Genes for Seed Tocopherol Content in 'Forrest' by 'Williams 82' Recombinant Inbred Line (RIL) Population of Soybean. *Plants* **2022**, *11*, 1258. <https://doi.org/10.3390/plants11091258>

Academic Editor: Matthew Nelson

Received: 4 April 2022

Accepted: 3 May 2022

Published: 6 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Soybean seeds are rich in secondary metabolites which are beneficial for human health, including tocopherols. Tocopherols play an important role in human and animal nutrition thanks to their antioxidant activity. In this study, the 'Forrest' by 'Williams 82' (F×W82) recombinant inbred line (RIL) population ($n = 306$) was used to map quantitative trait loci (QTL) for seed α -tocopherol, β -tocopherol, δ -tocopherol, γ -tocopherol, and total tocopherol contents in Carbondale, IL over two years. Also, the identification of the candidate genes involved in soybean tocopherols biosynthetic pathway was performed. A total of 32 QTL controlling various seed tocopherol contents have been identified and mapped on Chrs. 1, 2, 5, 6, 7, 8, 9, 10, 12, 13, 16, 17, and 20. One major and novel QTL was identified on Chr. 6 with an R^2 of 27.8, 9.9, and 6.9 for δ -tocopherol, α -tocopherol, and total tocopherol content, respectively. Reverse BLAST analysis of the genes that were identified in *Arabidopsis* allowed the identification of 37 genes involved in soybean tocopherol pathway, among which 11 were located close to the identified QTLs. The tocopherol cyclase gene (*TC*) *Glyma.06G084100* is located close to the QTLs controlling δ -tocopherol ($R^2 = 27.8$), α -tocopherol ($R^2 = 9.96$), and total-tocopherol ($R^2 = 6.95$). The geranylgeranyl diphosphate reductase (*GGDR*) *Glyma.05G026200* gene is located close to a QTL controlling total tocopherol content in soybean ($R^2 = 4.42$). The two methylphytylbenzoquinol methyltransferase (*MPBQ-MT*) candidate genes *Glyma.02G002000* and *Glyma.02G143700* are located close to a QTL controlling δ -tocopherol content ($R^2 = 3.57$). The two γ -tocopherol methyltransferase (γ -*TMT*) genes, *Glyma.12G014200* and *Glyma.12G014300*, are located close to QTLs controlling (γ + β) tocopherol content ($R^2 = 8.86$) and total tocopherol ($R^2 = 5.94$). The identified tocopherol seed QTLs and candidate genes will be beneficial in breeding programs to develop soybean cultivars with high tocopherol contents.

Keywords: soybean; RIL; forrest; Williams 82; linkage map; tocopherol; SNP

1. Introduction

Tocopherols and tocotrienols collectively constitute the tocochromanols family, known as Vitamin E. Tocochromanols are fat-soluble phenolic compounds, synthesized by photosynthetic organisms. In soybean, vitamin E is present almost exclusively as tocopherols. Tocopherols exist in four isoforms, α -tocopherol, γ -tocopherol, β -tocopherol, and δ -tocopherol

which differ from each other by the number and the location of the methyl groups. α -tocopherol possesses three methyl groups, followed by γ -tocopherol and β -tocopherol that have two methyl groups, and finally δ -tocopherol with only one methyl group [1].

Tocopherols have an important role in human and animal nutrition thanks to their vitamin E activity. However, from a nutritional perspective among the four tocopherol isoforms, α -tocopherol is the most important due to the high vitamin E activity [2]. It also has been reported to play a role in the prevention of cardiovascular diseases and cancer [3,4]. In the human body, α -tocopherol is preferentially accumulated due to its affinity with the liver α -tocopherol transfer protein (α -TTP), which enriches plasma with α -tocopherol [5].

Soybean (*Glycine max* Merr.) is not only one of the main sources of vegetable oil and animal feed worldwide, but also used for production of biofuel, aquaculture feed, and as source of protein for the human diet due to a high protein (40–42%) and oil contents (18–22%) [6], which make it an important crop worldwide.

Soybean seeds are rich in secondary metabolites beneficial for human health including tocopherols. Total tocopherol content is relatively high in soybean seeds compared to other oilseeds crops, and γ -tocopherol is the predominant form, while α -tocopherol content is less than 10% of the total tocopherol content [3,4,7].

Since soybean oil provides 30% of the total worldwide oil consumption and ~70% of the vitamin E in the American diet comes from soybean oil, developing soybean cultivars with high seed α -tocopherol contents could have tremendous positive effects on the health benefits associated with eating soybeans and their market value.

In soybean, tocopherol seed content and composition vary from one cultivar to another, being controlled by several genetic and environmental factors. These factors make it one of the most complex quantitative traits [8] and many studies have focused on investigating the genetic and molecular factors underlying this trait [9–12].

A 'TK780' by 'B04009' RIL population ($n = 94$) was used to identify six QTLs involved in α -tocopherol biosynthesis on Chr. 9, 11 and 12 [7]. Also, using a 'OAC Bayfield' \times 'OAC Shire' RIL population across three locations over 2 years, and 151 SSR markers, 26 SSR markers linked to QTLs with individual and total tocopherol content across 17 chromosomes were identified [13]. Likewise, using a 'Beifeng 9' by 'Freeborn' RIL population in six environments, 18, 13, 11, and 13 QTLs associated with seed α -tocopherol, γ -tocopherol, δ -tocopherol, and total tocopherol contents were identified respectively [14].

Tocopherol biosynthesis takes place at the plastid's envelope, where a combination of two precursors derived from different pathways occurs. The homogentisic acid (HGA), a product of the cytosolic shikimate pathway, is used to form the aromatic ring of tocopherols, while phytyl diphosphate (PDP), a product of either the methylerythritol phosphate (MEP) pathway or the phytol recycling pathway [15], forms the prenyl tail. The condensation of these two precursors is catalyzed by the homogentisate phytyl transferase (HPT) and creates 2-methyl-6-phytyl-1,4-benzoquinol (MPBQ), which can be further methylated by MPBQ methyltransferase (MPBQ-MT) to 2,3-dimethyl-6-phytyl-1,4-benzoquinone (DMPBQ). The cyclization of the MPBQ and DMPBQ by the tocopherol cyclase produces γ -tocopherol and δ -tocopherol, respectively. The conversion of γ -tocopherol and δ -tocopherol to α -tocopherol and β -tocopherol is catalyzed by the γ -tocopherol methyltransferase (γ -TMT) and it represents the last step of the tocopherol biosynthesis pathway.

Many studies have elucidated the genes involved in the tocopherol pathway in *Arabidopsis*. The tyrosine produced by the shikimate pathway is catalyzed by Tyrosine Aminotransferase (TAT) resulting in the formation of P-hydroxy Phenyl Pyruvate (HPP). The HPP will be catalyzed by p-hydroxyphenylpyruvate dioxygenase (HPPD) to produce the homogentisic acid, this enzyme is controlled by PDS1. In *A. thaliana*, mutants of *pds1* have shown a lack of tocopherols and plastoquinone with a lethal photobleached phenotype, this result showed the importance of PDS1 in the tocopherol biosynthesis pathway [16]. The overexpression of the *PDS1* gene in tobacco leaves or in *A. thaliana* seeds only gave moderately increased tocopherol concentrations [17,18]. The phytyl diphosphate (PDP) can be derived, either from the MEP pathway after reduction of geranylgeranyl diphosphate (GGDP) by the

Geranylgeranyl Diphosphate Reductase (GGDR) enzyme, or from the phytol recycling pathway. Many studies have investigated the phytol recycling pathway and have shown that mutants of *vte5-1* are devoid of phytol kinase. Also, *vte5-1* mutants have shown a reduction in total tocopherol content in seeds and leaves with 80% and 65% respectively, compared to the wild type [19]. The *VTE5* gene controls the phytol kinase that catalyzes the phytol phosphorylation producing Phytolmonophosphate which is catalyzed by Phytolphosphate kinase *VTE6* leading to phytyl diphosphate (PDP) formation. Arabidopsis *vte6* mutants have shown tocopherol deficiency in leaves and a reduction in plant growth and seed longevity. The overexpression of the *VTE6* gene resulted in a two-fold increase in PDP that resulted in higher γ -tocopherol accumulation in seeds [20]. Homogentisate phytyl transferase (HPT) catalyzes the condensation of HGA and PDP to produce 2-methyl-6-phytyl-1,4- benzoquinone (MPBQ). In *Arabidopsis*, the HPT enzyme is encoded by the *VTE2* gene [21,22].

The *Arabidopsis* devoid of *VTE2* have shown a complete deficiency in all tocopherol derivatives and all pathway precursors, which means that this is a crucial step in the tocopherol biosynthetic pathway [23]. The MPBQ-MT enzyme is encoded by the *VTE3* gene, which is a limiting step in producing α - and γ -tocopherol. In *Arabidopsis*, *vte3-2* mutants were lacking in α - and γ -tocopherol and exhibited a pale green phenotype, abnormal chloroplasts and did not survive beyond the seedling stage [24,25].

The *VTE1* gene catalyzes the conversion of 2-methyl-6-phytyl-1,4- benzoquinone (MPBQ) and 2,3-dimethyl-5-phytyl-1,4-benzoquinone (DMPBQ) to δ -tocopherol and γ -tocopherol, respectively. *vte1* mutants have a nonfunctional tocopherol cyclase enzyme (TC) and are totally devoid of all tocopherol forms, and accumulate DMPBQ, the γ -tocopherol precursor [26]. The overexpression of the *VTE1* gene in *Arabidopsis* plants showed an improvement in total tocopherol by 7-fold in leaves, as well as a major shift from α - to γ -tocopherol [27].

The *VTE4* gene encodes the γ -tocopherol methyltransferase (γ -TMT) that catalyzes the methylation of the γ -tocopherol and δ -tocopherol to produce α -tocopherol and β -tocopherol, respectively. The co-expression of both *At-VTE3* and *At-VTE4* in soybean showed an accumulation of >95% of α -tocopherol, in addition to a 5-fold increase of seed vitamin E activity [28].

In soybean, few genes have been reported to play a role in seed tocopherol content. These genes are γ -TMT1, γ -TMT2, and γ -TMT3 mainly [29]. Also, the Di-glucose binding protein with Leucine-rich repeat domain gene (*Glyma.02G099800*); Eukaryotic aspartyl protease family protein gene (*Glyma.02G100800*); Cytochrome b561 /ferric reductase transmembrane protein family gene (*Glyma.02G101300*); Transmembrane amino acid transporter family protein gene (*Glyma.02G098200*); and plant U-box 26 gene (*Glyma.02G102900*) were identified and determined to be significantly associated with α -tocopherol, γ -tocopherol, δ -tocopherol and total tocopherol in soybean seeds [30].

In this study, the genetic factors associated with tocopherol content in soybean were investigated, QTL for seed α -tocopherol, β -tocopherol, δ -tocopherol, γ -tocopherol, and total tocopherol contents were mapped, the link between the biosynthesis genes for tocopherol and soybean seed tocopherol content was studied, and the *in-Silico* tocopherol biosynthetic pathway in soybean was reconstructed.

2. Results

2.1. The SNP-Based Genetic Map

The SNP-Based genetic map used in this study was described previously and identified QTLs that control seed isoflavone contents [31]. The map which covered 4029.9 cM, was composed of 2075 SNP markers, and was based on 306 RILs of F × W82 [31].

2.2. Tocopherol Contents Frequency Distribution, Heritability, and Correlation

The frequency distributions among different tocopherol contents were not always normal in the FxW82 RIL population based on Shapiro–Wilk’s method for normality test.

Only total-tocopherol 2017 (T-Toc-2017) and δ -tocopherol 2020 (δ -Toc-2020) were normally distributed. The positive or negative skewness and kurtosis value (>3) were also identified in the RIL population (Table 1; Figure 1).

Table 1. Mean, range, CV (%), skewness, kurtosis, and value of Shapiro-Wilk normality test (W value) for seed tocopherol content of the RILs in Carbondale, IL. SE: Standard error.

Year	Trait	Mean ($\mu\text{g/g}$)	Range ($\mu\text{g/g}$)	CV (%)	SE	Skewness	Kurtosis	W Value ($p < 0.05$)
2017	δ -tocopherol17	95.86	94	19.73	1.09	0.01	2.21	0.98 **
	γ + β -tocopherol17	172.44	97.1	9.78	0.97	0.48	3.18	0.98 **
	α -tocopherol17	4.94	40.7	83.64	0.24	4.46	32.89	0.52 ***
	Total-tocopherol17	271.5	150.7	8.97	1.41	0.25	2.9	0.99
2020	δ -tocopherol20	94.5	76	14.67	0.81	0.2	3.1	0.99
	γ + β -tocopherol20	180.67	149.2	10.06	1.06	-0.36	5.59	0.97 ***
	α -tocopherol20	5.13	25.7	55.49	0.16	3.13	21.28	0.71 ***
	Total-tocopherol20	279.44	206.4	9.13	1.48	0.16	4.32	0.98 **

** $p < 0.01$, *** $p < 0.001$.

Each tocopherol component also showed a different degree of variation in the parameters of traits, and the variability appeared to not be greatly impacted by different environments. α -tocopherol 2017 (α -Toc-2017), displayed the highest coefficient of variation (CV) value (83.64%); however, the CV of (α -Toc-2020) was 55.49% indicating that phenotypic variability among tocopherol contents was constant year over these 2 years.

The broad sense heritability (h^2) of ($\mu\text{g/g}$ of dry seed weight) for seed α -tocopherol (α -Toc), δ -tocopherol (δ -Toc), γ + β -tocopherol ($(\gamma+\beta)$ -Toc), and total tocopherol (T-Toc) contents (in $\mu\text{g/g}$ of dry seed weight) across two different years appeared to be quite diverse. δ -Toc had the highest heritability (71%) and the h^2 for T-Toc was 41% (Table 2). However, the h^2 values for $(\gamma+\beta)$ -Toc and α -Toc were negative (-41% and -61%, respectively) implying that there was biologically meaningful phenotypic repulsion among these traits. The high heritability of seed δ -Toc contents suggested that a large portion of phenotypic variation could be detected in the mapped QTL. The RILs-Year interactions still played a significant role in the molecular formation among tocopherols in soybean seeds based on our two-way ANOVA analysis because the σ_{GE}^2 is relatively high (data not shown). It should be used as a parameter for trait improvement.

Due to cost effect of this undergraduate student-centered project, only technical replicates could be applied, and F value and probability could not be generated from the dataset (Table 2). Hence, we only calculated the Sum Sq and Mean Sq to determine σ_G^2 and σ_{GE}^2 for each trait (Table 2) using type I sum of squares (ANOVA (model)) function in R program but not σ_e^2 due to limited replicates.

2.3. Seed Tocopherol Contents QTL

We used both the interval mapping (IM) and composite interval mapping (CIM) methods of WinQTL Cartographer 2.5 [32] to identify QTLs that control seed α -Toc, δ -Toc, $(\gamma+\beta)$ -Toc, and T-Toc contents; however, only QTLs identified by CIM method with LOD scores >2.5 are reported here. A total of 32 QTL that control these seed tocopherols contents have been identified and mapped on Chr. 1, 2, 5, 6, 7, 8, 9, 10, 12, 13, 16, 17, and 20 in this RIL population grown in both years (2017 and 2020) (Table 3, Figures S1 and S2).

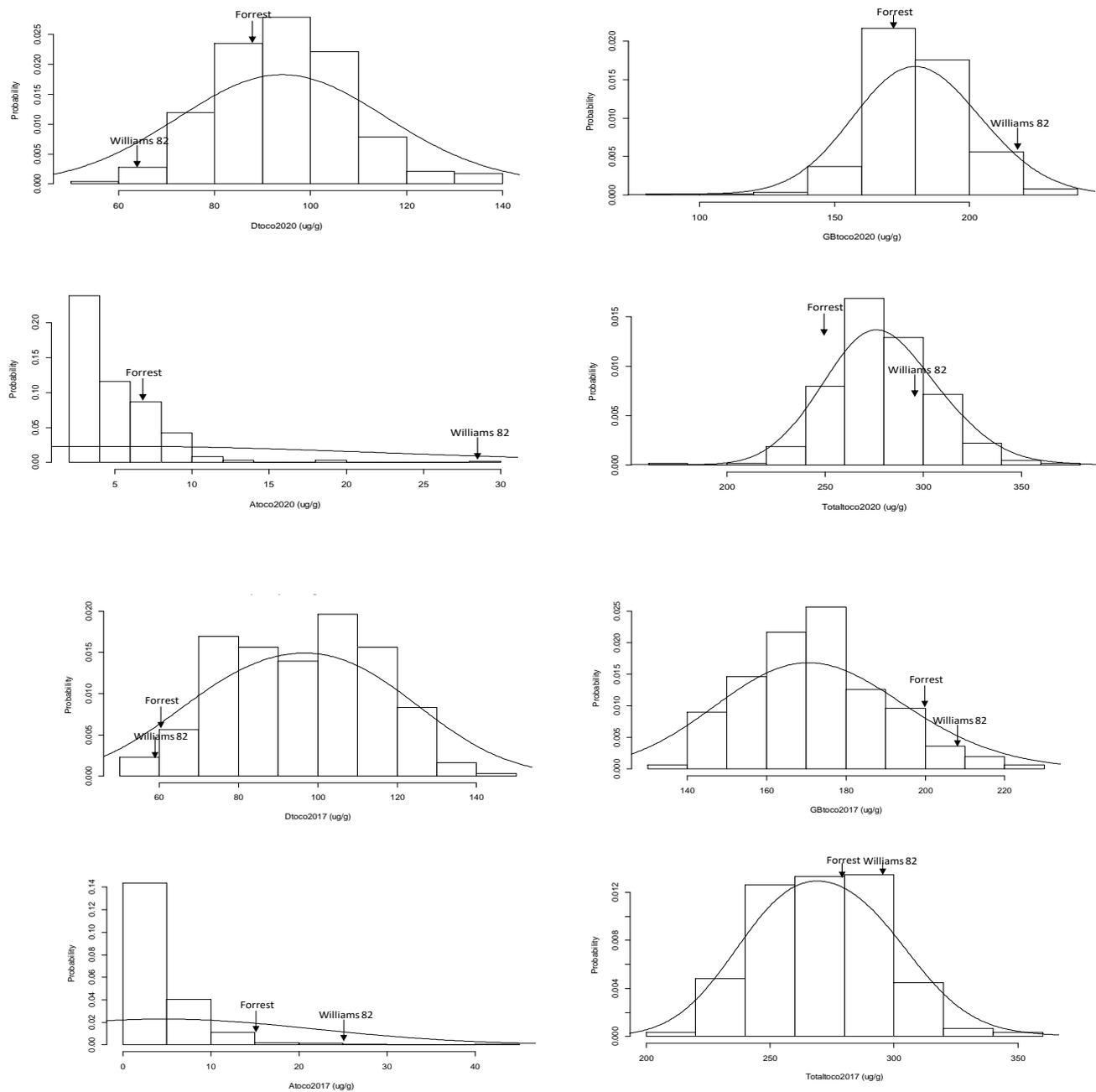


Figure 1. The distribution of seed tocopherol contents ($\mu\text{g/g}$ of seed weight) in the FxW82 RIL population. The seed α -Tocopherol (α -Toc), δ -Tocopherol (δ -Toc), $(\gamma+\beta)$ -Tocopherol ($(\gamma+\beta)$ -Toc), and Total-Tocopherols (T-Toc) contents were tested in the RILs harvested in Carbondale, IL 2017 and 2020, respectively.

Table 2. The broad sense heritability (h^2) of tocopherol traits (δ -tocopherol, $(\gamma+\beta)$ -tocopherol, α -tocopherol, and total-tocopherol) from the seeds harvested at Carbondale, IL in 2017 and 2020.

Response: δ -tocopherol				
	Df	Sum Sq	Mean Sq	H^2
Line	592	163,429	276.06	0.71
Year	1	391	391.28	
Line:Year	1	81	80.65	
Residuals	0	0	NA	

Table 2. Cont.

Response: γ + β -tocopherol				
	Df	Sum Sq	Mean Seq	H ²
Line	592	191,945	324.23	-0.41
Year	1	6	6.13	
Line:Year	1	457	456	
Residuals	0	0	NA	
Response: α -tocopherol				
	Df	Sum Sq	Mean Seq	H ²
Line	592	7479.9	12.635	-0.61
Year	1	20	20.041	
Line:Year	1	20	20.41	
Residuals	0	0	NA	
Response: Total-tocopherol				
	Df	Sum Sq	Mean Seq	H ²
Line	592	377,872	638.3	0.47
Year	1	205	205.28	
Line:Year	1	338	337.55	
Residuals	0	0	NA	

Table 3. QTLs that control seed α -Tocopherol (α -Toc), δ -Tocopherol (δ -Toc), (α + β)-Tocopherol ((α + β)-Toc), and Total-Tocopherols (T-Toc) contents in two environments over two years (A. 2017 and B. 2020). The two environments are in Carbondale, IL (2017) (A) and (2020) (B). Only solid QTL with LOD scores >2.5 and identified by CIM are reported.

A. QTL that Control Seed Tocopherols Contents in Carbondale, IL (2017)								
Trait	QTL	Chr.	Marker/Interval	Position (cM)	LOD	R ²	Additive	Environment
α -Tocopherol	q α -Toc-1	6	Gm06_1537675- Gm06_1570293	173.7–178.7	6.1	9.96	1.648195	Carbondale, IL
	q α -Toc-2	6	Gm06_1858327- Gm06_2048675	192.6–197.6	6.77	9.95	1.477826	Carbondale, IL
δ -Tocopherol	q δ -Toc-1	1	Gm01_1887205- Gm01_1653315	174.2–179.2	3.06	3.1	-3.30536	Carbondale, IL
	q δ -Toc-2	2	Gm02_1481798- Gm02_9925870	133.5–140.2	3.4	3.57	5.481172	Carbondale, IL
	q δ -Toc-3	6	Gm06_1674534- Gm06_4447485	183.8–207	23.01	27.9	10.14229	Carbondale, IL
γ + β -Tocopherol	q γ + β -Toc-1	6	Gm06_1674534- Gm06_4368839	185.8–203.2	5.13	6.16	4.161084	Carbondale, IL
	q γ + β -Toc-2	8	Gm08_3018731- Gm08_4266625	17.8–31.2	3.02	3.78	-3.23578	Carbondale, IL
	q γ + β -Toc-3	12	Gm12_3820261- Gm12_3818392	0.5–1	4.14	5.23	3.852091	Carbondale, IL
	q γ + β -Toc-4	12	Gm12_3805393- Gm12_3696093	2.5–18.5	7.18	8.86	5.050705	Carbondale, IL
	q γ + β -Toc-5	13	Gm13_2587196- Gm13_2048499	189.1–210.7	3.79	5.43	3.906039	Carbondale, IL
Total-Tocopherols	qTotal-Toc-1	5	Gm05_3674925- Gm05_3256515	29.4–32.2	3.63	4.42	8.695197	Carbondale, IL
	qTotal-Toc-2	6	Gm06_1739930- Gm06_2073990	188.9–197.6	4.07	5.05	-5.69939	Carbondale, IL
	qTotal-Toc-3	6	Gm06_3849946- Gm06_4447485	200.1–207	5.67	6.95	-6.57105	Carbondale, IL
	qTotal-Toc-4	7	Gm07_3635708- Gm07_1829304	81.4–88.9	2.88	3.46	-5.31938	Carbondale, IL
	qTotal-Toc-5	9	Gm09_3483063- Gm09_3544488	74.8–78	3.13	3.78	-4.56753	Carbondale, IL
	qTotal-Toc-6	12	Gm12_3820261- Gm12_3696093	2.5–18.5	4.84	5.94	4.753021	Carbondale, IL

Table 3. Cont.

B. QTLs that control seed tocopherols contents in Carbondale, IL (2020)								
Trait	QTL	Chr.	Marker	Position (cM)	LOD	R2	Additive	Envt.
α-Tocopherol	qα-Toc-1	1	Gm01_3466825- Gm01_5255151	4.1–10.1	5.81	0.35	2.35	Carbondale, IL
	qα-Toc-2	2	Gm02_5141136- Gm02_1020061	137.1–139.8	2.9	0.04	0.82	Carbondale, IL
	qα-Toc-3	6	Gm06_1954068- Gm06_2015292	195–197	2.01	0.03	0.5	Carbondale, IL
δ-Tocopherol	qδ-Toc-1	1	Gm01_4912170- Gm01_4852475	91.6–93.6	2.57	0.04	−2.66	Carbondale, IL
	qδ-Toc-2	8	Gm08_1810148- Gm08_2201336	125.7–130.8	2.42	0.04	2.86	Carbondale, IL
	qδ-Toc-3	10	Gm10_3943637- Gm10_3935014	79.2–81.4	2.4	0.03	−2.53	Carbondale, IL
	qδ-Toc-4	16	Gm16_1079308- Gm16_3673245	0.5–12.5	3.71	0.05	−6.65	Carbondale, IL
	qδ-Toc-5	20	Gm20_3665142- Gm20_1046460	174.9–176.9	2.68	0.04	−5.95	Carbondale, IL
(γ+β)- Tocopherol	q(γ+β)-Toc-2	2	Gm02_5155733- Gm02_4311734	130.5–132.5	2.15	0.04	−9.75	Carbondale, IL
	q(γ+β)-Toc-1	16	Gm16_1079308- Gm16_3673245	2.5–18.5	2.87	0.23	−10.56	Carbondale, IL
Total- Toopherol	qT-Toc-1	1	Gm01_3504836- Gm01_5566630	0.1–1.7	4.52	0.08	−14.62	Carbondale, IL
	qT-Toc-2	8	Gm08_2622664- Gm08_2852874	12.9–13.3	2.11	0.03	−4.44	Carbondale, IL
	qT-Toc-3	10	Gm10_3935014- Gm10_3890052	79.4–84.4	2	0.03	−4.27	Carbondale, IL
	qT-Toc-4	16	Gm16_1079308- Gm16_3673245	0.5–18.5	3.1	0.18	−13.92	Carbondale, IL
	qT-Toc-5	17	Gm17_3916734- Gm17_3929518	6.2–48.7	3.07	0.18	−11.81	Carbondale, IL
	qT-Toc-6	20	Gm20_3665142- Gm20_1046460	174.9–176.9	2.51	0.03	−10.65	Carbondale, IL

Five QTLs controlling α-tocopherol content in soybean were identified on Chrs. 6, 1 and 2 (Table 3, Figures S1 and S2). The qα-Toc-2-IL-2017 (192.6–197.6 cM) and qα-Toc-3-IL-2020 (195–197 cM) were collocated on Chr. 06. Additionally, eight QTLs underlying δ-tocopherol content were identified on Chrs. 1,2,6,8,16,20. The qδ-Toc-3-IL-2017 located on Chr.6 explains 27.9% of the phenotype (Table 3, Figures S1 and S2). For the γ+β tocopherol content, ten QTLs were identified on Chrs. 2,6,8,12,13, and 16 (Table 3, Figures S1 and S2). Twelve QTLs controlling total tocopherol content were identified and mapped on Chrs. 1, 5, 6, 7, 8, 9, 10, 12, 16, 17, 20 (Table 3, Figures S1 and S2).

2.4. In Silico Reconstruction of the Tocopherol Biosynthetic Pathway in Soybean

The tocopherol biosynthetic pathway has been investigated in the model plant *Arabidopsis thaliana*. The genes and compounds involved in that pathway were previously reported [33]. To reconstruct the tocopherol biosynthesis pathway in soybean, the reverse BLAST of these genes was conducted using SoyBase.

Thirty-seven candidate genes underlying the soybean's tocopherol pathway were identified (Figure 2). In the Shikimat pathway five candidate genes were identified for Tyrosine Aminotransferase (TAT) including *Glyma.06g235500*, *Glyma.06g235900*, *Glyma.12g161500*, *Glyma.12g205900*, and *Glyma.13g295000*. Two candidate genes were identified for hydroxyphenylpyruvate dioxygenase (HPPD) (PDS1), *Glyma.14G030400*, and *Glyma.02G284600*. In the last step of the MEP pathway three candidate genes underlying the geranylgeranyl diphosphate reductase (GGDR) that catalyzes the production of phytyl diphosphate (PDP) from geranylgeranyl diphosphate (GGDP), were identified, *Glyma.02G273800*, *Glyma.05G026200*, and *Glyma.17G100700* (Figure 2).

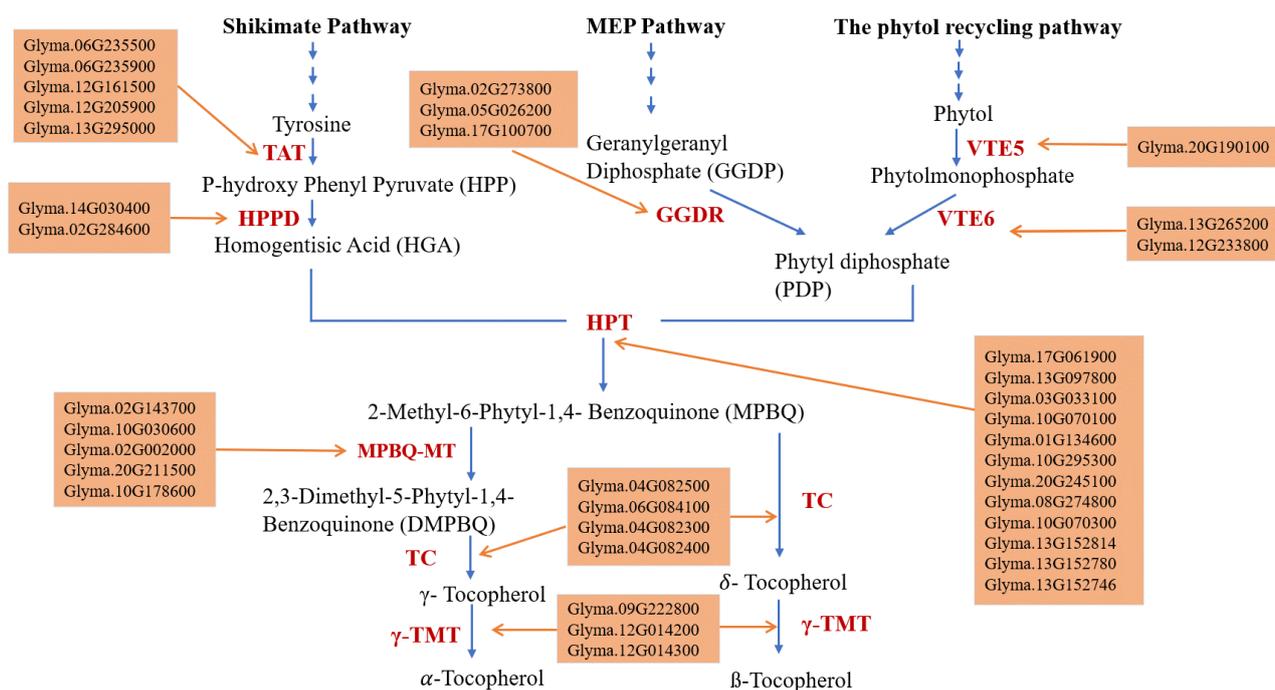


Figure 2. Tocopherol metabolic pathway [15] with identified candidate genes in soybean. TAT: Tyrosine Aminotransferase; HPPD: Hydroxyphenylpyruvate Dioxygenase; GGDR: Geranylgeranyl Diphosphate Reductase; HPT: Homogentisate Phytyl Transferase; MPBQ-MT: methylphytylbenzoquinol methyltransferase; TC: tocopherol cyclase; γ -TMT: gamma tocopherol methyltransferase. MEP Pathway: methylerythritol 4-phosphate pathway.

For the phytol recycling pathway, one candidate gene was identified to be the phytol kinase (VTE5) *Glyma.20G190100*, and two candidate genes were identified to be the phytol-phosphate kinase (VTE6) *Glyma.13G265200*, and *Glyma.12G233800* (Figure 2).

In the core tocopherol pathway, twelve candidate genes were identified for the HPT (VTE2), *Glyma.17G061900*, *Glyma.13G097800*, *Glyma.03G033100*, *Glyma.10G070100*, *Glyma.01G134600*, *Glyma.10G295300*, *Glyma.20G245100*, *Glyma.08G274800*, *Glyma.10G070300*, *Glyma.13G152814*, *Glyma.13G152780*, and *Glyma.13G152746*, four were identified for the TC (VTE1), *Glyma.04G082500*, *Glyma.06G084100*, *Glyma.04G082300*, and *Glyma.04G082400*. Five candidate genes were identified for the MPBQ-MT (VTE3), *Glyma.02G143700*, *Glyma.10G030600*, *Glyma.02G002000*, *Glyma.20G211500*, and *Glyma.10G178600*, in addition to three for the γ -TMT (VTE4), *Glyma.09G222800*, *Glyma.12G014200*, and *Glyma.12G014300* (Figure 2).

2.5. The Association between the Identified Tocopherol Pathway Candidate Genes and the Identified Tocopherol QTL

Among the identified candidate genes, 11 were located close to the identified QTLs on Chrs. 2, 5, 6, 10, 12, and 17 (Table 4, Figure 2). These candidate genes include the tocopherol cyclase candidate (TC) gene *Glyma.06G084100* that is located close to seven seed tocopherol QTLs controlling δ -tocopherol, α -tocopherol, and total tocopherol on Chr. 6 (Table 4).

On Chr. 2, the MPBQ-MT candidate genes *Glyma.02G002000* and *Glyma.02G143700* are located close to $q\delta$ -Toc-2-(2017). The *Glyma.02G002000* candidate gene is also located close to $q\alpha$ -Toc-2-(2020) and $q(\gamma+\beta)$ -Toc-1-(2020) (Table 4, Figure 2). The γ -TMT candidate genes *Glyma.12G014200* and *Glyma.12G014300* are positioned near to QTLs controlling $\gamma+\beta$ tocopherol and total tocopherol (Table 4, Figure 2). The HPT candidate gene *Glyma.17G061900*, and the GGDR candidate gene *Glyma.17G100700* are located close to QTLs controlling total tocopherol on Chr. 17 (Table 4, Figure 2). *Glyma.05G026200* is a GGDR candidate gene that is positioned near to a QTL underlying total tocopherol on Chr. 5 (Table 4, Figure 2).

Table 4. Tocopherol candidate genes located within or close to the tocopherol QTL identified in the FxW82 RIL population grown in Carbondale, IL over two years **A.** 2017 and **B.** 2020.

A. QTLs that Control Seed Tocopherols Contents in Carbondale, IL (2017)								
Trait	QTL	Wm82.a4. v1 Gene Models			Glyma1.0 Gene Models			Dist. (Mbp)
		Gene ID	Start	End	Gene ID	Start	End	
α-Tocopherol	qα-Toc-1	<i>Glyma.06G084100</i>	6,435,516	6,441,328	<i>Glyma06g08850</i>	6,460,802	6,466,636	4.8
	qα-Toc-2	<i>Glyma.06G084100</i>	6,435,516	6,441,328	<i>Glyma06g08850</i>	6,460,802	6,466,636	4.4
	qδ-Toc-1							
δ-Tocopherol	qδ-Toc-2	<i>Glyma.02G002000</i>	237,750	243,006	<i>Glyma02g00440</i>	237,612	245,017	1.2
		<i>Glyma.02G143700</i>	15,253,811	15,256,708	<i>Glyma02g16210</i>	14,623,815	14,626,862	4.6
	qδ-Toc-3	<i>Glyma.06G084100</i>	6,435,516	6,441,328	<i>Glyma06g08850</i>	6,460,802	6,466,636	2.01
	q(γ+β)-Toc-1	<i>Glyma.06G084100</i>	6,435,516	6,441,328	<i>Glyma06g08850</i>	6,460,802	6,466,636	2.09
(γ+β)-Tocopherol	q(γ+β)-Toc-2							
	q(γ+β)-Toc-3	<i>Glyma.12G014200</i>	1,020,484	1,023,995	<i>Glyma12g01680</i>	1,020,554	1,024,132	2.7
		<i>Glyma.12G014300</i>	1,028,051	1,031,954	<i>Glyma12g01690</i>	1,028,132	1,032,092	2.7
	q(γ+β)-Toc-4	<i>Glyma.12G014200</i>	1,020,484	1,023,995	<i>Glyma12g01680</i>	1,020,554	1,024,132	2.6
		<i>Glyma.12G014300</i>	1,028,051	1,031,954	<i>Glyma12g01690</i>	1,028,132	1,032,092	2.6
Total-Tocopherols	q(γ+β)-Toc-5							
	qT-Toc-1	<i>Glyma.05G026200</i>	2,284,067	2,286,242	<i>Glyma05g01000</i>	606,481	608,812	2.6
	qT-Toc-2	<i>Glyma.06G084100</i>	6,435,516	6,441,328	<i>Glyma06g08850</i>	6,460,802	6,466,636	4.3
	qT-Toc-3	<i>Glyma.06G084100</i>	6,435,516	6,441,328	<i>Glyma06g08850</i>	6,460,802	6,466,636	2.3
	qT-Toc-4							
	qT-Toc-5							
	qT-Toc-6	<i>Glyma.12G014200</i>	1,020,484	1,023,995	<i>Glyma12g01680</i>	1,020,554	1,024,132	2.6
		<i>Glyma.12G014300</i>	1,028,051	1,031,954	<i>Glyma12g01690</i>	1,028,132	1,032,092	2.6

B. QTLs that control seed tocopherols contents in Carbondale, IL (2020)								
Trait	QTL	Wm82.a4. v1 Gene Models			Glyma1.0 Gene Models			Dist. (Mbp)
		Gene ID	Start	End	Gene ID	Start	End	
α-Tocopherol	qα-Toc-1							
	qα-Toc-2	<i>Glyma.02G002000</i>	237,689	243,112	<i>Glyma02g00440</i>	237,612	245,017	0.7
	qα-Toc-3	<i>Glyma.06G084100</i>	6,466,090	6,471,839	<i>Glyma06g08850</i>	6,460,802	6,466,636	4.5
	qδ-Toc-1							
δ-Tocopherol	qδ-Toc-2							
		<i>Glyma.10G030600</i>	2,658,064	2,661,302	<i>Glyma10g03590</i>	2,650,012	2,653,309	1.28
	qδ-Toc-3	<i>Glyma.10G070100</i>	6,923,409	6,931,780	<i>Glyma10g08080</i>	6,888,551	6,893,731	2.95
		<i>Glyma.10G070300</i>	7,023,173	7,029,710	<i>Glyma10g08150</i>	6,986,426	6,992,505	3.04
(γ+β)-Tocopherol	qδ-Toc-4							
	qδ-Toc-5							
	q(γ+β)-Toc-2	<i>Glyma.02G002000</i>	237,689	243,112	<i>Glyma02g00440</i>	237,612	245,017	4.06
	q(γ+β)-Toc-1							
Total-Tocopherol	qT-Toc-1							
	qT-Toc-2							
		<i>Glyma.10G030600</i>	2,658,064	2,661,302	<i>Glyma10g03590</i>	2,650,012	2,653,309	1.23
	qT-Toc-3	<i>Glyma.10G070100</i>	6,923,409	6,931,780	<i>Glyma10g08080</i>	6,888,551	6,893,731	2.95
		<i>Glyma.10G070300</i>	7,023,173	7,029,710	<i>Glyma10g08150</i>	6,986,426	6,992,505	3.05
	qT-Toc-4							
Total-Tocopherol	qT-Toc-5	<i>Glyma.17G061900</i>	4,728,685	4,734,790	<i>Glyma17g06940</i>	4,998,801	5,004,742	1.06
		<i>Glyma.17G100700</i>	7,920,291	7,923,450	<i>Glyma17g10890</i>	8,190,830	8,194,219	4.2
	qT-Toc-6							

2.6. Association between the Identified Candidate Genes and the Previously Reported QTL

Mapping the identified genes to the previously reported QTL regions associated with soybean seeds tocopherols was done using data from SoyBase and previous studies describing the QTL underlying tocopherol contents in soybean [7,13,34,35]. Six candidate genes were located within the identified seed tocopherol QTLs and ten were very close to some of these regions (Table 5).

Among these QTLs, qαTC-9 QTL was collocated with the γ-TMT3 (*Glyma.09G222800*) [7]. Also, γ-TMT2 (*Glyma.12G014300*) and γ-TMT1 (*Glyma.12G014200*) are located 703 kb and 711 kb, respectively, apart from qαTC-12 QTL associated with α-tocopherol content [7]. Whereas the fourth tocopherol candidate gene (*Glyma.06G084100*) was located 9.7 Mbp from the QTVEC2_2 QTL underlying total seed tocopherol content identified earlier [34]. The HPT gene (*Glyma.17G061900*) is located 4 Mbp apart from qαγR-17 QTL associated

with seed α -tocopherol content [7]. Moreover, *Glyma.04G082500*, *Glyma.04G082300*, and *Glyma.04G082400*, the tocopherol cyclase candidate genes, were located within q δ TC-4 QTL associated with δ -tocopherol [7]. The MPBQ-MT genes *Glyma.02G002000* and *Glyma.02G143700* were located within two QTL controlling the seed γ -tocopherol content. The first one is the seed tocopherol, γ -1-5 QTL [34] (<https://soybase.org/>; accessed on 3 April 2022) and the second one is the seed tocopherol, γ -2-5 [35] (<https://soybase.org/>; accessed on 3 April 2022). The two HPT candidate genes *Glyma.10G070100* and *Glyma.10G070300* and the MPBQ-MT candidate gene *Glyma.10G030600* are located 3.3, 3.18, and 7.6 Mbp, respectively from the seed total tocopherol, T-Toc 2-2 QTL [35] (<https://soybase.org/>; accessed on 3 April 2022). The MPBQ-MT candidate gene *Glyma.20G211500* and the HPT candidate gene (*Glyma.20G245100*) were located 0.22 and 0.9 Mbp apart from the seed tocopherol, QTVEC2_2 QTL controlling total tocopherol content [35] (Table 5).

Table 5. Tocopherol candidate genes associated to the previously reported QTLs.

Gene ID	Start	End	QTL	QTL Start	QTL End	Parents	Number Loci Tested	Lod Score	Interval Length	Reference
<i>Glyma.09G222800</i>	44,341,974	44,346,311	q α TC-9	43,927,286	44,366,371	TK780 X B04009	ND	13.1	ND	[7]
<i>Glyma.12G014200</i>	1,026,615	1,029,095	q α TC-12	1,507,927	1,790,872	TK780 X B04009	ND	7.8	ND	[7]
<i>Glyma.12G014300</i>	1,033,151	1,037,054	q α TC-12	1,507,927	1,790,872	TK780 X B04009	ND	7.8	ND	[7]
<i>Glyma.04G082500</i>	6,948,445	6,954,177	q δ TC-4	6,780,105	7,188,146	TK780 X B04009	ND	5.5	ND	[7]
<i>Glyma.04G082400</i>	6,946,447	6,947,480	q δ TC-4	6,780,105	7,188,146	TK780 X B04009	ND	5.5	ND	[7]
<i>Glyma.04G082300</i>	6,945,685	6,946,469	q δ TC-4	6,780,105	7,188,146	TK780 X B04009	ND	5.5	ND	[7]
<i>Glyma.06G084100</i>	6,466,090	6,471,839	Seed tocopherol, alpha 1-2	16,106,296	16,256,544	Bayfield X Hefeng 25	107	ND	ND	[35]
<i>Glyma.14G030400</i>	2,204,142	2,206,424	Seed tocopherol, alpha 2-1	675,214	2204,996	X OAC Bayfield OAC	606	ND	ND	[34]
<i>Glyma.02G143700</i>	14,826,295	14,829,286	Seed tocopherol, gamma 1-5	13,316,369	37,285,448	Bayfield X Hefeng 25	107	ND	ND	[35]
			Seed tocopherol, gamma 2-5	14,288,241	45,267,040	X OAC Bayfield OAC	606	ND	56.73	[34]
<i>Glyma.02G002000</i>	237,689	243,112	Seed tocopherol, gamma 1-5	13,316,369	37,285,448	Bayfield X Hefeng 25	107	ND	ND	[35]
			Seed tocopherol, gamma 2-5	14,288,241	45,267,040	X OAC Bayfield OAC	606	ND	56.73	[34]
<i>Glyma.13G097800</i>	21,299,008	21,305,797	Seed tocopherol, delta 1-3	15,248,933	15,306,234	Bayfield X Hefeng 25	107	ND	ND	[35]
<i>Glyma.17G061900</i>	4,728,685	4,734,790	q α γ R-17	8,786,113	9,025,866	TK780 X B04009	ND	4.1	ND	[7]
<i>Glyma.17G100700</i>	792,0291	7,923,450	Seed tocopherol, gamma 3-6	5,891,979	36,718,722	OAC Bayfield X OAC Shire	550	2.6	67.66	[13]
<i>Glyma.12g161500</i>	30,805,424	30,815,155	Seed tocopherol, total 3-5	24,129,662	37,556,592	OAC Bayfield X OAC Shire	550	3.4	29.62	[13]
<i>Glyma.12g205900</i>	38,082,220	38,086,113	Seed tocopherol, alpha 3-3	24,129,662	37,556,592	OAC Bayfield X OAC Shire	550	3.5	29.62	[13]
			Seed tocopherol, total 3-5	24,129,662	37,556,592	OAC Bayfield X OAC Shire	550	3.4	29.62	[13]
<i>Glyma.13g295000</i>	38,800,738	38,805,839	Seed tocopherol, delta 3-2	37,603,911	40,131,770	OAC Bayfield X OAC Shire	550	2.6	17.11	[13]
			Seed tocopherol, delta 3-3	31,449,060	43,325,731	OAC Bayfield X OAC Shire	550	3.8	52.94	[13]

2.7. Organ-Specific Expression of the Identified Candidate Genes

To investigate the role of the identified 37 candidate genes, the expression analysis of these genes was performed in Williams 82 cv. using the publicly available RNA-seq

database at SoyBase (<https://soybase.org/>; accessed on 3 April 2022). The tissues that were included in this dataset were leaves, nodules, roots, pods, and seeds. Amongst the 37 candidate genes, no RNAseq data was available for the TC candidate gene *Glyma.04G082400*, the HPT candidate gene *Glyma.08G274800*, and the MPBQ-MT candidate gene *Glyma.10G178600*. The rest of the tocopherol candidate genes presented different gene expression patterns. Most of the identified candidate genes were expressed in all the analyzed tissues except for the HTP candidate gene, *Glyma.03G033100*, that was not expressed in any of the tissues. While the two GGDR candidate genes, *Glyma.05G026200* and *Glyma.17G100700*, the HPT candidate gene, *Glyma.13G097800*, and the MPBQ-MT candidate gene, *Glyma.02G143700*, were highly expressed in flowers. The two GGDR candidate genes *Glyma.05G026200* and *Glyma.17G100700*, the MPBQ-MT candidate genes *Glyma.02G143700*, *Glyma.02G002000*, and *Glyma.10G030600*, the TC candidate gene *Glyma.06G084100*, and the γ -TMT candidate gene *Glyma.12G014300* were abundantly expressed in leaves. The GGDR candidate genes *Glyma.05G026200*, *Glyma.02G273800* and *Glyma.17G100700*, and the TAT candidate gene *Glyma.12G161500* were highly expressed in seeds. The TAT candidate genes *Glyma.06G235900* and *Glyma.12G205900*, the two GGDR candidate genes *Glyma.05G026200* and *Glyma.17G100700*, and the MPBQ-MT candidate genes *Glyma.02G143700* and *Glyma.10G030600* were highly expressed in pods. The γ -TMT candidate gene *Glyma.09G222800*, the GGDR candidate genes *Glyma.05G026200* and *Glyma.02G273800* were highly expressed in roots. Whereas the TAT candidate genes *Glyma.12G161500* and *Glyma.13G295000*, the GGDR candidate genes *Glyma.05G026200* and *Glyma.02G273800*, and the MPBQ-MT candidate gene *Glyma.10G030600* were highly expressed in the nodules (Figure 3A, Table S1).

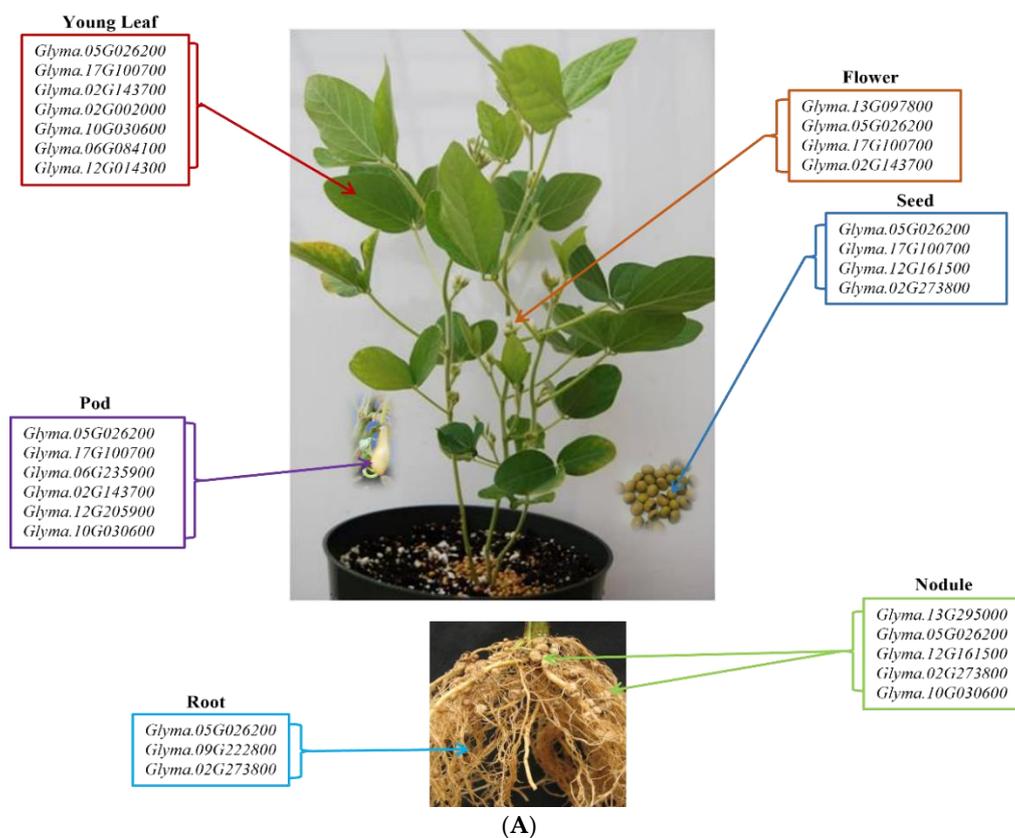


Figure 3. Cont.

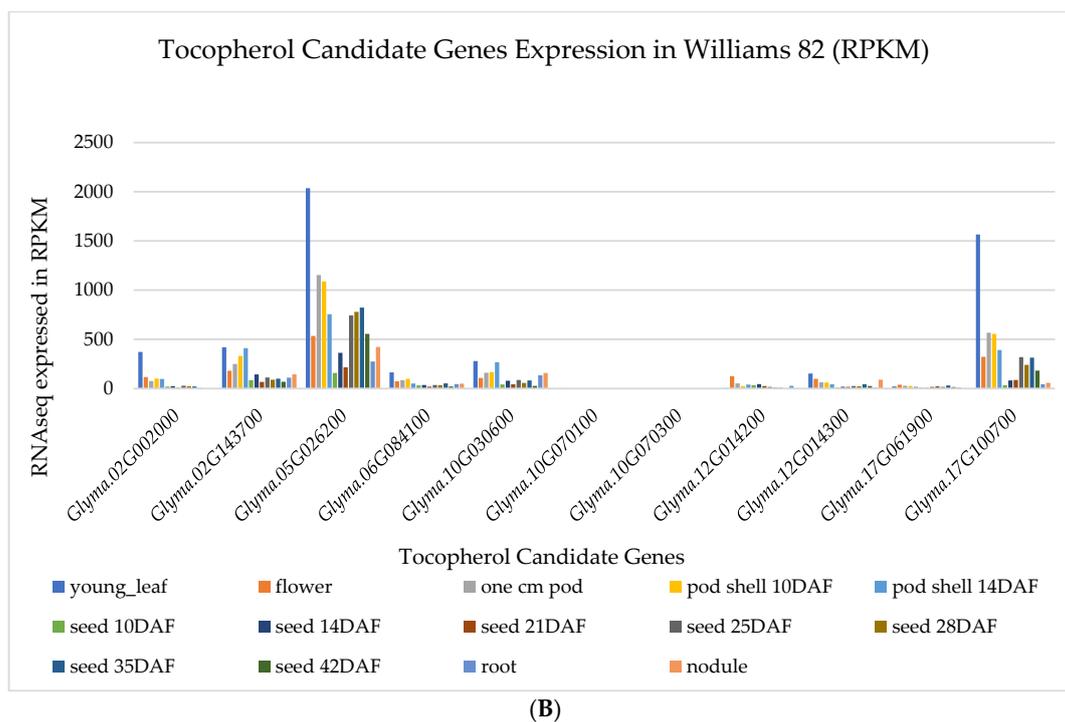


Figure 3. (A). Tissue specific expression of the identified tocopherol candidate genes. (B). Expression pattern of the 11 tocopherol candidate genes located within tocopherol QTL in Williams 82 (RPKM) were retrieved from publicly available RNA-seq data from Soybase database (<http://www.soybase.org/soyseq>; accessed on 3 April 2022).

Amongst the identified candidate genes, eleven were located close to the tocopherol seed QTLs identified in FxW82 RIL population, in tocopherol seed content in soybean. *Glyma.05G026200* and *Glyma.17G100700* are highly expressed in seeds in Williams 82 cv., followed by *Glyma.02G002000*, *Glyma.02G143700*, *Glyma.10G030600*, *Glyma.12G014300*, and *Glyma.06G084100* that have moderate expression profiles in seeds. The rest of the genes have a low expression profile in seeds except for *Glyma.10G070100* and *Glyma.10G070300* that have a limited expression profile, with very low to no expression in seeds (Figure 3B).

3. Discussion

Tocopherols are lipophilic antioxidants that are important for human health due to their ability to prevent the oxidation of unsaturated fatty acids by scavenging the free radicals and prevent cell membrane damage [13]. Soybean seeds contain the highest tocopherol concentrations among all legume species [36]. The dominant tocopherol isoform in soybean seeds is γ -tocopherol with amounts reaching up to 70% of the total tocopherol content, while α -tocopherol isoform has a lower concentration of about 10% of the total tocopherol content. The α -tocopherol isoform has the highest vitamin E activity [4] and has the highest affinity with the hepatic tocopherol transfer protein. Therefore, improving soybean seed tocopherol composition and content is crucial.

Several studies have revealed the genetic and molecular bases underlying tocopherol content in soybean [7,13,14,29,30,34,35] as summarized recently in [37].

Among the QTLs identified in these studies, $q\alpha$ TC-9 QTL was collocated with the γ -TMT3 *Glyma.09G222800*; [7], however, the QTL identified in this study on Chr. 9 was more than 40 Mbp apart from this gene. Also, γ -TMT2 (*Glyma.12G014300*) and γ -TMT1 (*Glyma.12G014200*) are located 703 kb and 711 kb, respectively, apart from the $q\alpha$ TC-12 QTL associated with α -tocopherol content [7]. Similarly, in this study, these candidate genes are located 2.7, 2.6 and 2.6 Mbp apart from $q(\gamma+\beta)$ -Toc-3-(2017), $q(\gamma+\beta)$ -Toc-4-(2017), and qT -Toc-6-(2017), respectively, on Chr. 12 (Table 4, Figures S1 and S2). Also, the TC candidate

gene (*Glyma.06G084100*) was located 9.7 Mbp from the QTVEC2_2 QTL underlying the total tocopherol identified earlier [34]. Likewise, this gene is located close to seven seed tocopherol QTL, 2 Mbp from the *q δ -Toc-3-(2017)* ($R^2 = 27.8$), 4.8 and 4.4 Mbp from *q α -Toc-1-(2017)* and *q α -Toc-2-(2017)*, respectively, 4.3 and 2.3 Mbp from *qT-Toc-2-(2017)* and *qT-Toc-3-(2017)*, respectively, 2.09 Mbp from *q(γ + β)-Toc-1-(2017)*, and 4.5 Mbp from *q α -Toc-2-(2020)* on Chr. 6 (Table 4, Figures S1 and S2). On Chr. 2, the MPBQ-MT candidate genes *Glyma.02G002000* and *Glyma.02G143700* are located 1.2 and 4.6 Mbp, respectively from *q δ -Toc-2-(2017)*. Also, the *Glyma.02G002000* candidate gene is located 0.7 and 4.06 Mbp apart from *q α -Toc-2-(2020)* and *q(γ + β)-Toc-1-(2020)*, respectively (Table 4, Figures S1 and S2). This is coherent with previous studies [34,35], where these two candidate genes were located within seed tocopherol, gamma 1-5 [35] (<https://soybase.org/>; accessed on 3 April 2022) and seed tocopherol, gamma 2-5 [34] (<https://soybase.org/>; accessed on 3 April 2022).

The QTL associate with δ -tocopherol explains 27.87% of the phenotype, and the one associated with α -tocopherol explains only 9.96% of the phenotype. A TC (*Glyma.06G084100*) gene was identified close to these QTLs, the TC enzyme is involved directly in the conversion of MPBQ to δ -tocopherol, and indirectly in the conversion to α -tocopherol (Table 4, Figure 2). α -tocopherol is the most known potent fat-soluble antioxidant, it is preferentially absorbed and accumulated in humans [38], its activity has been demonstrated in the prevention and treatment of heart disease, cancer and Alzheimer's disease [39]. Alpha-tocopherol has been designated as the most beneficial tocopherol compound among health professionals. Unfortunately, this compound is present in small amount in soybean oil when compared to sunflower, canola or corn oil [40]. Therefore, improving α -tocopherol content in soybean is a priority for the soy-industry, the identification in the two years data of the *q α -Toc-2-IL-2017* (192.6–197.6 cM) and *q α -Toc-3-IL-2020* (195–197 cM) that were collocated on Chr. 06 will provide an opportunity for breeding lines with high α -tocopherol.

Interestingly, the HPT candidate gene (*Glyma.17G061900*) is located 4 Mbp apart from *q α γ R-17* QTL associated with α -tocopherol [7]. Similarly, this candidate gene is located 1.06 Mbp apart from *qT-Toc-5-(2020)* identified here on Chr. 17 (Table 4, Figures S1 and S2). The two HPT candidate genes *Glyma.10G070100* and *Glyma.10G070300* and the MPBQ-MT candidate gene *Glyma.10G030600* are located 3.3, 3.18, and 7.6 Mbp, respectively from the seed tocopherol, total 2-2 QTL [34,35] (<https://soybase.org/>; accessed on 3 April 2022). Likewise, in this study these genes are located 2.95, 3.04 and 1.28, respectively from *q δ -Toc-2-(2020)*, and *qT-Toc-3-(2020)* on Chr. 10 (Table 4, Figures S1 and S2). The MPBQ-MT candidate gene *Glyma.20G211500* and the HPT candidate gene (*Glyma.20G245100*) were located 0.22 and 0.9 Mbp apart from seed tocopherol, QTVEC2_2 QTL controlling total tocopherol content identified earlier [34], however, the QTL identified in this study on Chr. 20 was located more than 40 Mbp apart from these genes. Moreover, *Glyma.04G082500*, *Glyma.04G082300*, and *Glyma.04G082400* tocopherol cyclase candidate genes were located within *q δ TC-4* QTL associated with δ -tocopherol and identified earlier [7].

Although previous studies have reported some soybean genes as candidates for the tocopherol biosynthesis pathway [41], the present study shows the most comprehensive analysis of the whole soybean genome, showing the potential candidate genes for the tocopherol biosynthetic pathway in soybean.

Most QTL regions that were identified in 2017 were not found in 2020 except the *q α -Toc-2-IL-2017* (192.6–197.6 cM) and *q α -Toc-3-IL-2020* (195–197 cM) that were collocated on Chr. 06. This could be explained by the difference in weather conditions between 2017 and 2020. In August 2017 the temperature ranged between 8 and 33.3 °C, while in August 2020 the temperature ranged between 13.3 and 32.8 °C (<https://www.extremeweatherwatch.com/>; accessed on 3 April 2022). It has been proven that temperature stress during all stages of development affect soybean seed tocopherol concentrations [42].

The QTL region identified on Chr.7, *qTotal-Toc-4-IL-2017* (81.4–88.9 cM), is 7.6 cM from a QTL region previously identified [7]. Likewise, the QTL region identified on Chr. 20 is 61.21 cM away from a QTL region identified in previous studies [34,35]. Interestingly, the QTL region identified on Chr.12 *q γ + β -Toc-3-IL-2017*, *q γ + β -Toc-4-IL-2017*, *qTotal-Toc-6-IL-*

2017 (0.5–18.5 cM) and Chr.8 q γ + β -Toc-2-IL-2017 and qT-Toc-2-IL-2020 (12.9–31.2 cM) were reported in previous studies [7,34]. Which make them important regions to investigate further for candidate genes. The rest of the QTLs are novel (Table 4, Figures S1 and S2).

Although previous studies identified QTL regions for soybean seed tocopherol content on Chr.6, all the identified QTLs map to the region between 74.5 and 118.5 cM (Table S2) [7,34,35]. The QTL regions identified in this study on Chr.6 clusters between 173.7 and 207 cM, which is the region that encompass an important gene in the tocopherol biosynthesis pathway, namely the tocopherol cyclase candidate gene, *Glyma.06G084100*. This QTL on Chr.6 is responsible for 27.8% of δ -tocopherol, 9.96% of α -tocopherol, 6.16% of γ + β -tocopherol, and 6.95% of total tocopherol content

4. Materials and Methods

4.1. Plant Materials

The F_{6:13} ‘Forrest’ \times ‘Williams 82’ RIL population ($n = 306$) described previously was used in this study [30,43]. The parents and RILs were grown in Carbondale, southern Illinois in 2017 and 2020, and seeds were harvested at maturity of all RILs and parents.

4.2. Tocopherols Quantification

At maturity, seeds of the parents and RILs were harvested and analyzed for α -Tocopherol (α -Toc), δ -Tocopherol (δ -Toc), α + β -Tocopherol ($(\gamma$ + β)-Toc), and total-Tocopherols (T-Toc) using a protocol developed and validated in the Nguyen Lab, the University of Missouri. Briefly, approx. 1gr. of soybean seeds were ground to fine powders with a Thomas Wiley Mini-Mill, followed by lyophilizing for 48 hrs. Approx. 200mg of powder were mixed with 2mL 200-proof ethanol and vortexed, followed by an incubation with agitation at 75 °C for 2 hrs. The products were then filtered into HPLC vials for analysis along with standard solutions of tocopherols. Quantification of tocopherols was performed by employing an external calibration curve method, in which each curve was created with the six standard solutions of 0.62, 1.25, 2.5, 5, 10, and 20 μ g/mL.

4.3. DNA Isolation, SNP Genotyping, and Genetic Map Construction

DNA Isolation, SNP Genotyping, and the construction of the F \times W82 genetic linkage map have been described earlier [30]. Briefly, SNP genotyping was performed with Illumina Infinium SoySNP6K BeadChips (Illumina, Inc. San Diego, CA, USA) and the genetic map was constructed with JoinMap 4.0 software with a LOD score of 3.0 and maximum distance of 50 cM as described earlier [30].

4.4. Seed Tocopherols QTL Detection

We used WinQTL Cartographer 2.5 [31] and both interval mapping (IM) and composite interval mapping (CIM) methods to identify QTL that control seed α -Toc, δ -Toc, γ + β -Toc, and T-Toc in this RIL population; however, only QTL detected with CIM are reported here. QTL identified via IM are reported in the supplementary data section (Table S4A,B). MapChart 2.2 [32] was used to draw chromosomes with CIM tocopherols QTL locations.

4.5. Tocopherols Candidate Genes Identification

The reverse blast of the genes underlying the tocopherol pathway in *Arabidopsis* was conducted using the available data at SoyBase (<https://soybase.org/>; accessed on 3 April 2022). The sequences of the *Arabidopsis* genes were obtained from the Phytozome database (<https://phytozome-next.jgi.doe.gov>; accessed on 3 April 2022), these sequences were used for Blast in SoyBase. The obtained genes that control the tocopherol biosynthetic pathway were mapped to the identified tocopherol QTL.

4.6. Expression Analysis

The expression analysis of the identified tocopherol candidate genes that are located within or close the identified seed tocopherol QTLs was performed using the publicly

available data from SoyBase (<https://soybase.org/>; accessed on 3 April 2022) to produce the expression profiles of these candidate genes in the soybean reference genome Williams 82 in Glyma1.0 Gene Models version.

5. Conclusions

In conclusion, 32 QTL controlling seed tocopherol contents on Chr. 1, 2, 5, 6, 7, 8, 9, 10, 12, 13, 16, 17, and 20 were identified. 37 candidate genes involved in soybean tocopherol biosynthetic pathway have also been identified among which 11 were located close to the QTL regions identified in this study. Two of these candidate genes were highly expressed in seeds *Glyma.05G026200* and *Glyma.17G100700*, followed by *Glyma.06G084100*, *Glyma.02G002000*, *Glyma.02G143700*, *Glyma.10G030600*, and *Glyma.12G014300* with moderate expression profiles in seeds (Figure 3B).

Forrest and Williams 82 sequences of the eleven candidate genes located close to the identified QTLs were compared, and the results have shown that three of them have SNPs between the Forrest and Williams 82 sequences, *Glyma.06G084100*, *Glyma.17G061900* and *Glyma.17G100700* (Figure 4). The TC candidate gene *Glyma.06G084100* has 5 SNPs in the coding sequence, one of them caused a missense mutation (T379A) (Figure 4) in addition to 12 SNPs and 2 InDels in the 5'UTR region (Table S3). The HPT candidate gene, *Glyma.17G061900*, has only one SNP located in the coding sequence that caused a missense mutation (G326A) (Figure 4). For the GGDR candidate gene, *Glyma.17G100700*, there is also only one SNP that caused a silent mutation (Figure 4). These SNPs could play a role in the difference of tocopherol content between Forrest and Williams 82 cultivars. *Glyma.06G084100* is associated with the q δ -Toc-3-(2017) (R² = 27.8), q α -Toc-1-(2017), q α -Toc-2-(2017), qT-Toc-2-(2017), qT-Toc-3-(2017), q(γ + β)-Toc-1-(2017), and q α -Toc-2-(2020) on Chr. 6 (Table 3, Figures S1 and S2). While *Glyma.17G061900* and *Glyma.17G100700* are associated to qT-Toc-5-(2020) on Chr. 17 (Table 2, Figures S1 and S2). These genes could be used in breeding programs or gene editing technology to develop soybean lines and cultivars that produce high amounts of the beneficial tocopherols (vitamin E) for human consumption.

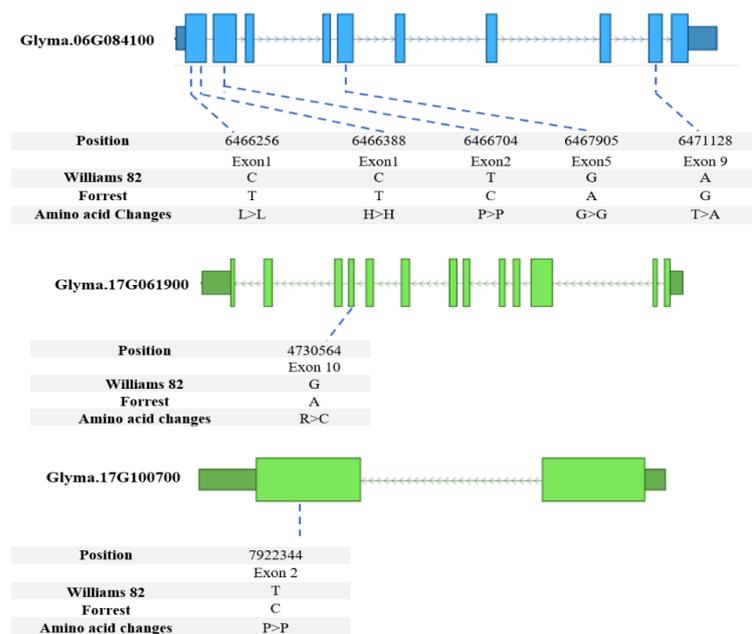


Figure 4. Positions of SNPs between Forrest and Williams 82 cultivars in *Glyma.06G084100*, *Glyma.17G061900* and *Glyma.17G100700* coding sequences. In the gene model diagram, the light blue/light green boxes represent exons, blue/green bars represent introns, dark blue/dark green boxes represent 3'UTR or 5'UTR. SNPs were positioned relative to the genomic position in the genome version W82.a2.

6. Patents

Patent resulting from this work is under submission.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants11091258/s1>, Table S1. Expression profiles of the tocopherol biosynthesis candidate genes in soybean based on RNAseq data available from RNAsequencing data (<http://www.soybase.org/soyseq>; accessed on 3 April 2022). Table S2. QTLs identified in previous studies on Chr. 6. Table S3. Positions of SNPs between Forrest and Williams 82 cultivars in the promoter of the tocopherol cyclase candidate gene (*GmTC06*, *Glyma.06G084100*). Table S4. A. QTLs controlling seed tocopherols contents in Carbondale, IL (2017)—Identified by Interval Mapping Method. B. QTLs controlling seed tocopherols contents in Carbondale, IL (2020)—Identified by Interval Mapping Method. Figure S1. Positions of QTL that control seed α -Tocopherol (α -Toc), δ -Tocopherol (δ -Toc), $(\gamma+\beta)$ -Tocopherol ($(\gamma+\beta)$ -Toc), and Total-Tocopherols (T-Toc) contents on chromosomes 1, 2, 5, 6, 7, 8, 9, 10, 12, 13, 16, 17, and 20. The QTL have been identified in F \times W82 grown in two environments in Carbondale, IL over two years (2017 and 2020). Legend: $(\gamma + \beta) = (G + \beta)$ and (Carb-IL) = Carbondale, IL. Figure S2. QTL that control seed α -Tocopherol (α -Toc), δ -Tocopherol (δ -Toc), $(\gamma+\beta)$ -Tocopherol ($(\gamma+\beta)$ -Toc), and Total-Tocopherols (T-Toc) contents identified by IM and CIM methods in the F \times W82 RIL population grown in two environments in Carbondale, IL over two years (2017 and 2020).

Author Contributions: Conceptualization, K.M., H.T.N. and M.A.K.; D.K. conceived and wrote the manuscript, methodology, D.K., J.Y., T.V. and N.L.; validation, K.M., M.A.K. and H.T.N.; formal analysis, D.K. and J.Y.; investigation, K.M. and D.K.; resources and data curation, K.M., M.A.K. and H.T.N.; review and editing, D.K., J.Y., N.L., M.C., A.E.B., T.V., H.M., M.A.K., K.M. and H.T.N.; visualization, J.Y.; supervision, K.M., M.A.K., J.Y., H.M. and H.T.N.; project administration, K.M., M.A.K., H.T.N. and K.M. supervised and conceived the work, performed data interpretation, designed the experiment, and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the United Soybean Board: Development and Releases of High Tocopherol Soybean Germplasm Project USB # 2220-162-0116.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data supporting reported results are available on request from the corresponding author.

Acknowledgments: We thank all the student workers at Southern Illinois University Carbondale, who assisted in planting the recombinants inbred lines, DNA extraction, and greenhouse and field work. Technical assistance in the HPLC analysis by Haiying Shi at the University of Missouri is greatly appreciated.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rani, A.; Kumar, V.; Verma, S.K.; Shakya, A.K.; Chauhan, G.S. Tocopherol Content and Profile of Soybean: Genotypic Variability and Correlation Studies. *J. Am. Oil Chem. Soc.* **2007**, *84*, 377–383. [[CrossRef](#)]
2. Kamal-Eldin, A.; Appelqvist, L.-Å. The chemistry and antioxidant properties of tocopherols and tocotrienols. *Lipids* **1996**, *31*, 671–701. [[CrossRef](#)] [[PubMed](#)]
3. Grusak, M.A.; DellaPenna, D. Improving the nutrient composition of plants to enhance human nutrition and health. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1999**, *50*, 133–161. [[CrossRef](#)] [[PubMed](#)]
4. Bramley, P.M.; Elmadfa, I.; Kafatos, A.; Kelly, F.J.; Manios, Y.; Roxborough, H.; Schuch, W.; Sheehy, P.J.A.; Wagner, K.H. Vitamin E: A critical review. *J. Sci. Food Agric.* **2000**, *80*, 913–938. [[CrossRef](#)]
5. Kono, N.; Ohto, U.; Hiramatsu, T.; Urabe, M.; Uchida, Y.; Satow, Y.; Arai, H. Impaired α -TTP-PIPs interaction underlies familial vitamin E deficiency. *Science* **2013**, *340*, 1106–1110. [[CrossRef](#)]
6. Pagano, M.C.; Miransari, M. 1—The importance of soybean production worldwide. In *Abiotic and Biotic Stresses in Soybean Production*; Miransari, M., Ed.; Academic Press: San Diego, CA, USA, 2016; pp. 1–26.
7. Park, C.; Dwiyantri, M.S.; Nagano, A.J.; Liu, B.; Yamada, T.; Abe, J. Identification of quantitative trait loci for increased α -tocopherol biosynthesis in wild soybean using a high-density genetic map. *BMC Plant Biol.* **2019**, *19*, 510. [[CrossRef](#)]

8. Seguin, P.; Tremblay, G.; Pageau, D.; Liu, W. Soybean Tocopherol Concentrations Are Affected by Crop Management. *J. Agric. Food Chem.* **2010**, *58*, 5495–5501. [[CrossRef](#)]
9. Marwede, V.; Schierholt, A.; Möllers, C.; Becker, H.C. Genotype × Environment Interactions and Heritability of Tocopherol Contents in Canola. *Crop Sci.* **2004**, *44*, 728–731. [[CrossRef](#)]
10. Ujiie, A.; Yamada, T.; Fujimoto, K.; Endo, Y.; Kitamura, K. Identification of Soybean Varieties with High α -Tocopherol Content. *Breed. Sci.* **2005**, *55*, 123–125. [[CrossRef](#)]
11. Carrera, C.S.; Seguin, P. Factors Affecting Tocopherol Concentrations in Soybean Seeds. *J. Agric. Food Chem.* **2016**, *64*, 9465–9474. [[CrossRef](#)]
12. Dwiyantri, M.S.; Maruyama, S.; Hirono, M.; Sato, M.; Park, E.; Yoon, S.H.; Yamada, T.; Abe, J. Natural diversity of seed α -tocopherol ratio in wild soybean (*Glycine soja*) germplasm collection. *Breed. Sci.* **2016**, *66*, 653–657. [[CrossRef](#)] [[PubMed](#)]
13. Shaw, E.J.; Rajcan, I. Molecular mapping of soybean seed tocopherols in the cross ‘OAC Bayfield’ × ‘OAC Shire’. *Plant Breed.* **2017**, *136*, 83–93. [[CrossRef](#)]
14. Liu, H.; Cao, G.; Wu, D.; Jiang, Z.; Han, Y.; Li, W. Quantitative trait loci underlying soybean seed tocopherol content with main additive, epistatic and QTL × environment effects. *Plant Breed.* **2017**, *136*, 924–938. [[CrossRef](#)]
15. Muñoz, P.; Munné-Bosch, S. Vitamin E in Plants: Biosynthesis, Transport, and Function. *Trends Plant Sci.* **2019**, *24*, 1040–1051. [[CrossRef](#)]
16. Norris, S.R.; Shen, X.; Della Penna, D. Complementation of the Arabidopsis pds1 Mutation with the Gene Encoding p-Hydroxyphenylpyruvate Dioxygenase. *Plant Physiol.* **1998**, *117*, 1317–1323. [[CrossRef](#)]
17. Falk, J.; Andersen, G.; Kernebeck, B.; Krupinska, K. Constitutive overexpression of barley 4-hydroxyphenylpyruvate dioxygenase in tobacco results in elevation of the vitamin E content in seeds but not in leaves1. *FEBS Lett.* **2003**, *540*, 35–40. [[CrossRef](#)]
18. Li, Y.; Wang, Z.; Sun, X.; Tang, K. Current Opinions on the Functions of Tocopherol Based on the Genetic Manipulation of Tocopherol Biosynthesis in Plants. *J. Integrat. Plant Biol.* **2008**, *50*, 1057–1069. [[CrossRef](#)]
19. Valentin, H.E.; Lincoln, K.; Moshiri, F.; Jensen, P.K.; Qi, Q.; Venkatesh, T.V.; Karunanandaa, B.; Baszis, S.R.; Norris, S.R.; Savidge, B.; et al. The Arabidopsis vitamin E pathway gene5-1 Mutant Reveals a Critical Role for Phytol Kinase in Seed Tocopherol Biosynthesis. *Plant Cell* **2005**, *18*, 212–224. [[CrossRef](#)]
20. Vom Dorp, K.; Hölzl, G.; Plohm, C.; Eisenhut, M.; Abraham, M.; Weber, A.P.M.; Hanson, A.D.; Dörmann, P. Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis. *Plant Cell* **2015**, *27*, 2846–2859. [[CrossRef](#)]
21. Collakova, E.; DellaPenna, D. Isolation and functional analysis of homogenisate phytyltransferase from Synechocystis sp. PCC 6803 and Arabidopsis. *Plant Physiol.* **2001**, *127*, 1113–1124. [[CrossRef](#)]
22. Schledz, M.; Seidler, A.; Beyer, P.; Neuhaus, G. A novel phytyltransferase from Synechocystis sp. PCC 6803 involved in tocopherol biosynthesis. *FEBS Lett.* **2001**, *499*, 15–20. [[CrossRef](#)]
23. Sattler, S.E.; Cheng, Z.; DellaPenna, D. From Arabidopsis to agriculture: Engineering improved Vitamin E content in soybean. *Trends Plant Sci.* **2004**, *9*, 365–367. [[CrossRef](#)] [[PubMed](#)]
24. Motohashi, R.; Ito, T.; Kobayashi, M.; Taji, T.; Nagata, N.; Asami, T.; Yoshida, S.; Yamaguchi-Shinozaki, K.; Shinozaki, K. Functional analysis of the 37 kDa inner envelope membrane polypeptide in chloroplast biogenesis using a Ds-tagged Arabidopsis pale-green mutant. *Plant J.* **2003**, *34*, 719–731. [[CrossRef](#)]
25. Cheng, Z.; Sattler, S.; Maeda, H.; Sakuragi, Y.; Bryant, D.A.; DellaPenna, D. Highly Divergent Methyltransferases Catalyze a Conserved Reaction in Tocopherol and Plastoquinone Synthesis in Cyanobacteria and Photosynthetic Eukaryotes. *Plant Cell* **2003**, *15*, 2343–2356. [[CrossRef](#)] [[PubMed](#)]
26. Porfirova, S.; Bergmüller, E.; Tropf, S.; Lemke, R.; Dörmann, P. Isolation of an Arabidopsis mutant lacking vitamin E and identification of a cyclase essential for all tocopherol biosynthesis. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12495–12500. [[CrossRef](#)]
27. Kanwischer, M.; Porfirova, S.; Bergmüller, E.; Dörmann, P. Alterations in Tocopherol Cyclase Activity in Transgenic and Mutant Plants of Arabidopsis Affect Tocopherol Content, Tocopherol Composition, and Oxidative Stress. *Plant Physiol.* **2005**, *137*, 713–723. [[CrossRef](#)] [[PubMed](#)]
28. Van Eenennaam, A.L.; Lincoln, K.; Durrett, T.P.; Valentin, H.E.; Shewmaker, C.K.; Thorne, G.M.; Jiang, J.; Baszis, S.R.; Levering, C.K.; Aasen, E.D.; et al. Engineering vitamin E content: From Arabidopsis mutant to soy oil. *Plant Cell* **2003**, *15*, 3007–3019. [[CrossRef](#)] [[PubMed](#)]
29. Dwiyantri, M.S.; Yamada, T.; Sato, M.; Abe, J.; Kitamura, K. Genetic variation of γ -tocopherol methyltransferase gene contributes to elevated α -tocopherol content in soybean seeds. *BMC Plant Biol.* **2011**, *11*, 152. [[CrossRef](#)]
30. Sui, M.; Jing, Y.; Li, H.; Zhan, Y.; Luo, J.; Teng, W.; Qiu, L.; Zheng, H.; Li, W.; Zhao, X.; et al. Identification of Loci and Candidate Genes Analyses for Tocopherol Concentration of Soybean Seed. *Front. Plant Sci.* **2020**, *11*, 539460. [[CrossRef](#)]
31. Knizia, D.; Yuan, J.; Bellaloui, N.; Vuong, T.; Usovsky, M.; Song, Q.; Betts, F.; Register, T.; Williams, E.; Lakhssassi, N.; et al. The Soybean High Density ‘Forrest’ by ‘Williams 82’ SNP-Based Genetic Linkage Map Identifies QTL and Candidate Genes for Seed Isoflavone Content. *Plants* **2021**, *10*, 2029. [[CrossRef](#)]
32. Wang, Z.; Chen, Z.; Cheng, J.; Lai, Y.; Wang, J.; Bao, Y.; Huang, J.; Zhang, H. QTL Analysis of Na⁺ and K⁺ Concentrations in Roots and Shoots under Different Levels of NaCl Stress in Rice (*Oryza sativa* L.). *PLoS ONE* **2012**, *7*, e51202. [[CrossRef](#)] [[PubMed](#)]
33. Fritsche, S.; Wang, X.; Jung, C. Recent Advances in our Understanding of Tocopherol Biosynthesis in Plants: An Overview of Key Genes, Functions, and Breeding of Vitamin E Improved Crops. *Antioxidants* **2017**, *6*, 99. [[CrossRef](#)] [[PubMed](#)]

34. Li, H.; Wang, Y.; Han, Y.; Teng, W.; Zhao, X.; Li, Y.; Li, W. Mapping quantitative trait loci (QTLs) underlying seed vitamin E content in soybean with main, epistatic and QTL × environment effects. *Plant Breed.* **2016**, *135*, 208–214. [[CrossRef](#)]
35. Li, H.; Liu, H.; Han, Y.; Wu, X.; Teng, W.; Liu, G.; Li, W. Identification of QTL underlying vitamin E contents in soybean seed among multiple environments. *Theor. Appl. Genet.* **2010**, *120*, 1405–1413. [[CrossRef](#)] [[PubMed](#)]
36. Grela, E.R.; Günter, K.D. Fatty acid composition and tocopherol content of some legume seeds. *Anim. Feed Sci. Technol.* **1995**, *52*, 325–331. [[CrossRef](#)]
37. Kassem, M.A. *Soybean Seed Composition: Protein, Oil, Fatty Acids, Amino Acids, Sugars, Mineral Nutrients, Tocopherols, and Isoflavones*; Springer International Publishing AG: Cham, Switzerland, 2021.
38. Rigotti, A. Absorption, transport, and tissue delivery of vitamin E. *Mol. Aspects Med.* **2007**, *28*, 423–436. [[CrossRef](#)]
39. Talegawkar, S.A.; Johnson, E.J.; Carithers, T.; Taylor, H.A.; Bogle, M.L.; Tucker, K.L. Total α-Tocopherol Intakes Are Associated with Serum α-Tocopherol Concentrations in African American Adults. *J. Nutr.* **2007**, *137*, 2297–2303. [[CrossRef](#)]
40. Grilo, E.C.; Costa, P.N.; Gurgel, C.S.S.; Beserra, A.F.D.L.; Almeida, F.N.D.S.; Dimenstein, R. Alpha-tocopherol and gamma-tocopherol concentration in vegetable oils. *Food Sci. Technol.* **2014**, *34*, 379–385. [[CrossRef](#)]
41. Vinutha, T.V.; Bansal, N.; Kumari, K.; Prashat, G.R.; Sreevathsa, R.; Krishnan, V.; Kumari, S.; Dahuja, A.; Lal, S.K.; Sachdev, A.; et al. Comparative Analysis of Tocopherol Biosynthesis Genes and Its Transcriptional Regulation in Soybean Seeds. *J. Agric. Food Chem.* **2017**, *65*, 11054–11064. [[CrossRef](#)]
42. Chennupati, P.; Seguin, P.; Liu, W. Effects of High Temperature Stress at Different Development Stages on Soybean Isoflavone and Tocopherol Concentrations. *J. Agric. Food Chem.* **2011**, *59*, 13081–13088. [[CrossRef](#)]
43. Wu, X.; Vuong, T.D.; Leroy, J.A.; Grover Shannon, J.; Sleper, D.A.; Nguyen, H.T. Selection of a core set of RILs from Forrest × Williams 82 to develop a framework map in soybean. *Theor. Appl. Genet.* **2011**, *122*, 1179–1187. [[CrossRef](#)] [[PubMed](#)]