BMC Plant Biology

RESEARCH

Open Access

Total tocopherol levels in maize grain depend on chlorophyll biosynthesis within the embryo

Sam Herr[1†], Xiaowei Li[1†], Di Wu[1†], Charles T. Hunter[2], Maria Magallanes-Lundback[3], Joshua C. Wood[4], Nicholas Kaczmar[1], C. Robin Buell[4,5], Dean DellaPenna[3] and Michael A. Gore[1*]

Abstract

Background Tocopherols are a class of lipid-soluble compounds that have multiple functional roles in plants and exhibit vitamin E activity, an essential nutrient for human and animal health. The tocopherol biosynthetic pathway is conserved across the plant kingdom, but source of the key tocopherol pathway precursor, phytol, is unclear. Two protochlorophyllide reductases (*POR1* and *POR2*) were previously identified as loci controlling the natural variation of total tocopherols in maize grain, a non-photosynthetic tissue. *POR1* and *POR2* are key genes in chlorophyll biosynthesis yet the contribution of the chlorophyll biosynthetic pathway to tocopherol biosynthesis is still not understood.

Results We took two approaches to alter the activity of these two *POR* genes within kernel tissue, physiological treatments and CRISPR/Cas9-mediated knockouts, to determine the role of chlorophyll biosynthesis for tocopherol content. Since light is required for POR enzymatic activity, we imposed a dark treatment on developing kernels, which reduced chlorophyll *a* and tocopherols levels in embryo tissue by 92–99% and 87–90%, respectively, compared to the light treatment. In CRISPR/Cas9-mediated knockouts, the levels of chlorophyll *a* and tocopherols in embryos of the *por1 por2* double homozygous mutant were reduced by 98–100% and 76–83%, respectively, compared to WT.

Conclusion These findings demonstrate that tocopherol synthesis in maize grain depends almost entirely on phytol derived from chlorophyll biosynthesis within the embryo. *POR1* and *POR2* activity play crucial roles in chlorophyll biosynthesis, underscoring the importance of *POR* alleles and their activity in the biofortification of vitamin E levels in non-photosynthetic grain of maize.

Keywords Biofortification, Chlorophyll, Maize, CRISPR/Cas9, Tocopherol, Vitamin E

†Sam Herr, Xiaowei Li and Di Wu considered joint first authors.

*Correspondence:
Michael A. Gore
mag87@cornell.edu
[1] Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA
[2] USDA-ARS Chemistry Research Unit, Gainesville, FL 32608, USA
[3] Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA
[4] Center for Applied Genetic Technologies, University of Georgia, Athens, GA 30602, USA
[5] Department of Crop & Soil Sciences and Institute of Plant Breeding, Genetics, & Genomics, University of Georgia, Athens, GA 30602, USA

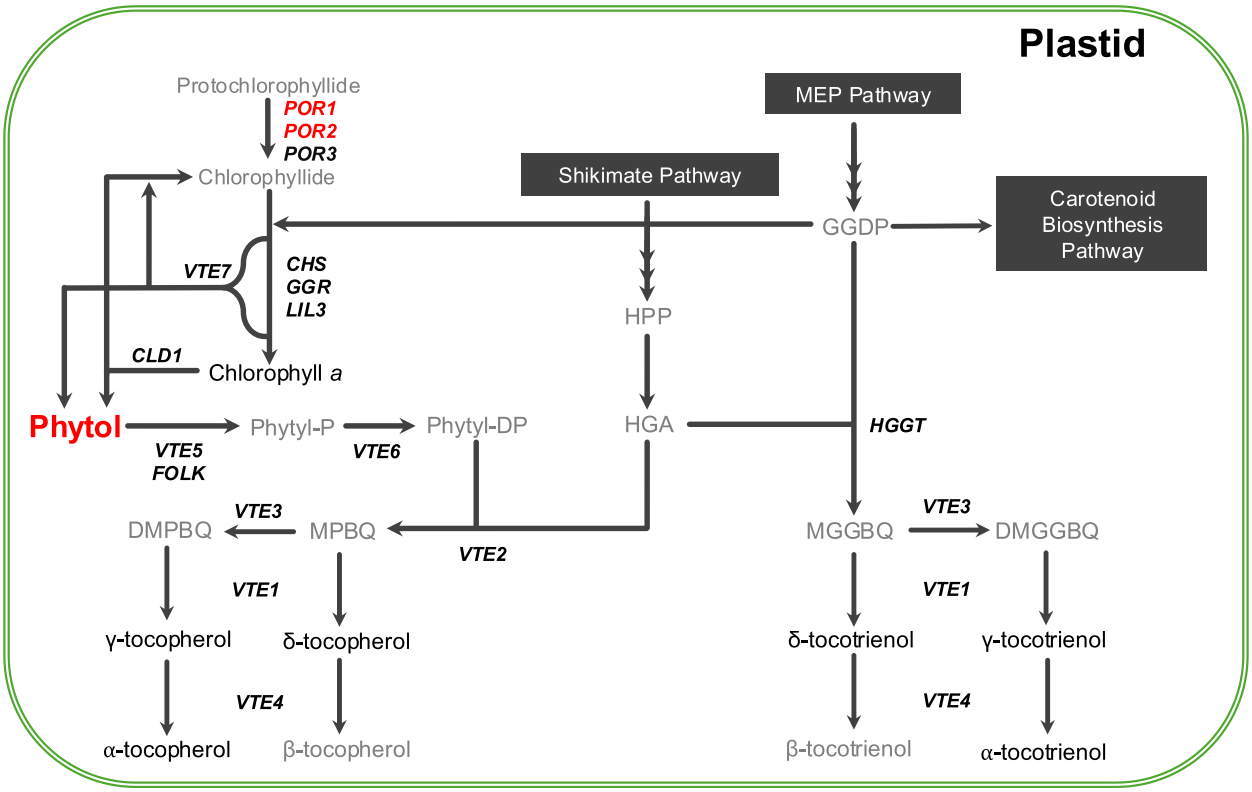

Fig. 1 Tocochromanol biosynthesis pathway. The two precursor pathways and the carotenoid pathway are represented as black boxes. The six quantified tocochromanols and chlorophyll a are indicated in black nonitalicized text. The names of key a priori genes are bolded and italicized at the pathway step(s) catalyzed by their encoded enzymes, with the *POR* genes in red text. The key tocopherol precursor phytol is bolded in red. Compound abbreviations: 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, p-hydroxyphenylpyruvate; MGGBQ, 2-met hyl-6-geranylgeranyl-1,4-benzoquinol; MPBQ, 2-methyl-6-phytyl-1,4-benzoquinol; Phytyl-DP, phytyl diphosphate; Phytyl-P, phytyl monophosphate. Gene abbreviations: chlorophyll synthase (*CHS*); chlorophyll dephytylase1 (*CLD1*); farnesol kinase (*FOLK*); geranylgeranyl reductase (*GGR*); homogentisate geranylgeranyl transferase (*HGGT*); light harvesting protein-like 3 (*LIL3*); protochlorophyllide reductase (*POR1*, *POR2*, and *POR3*); tocopherol cyclase (*VTE1*); homogentisate phytyltransferase (*VTE2*); MPBQ/MGGBQ methyltransferase (*VTE3*); γ-tocopherol methyltransferase (*VTE4*); phytol kinase (*VTE5*); phytol phosphate kinase (*VTE6*); α-/β-hydrolase (*VTE7*). Figure updated from [55]

Background

Suboptimal dietary intake of vitamin E, an essential micronutrient, can be prevalent in countries where maize grain significantly contributes to the caloric requirements of human populations [8, 15]. Increased dietary vitamin E intake has been associated with reduced mortality from cardiovascular disease and cancer [22, 44]. Improved crop nutritional quality through agronomic, conventional breeding, or bioengineering approaches, collectively known as biofortification, represents a potential avenue to enhance dietary vitamin E by increasing tocochromanols, a group of biosynthetically related compounds with varying levels of vitamin E activity [37]. Biofortification efforts can be maximized through a comprehensive understanding of the genetic and molecular basis underlying tocochromanol synthesis and accumulation in maize grain.

Tocochromanols, lipid-soluble compounds synthesized in plants, algae, and some cyanobacteria, protect lipids by quenching both reactive oxygen species and lipid peroxyl radicals [4, 30, 42]. They are structurally separated into two classes: tocotrienols and tocopherols. Each class has four types: α, β, δ, and γ (Fig. 1). Tocotrienols, predominantly produced in the endosperm in monocots and a few dicots [16, 21, 57], are associated with numerous health benefits due to their potent antioxidant properties [45]. Tocopherols, primarily produced in the embryo [16], generally have higher vitamin E activity, with α-tocopherol providing the highest vitamin E activity on a molar basis [26].

The core tocopherol biosynthesis pathway has been well-characterized in *Arabidopsis thaliana* [34] and is conserved across the plant kingdom (Fig. 1). The committed step of tocopherol synthesis is catalyzed by VTE2

(homogentisate phytyltransferase), which condenses homogentisic acid (HGA) and phytyl diphosphate (PDP) to produce 2-methyl-6-phytyl-1,4-benzoquinol (MPBQ), the immediate precursor of the four tocopherol types (Fig. 1). This condensation reaction and the availability of its two substrates, HGA and PDP, primarily determine the level of tocopherol produced within a tissue [4, 37]. HGA is derived from the aromatic amino acid pathway, which is tightly controlled by feedback inhibition [50]. In Arabidopsis, the PDP required for tocopherol synthesis is derived from a two-step phosphorylation of free phytol. Phytyl-P synthesis depends on two kinases, VTE5 and FOLK, which when knocked out results in tocopherol deficiency in 4-week-old Arabidopsis seed and leaf tissues [40]. Phytyl-P is subsequently phosphorylated to PDP by VTE6. *vte6* null alleles result in non-detectable levels of tocopherol [51]. While the synthesis of PDP from phytol is well understood, the mechanism of phytol production for tocopherol synthesis remains unclear.

In photosynthetic tissues such as Arabidopsis leaf and seed, the phytol supply for tocopherol synthesis was thought to depend on chlorophyll degradation, during which chlorophyll *a* is converted to pheophytin *a* and then hydrolyzed by pheophytin pheophorbide hydrolase (PPH) to produce pheophorbide and phytol. However, in an Arabidopsis *pph* null mutant, tocopherol levels did not differ from the wild type (WT) in leaf and seed tissue [60]. An alternative pathway for phytol could be the chlorophyll-salvage cycle, which releases both chlorophyllide *a* and phytol through the dephytylation of chlorophyll *a*. Yet when a key gene in the chlorophyll-salvage cycle, *CHLOROPHYLL DEPHYTYLASE1* (*CLD1*), was silenced there was no significant difference in tocopherol levels compared to the WT [27].

A recent study of natural variation in total seed tocopherols in Arabidopsis identified a strong dependence of tocopherol content on chlorophyll biosynthesis, with phytol being hydrolyzed from chlorophyll biosynthetic intermediates rather than from the degradation of the bulk chlorophyll pool [1]. Specifically, Albert et al. [1] found that null alleles of a seed-specific, plastid-localized alpha/beta hydrolase (*VTE7*) decreased seed tocopherol content by 55% with concomitant increases in specific chlorophyll biosynthetic intermediates with partially reduced tails. There was no impact on tocopherol content of Arabidopsis leaves, but in maize a leaky *vte7* Mu-allele decreased kernel and leaf tocopherol content by 38% and 50%, respectively, indicating differences in phytol provision for tocopherol synthesis between monocot and dicot leaves [1].

Despite maize grain being a non-green, non-photosynthetic tissue, tocopherol synthesis in grain is also strongly dependent on chlorophyll synthesis. Two protochlorophyllide reductases (*POR1* and *POR2*), key chlorophyll biosynthetic genes, underlie the largest effect QTL for total tocopherols in maize grain [5]. Their involvement is supported by near-isogenic and single knockout lines in maize, further highlighting the role of chlorophyll in tocopherol synthesis in non-photosynthetic tissue [29, 56]. In addition to its expression being regulated by light, POR also requires light as a co-substrate to catalyze the reduction of protochlorophyllide to chlorophyllide [14, 47]. POR is part of a complex that includes geranylgeranyl reductase (GGR), chlorophyll synthase (CHS), and a non-enzymatic protein, light harvesting protein-like 3 (LIL3), which catalyzes the final steps of chlorophyll *a* synthesis [19, 49]. The dependence of tocopherol synthesis on chlorophyll synthesis is further supported by knockouts of LIL3, GGR, and CHS, which result in severely reduced levels of chlorophylls and tocopherols in leaf tissue [18, 19, 48, 49, 58]. Given the large molar ratio of tocopherol to chlorophyll in developing maize embryos, Diepenbrock et al. [5] hypothesized that instead of chlorophyll degradation, a chlorophyll-based biosynthetic cycle provides phytol for tocopherol synthesis through the repeated removal of phytol during chlorophyll *a* synthesis. The identification of *VTE7* and its impact on tocopherol abundance and chlorophyll biosynthetic intermediates in both leaf and grain tissue in maize support this model [1].

In this study, we took two complementary experimental approaches to assess the role(s) of POR1 and POR2 and therefore, the contribution of chlorophyll biosynthesis, in tocopherol biosynthesis within maize kernels. In physiological experiments, we imposed light and dark treatments on developing kernels to alter the in vivo light-dependent activities of all PORs [17] and measured the impacts on tocochromanols, chlorophyll *a,* and the transcriptome within embryo tissue. In CRISPR/Cas9 knockout experiments, we generated null alleles of *POR1* and *POR2* and evaluated the metabolomic consequences in embryos of the single and double mutants. From these experiments, we demonstrated the central role of chlorophyll biosynthesis in tocopherol synthesis within the maize embryo.

Materials and methods

Experimental design for evaluation of physiological treatments

We conducted a set of physiological experiments by exposing developing kernels to light and dark treatments to alter the in vivo activities of *POR1* and *POR2*. The field experiment conducted in 2018 included founder lines of the U.S. maize nested association mapping (NAM) panel with contrasting *POR1* and/or *POR2* allelic effects [5]. Relative to the common parent B73, these included B97

and M37W with large *POR1* effects, NC358 and Ki11 with large *POR2* effects, and MS71 and OH7B with large effects for both *POR1* and *POR2.* We also included B73 as a reference line, given that it is the common parent for this panel. The seven genotypes were planted in a split-plot design with two replications at Cornell University's Musgrave Research Farm in Aurora, NY. Each main plot of a single genotype comprised three one-row subplots, each with a length of 5.33 m. The light treatment, dark treatment, and control condition were randomly assigned to each subplot. Each subplot had an average of 20 plants, with multiple plants in each subplot having their primary ears self-pollinated and covered with medium-sized pollination bags (MIDCO Global, St. Louis, MO). The light and dark treatments were applied to self-pollinated ears at 12 DAP. For each light-treated ear, the pollination bag was removed, and the outer husks were peeled, leaving only three to four layers, allowing overexposure to light while still retaining moisture for kernel development. A foil bag was used to completely cover the pollination bag of each dark-treated ear, preventing the penetration of light to the developing kernels. The control ears remained in their pollination bags to represent typical self-pollinated ear conditions (Supporting Information Fig. S1).

We harvested self-pollinated ears from each subplot at two kernel developmental time points for RNA-seq (at 24 DAP) and metabolite profiling (at 24 DAP and at maturity). At 24 DAP, three ears per subplot were collected to provide developing kernels for dissection. The harvested ears were immediately placed on wet ice in a cooler with a sealed lid for transportation. For each subplot, 20 kernels from the middle section of each of the three dehusked ears were manually dissected with a scalpel to separate embryos from other seed tissues. This procedure was conducted in a dark room, with each dissector wearing a green headlamp to minimize the initiation of light-regulated reactions. The embryos from all three ears were bulked to form a representative sample, frozen in liquid nitrogen, and stored at − 80 °C. The remaining selfed ears in each subplot (3–13 ears) were harvested at physiological maturity, dried, individually shelled, and then stored separately as mature kernel samples.

In 2019, to generate an additional year of metabolite data from mature kernels, we repeated the physiological experiment identical to that of 2018. However, this time, we only included four of the seven NAM founders —B73, B97, Ki11, and OH7B— planted in a split-plot design at Musgrave Research Farm. Due to a weather-delayed field planting, all six subplots of Ki11 and one subplot under light treatment of OH7B matured too late and were consequently not harvested. Self-pollinated ears from each subplot (2–7 ears) were harvested at physiological maturity and processed individually to generate mature kernel samples, following the same procedure as in 2018.

Metabolite analysis of physiologically treated embryo and kernel samples

The concentrations of tocochromanols, carotenoids, and chlorophyll *a* were measured in the 24 DAP embryo samples following previously described methods [28, 36]. Briefly, metabolites were extracted from ~ 15 mg of ground 24 DAP embryos, and their quantification was performed using HPLC and fluorometry. Similarly, to assess only the tocochromanol profile of the mature kernel samples, tocochromanols were extracted from ~ 30 mg of ground kernels and quantified using the same method as for the embryo samples. The tocochromanol compounds measured were α-tocopherol (α-T), δ-tocopherol (δ-T), γ-tocopherol (γ-T), α-tocotrienol (α-T3), δ-tocotrienol (δ-T3), and γ-tocotrienol (γ-T3). Additionally, the sum traits total tocopherols (ΣT, ΣT = α-T + δ-T + γ-T), total tocotrienols (ΣT3, ΣT3 = α-T3 + δ-T3 + γ-T3), and total tocochromanols (ΣTT3, ΣTT3 = ΣT + ΣT3) were calculated. The carotenoid compounds measured were neoxanthin, violaxanthin, lutein, zeaxanthin, β-carotene, and other carotenes. In addition, the total carotenoid content was calculated by summing all individual carotenoids.

The mature kernel samples were obtained from individual ears rather than multiple ears. This approach helped in identifying potential outliers at the level of an individual ear, which could result from factors such as premature drying of immature kernels or light leakage due to improper application of the light and dark treatments. Consequently, the raw data underwent visual inspection for extreme outliers using box plots. No samples, when considered collectively on a sub-plot basis, indicated human errors.

To test the degree to which light and dark treatments affected the metabolite profiles of developing and mature kernels, we conducted statistical analyses on three datasets (Dataset S1): metabolite profiles of mature kernel samples averaged by subplot from 2018 and 2019, and embryo samples from 2018. The two datasets of mature kernel metabolites were treated as separate experiments due to the varying number of genotypes evaluated in 2018 and 2019. Levene's test was conducted to test for equality of variances with the SAS PROC GLM procedure in SAS studio release 3.8 (https://www.sas.com/en_us/software/on-demand-for-academics.html). The genotype and treatment terms were tested separately for each phenotype, with the most significant term ($\alpha = 0.01$) selected to correct for unequal variances. With the metabolite levels for each subplot from the 2018 and 2019 mature grain (averaged by subplot) and 2018 embryo samples, we

screened the three datasets separately for outliers. Studentized deleted residuals [35] were estimated using the SAS PROC GLIMMIX procedure, with correction for unequal variances of each phenotype in a mixed linear model (Eq. 1), as follows [9, 46].

$$Y_{ijk} = \mu + \text{genotype}_i + \text{rep}_j + \text{genotype} \times \text{rep}_{ij} + \text{treatment}_k + \text{genotype} \times \text{treatment}_{ik} + \varepsilon_{ijk} \quad (1)$$

in which Y_{ijk} is an individual phenotypic observation; μ is the grand mean; genotype_i is the fixed effect of the ith genotype; rep_j is the random effect of the jth replicate; $\text{genotype} \times \text{rep}_{ij}$ is the random effect of the interaction between the ith genotype and jth replicate; treatment_k is the fixed effect of the kth treatment, $\text{genotype} \times \text{treatment}_{ik}$ is the fixed effect of the interaction between the ith genotype and kth treatment; and ε_{ijk} is the residual error effect assumed to be independently and identically distributed according to a normal distribution with mean zero and variance σ_ε^2, that is ~ iid $N(0, \sigma_\varepsilon^2)$. No significant outlier was identified after a Bonferroni correction of $\alpha = 0.01$. An ANOVA test for the significance of model main effects (Table S1) and Tukey's Honestly Significant Difference (HSD) test for pairwise differences among treatments within each genotype (Table S2) were subsequently performed using Eq. 1 in the SAS PROC GLIMMIX procedure with a correction for unequal variances. For each metabolite phenotype, the Eq. 1 model was used to separately generate best linear unbiased estimator (BLUE) values for the 2018 (mature grain and 24 DAP embryo) and 2019 (mature grain) datasets (Table S3). We calculated a Pearson's correlation coefficient (r) between BLUE values of each pair of metabolites measured in the 24 DAP embryo samples using the 'cor' function and visualized the results with the 'corrplot' package version 0.92 [53] in R version 4.4.0 [38]. All other visualizations were created with the package "ggplot2" [54].

RNA-seq analysis of physiologically treated embryo samples

The embryo samples collected from the physiological experiment conducted in 2018 were utilized for RNA-seq analysis. We ground 30 embryos in liquid nitrogen and then subsampled 100–200 mg of the ground tissue for RNA extraction using a modified hot borate method [52]. RNA was treated with DNase and assessed for purity and quality using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE). High-quality RNA was then used to construct Illumina TruSeq Stranded mRNA libraries (Illumina, San Diego, CA), which were subsequently sequenced in single-end 50 nt mode on an Illumina HiSeq4000 (Illumina, San Diego, CA) at the Resource Technology Support Facility at Michigan State University.

RNA-seq reads were trimmed using Cutadapt version 1.18 [33] with a quality cutoff score of 20 and a minimum read length of 30 nt. Trimmed RNA-seq reads were aligned to the B73 RefGen_v4 reference genome [23] using HISAT2 version 2.1.0 [24] with the following parameters: –min-intronlen 20' and –max-intronlen 60,000 in stranded mode. Read counts per gene were calculated using the count function of HTSeq version 0.6.1p1 [3] with the following parameters: –format = bam, –order = pos, –minaqual = 10, –idattr = ID, –type = gene, and –mode = union in stranded mode.

Differentially expressed genes (DEGs) were identified using the R package DESeq2 version 1.40.2 [32] for each pair of comparisons (control vs. dark, light vs. dark, and control vs. light) within each of the seven genotypes, resulting in a total of 21 comparisons. A false discovery rate corrected *P*-value threshold of 0.05 was applied. Principal component analysis (PCA) was performed on the standardized rlog values obtained from the DESeq2 package [32] with the 'prcomp' function from the R package 'stats' version 4.3.0 and visualized with ggbiplot version 0.55.

CRISPR/Cas9 vector construction and transformation for *POR* knockout generation

We employed CRISPR/Cas9 mutagenesis to create functional knockouts of both *POR1* and *POR2*. In maize, *POR1* (*Zm00001d032576*) and *POR2* (*Zm00001d013937*) have 95.4% CDS sequence identity and 95.2% amino acid sequence identity, assessing sequences obtained from http://maize.uga.edu/ with Blast 2 sequences [2]. The high similarity enabled us to design a single gRNA construct (CRISPR-A) and a double gRNA (CRISPR-B) construct targeting conserved regions within the third exon of both genes (Fig. S2, Methods S1). Agrobacterium-mediated transformation of B104 embryonic callus was performed with the assembled vector (Methods S1) at Iowa State University's Plant Transformation Center as described [12, 13]. The 24 T_0 seedlings from nine events (CRISPR-A generated six events and CRISPR-B generated three events) were shipped to Cornell University where they were self-pollinated or backcrossed with B104 to create BC_1F_1 individuals in a greenhouse at Cornell University's Guterman Bioclimatic Lab in Ithaca, NY, in 2018.

Identification of homozygous *POR* knockout mutants

We used PCR to identify six transgene-free BC_1F_1 individuals with the desired CRISPR-A-induced knockout (Fig. S3, Methods S1), which were then used to generate BC_1F_2 plants homozygous for *por1* and/or *por2* knockout alleles at Cornell University's Musgrave Research Farm in Aurora, NY, in 2019. Twelve seeds were planted in each plot, and plots were 3.05-m long. Sanger sequencing was employed to analyze purified PCR fragments obtained using *POR1*-specific and *POR2*-specific primer pairs (Tables S4, S5), revealing not only the presence or absence but also the type of induced mutations in each BC_1F_2 plant. The primary ear of each BC_1F_2 plant was self-pollinated following rigorously controlled pollination procedures (Methods S1). Selfed ears were harvested at physiological maturity, shelled, and the kernels of each ear were individually maintained for metabolite analysis. Among the six BC_1F_2 families, only CR9.3–5 and CR9.3–41 provided a sufficient number of samples for the different genotype classes, so they were prioritized for HPLC analysis. However, these two families had double mutant (*por1/por1;por2/por2*) plants that produced harvestable ears with only minimal seed.

Experimental design for evaluation of *POR* double knockout mutants

Due to the severity of the double knockout mutant phenotype, we planted seeds from self-pollinated *por1/+;por2/por2* plants of the CR9.3–5 and CR9.3–41 families in a greenhouse environment. The *por1/+;por2/por2* mutant was selected for the experiment due to its visibly stronger leaf color mutant phenotype compared to the *por1/por1;por2/+* mutant. In 2020, seeds from self-pollinated BC_1F_2 WT (*POR1/POR1;POR2/POR2*) and *por1/+;por2/por2* plants of the CR9.3–5 and CR9.3–41 families were sown in pots in the Guterman greenhouse to generate BC_1F_3 plants. For each of the two families, leaf tissue from seedlings derived from seeds of self-pollinated ears of *por1/+;por2/por2* BC_1F_2 plants was collected for genomic DNA isolation, followed by PCR and Sanger sequencing as previously described, to classify BC_1F_3 plants as *por2/por2*, *por1/+;por2/por2*, or *por1/por1;por2/por2*. All individual BC_1F_3 plants of the four genotype classes, including WT, from both families, were self-pollinated. For each of the two families, a self-pollinated primary ear was harvested from each of four WT and four p*or1/+;por2/por2* BC_1F_3 plants at 24 DAP, followed by immediate freezing in liquid nitrogen after removing the husk leaves. For each frozen ear, the middle section was hand-shelled on dry ice, and the collected kernels were stored at − 80 °C. The remaining BC_1F_3 plants were retained for mature kernel harvest. However, only the CR9.3–5 family had *por1/por1;por2/por2* plants with harvestable ears. Consequently, only the selfed primary ears of plants from this family (4 to 11 plants per genotypic class) were harvested at physiological maturity and processed individually to generate mature kernel samples for analyzing metabolite levels.

After genotyping the endosperm tissue (Methods S1), frozen kernels were dissected on a metal plate over dry ice to collect embryo and endosperm tissues. Kernels from each plant contributed 15 embryos and five endosperms yielding four sets of samples for every genotype class and tissue type. To ensure ample tissue for metabolite extraction, the first two sets and the last two sets were combined, resulting in the creation of two biological replicates for each genotype class, each comprising either 30 embryos or 10 endosperms. For RNA extraction, 10 embryos were collected from dissected frozen kernels from each plant in the CR9.3–5 family.

Metabolite analysis of *POR* knockout mutant embryo, endosperm, and kernel samples

We performed a metabolite analysis of embryo, endosperm, and mature kernel samples collected from field and greenhouse evaluations of WT and *POR* knockout mutants (Dataset S1). The procedures for tocochromanol, carotenoid, and chlorophyll *a* quantification are described in an earlier section. The only modification is that the total carotenoids phenotype was quantified as total area/mg fresh tissue.

Statistical analysis was performed to assess differences between CRISPR/Cas9 mutants and/or WT. All tests were performed in R (version 4.4.0). The developing embryos and endosperms from the 2020 greenhouse experiment and the mature kernels from the 2019 field experiment were analyzed with the following model:

$$Y_{ij} = \mu + \text{family}_i + \text{genotype}_j + \text{family}_i \times \text{genotype}_j + \varepsilon_{ij} \quad (2)$$

in which Y_{ij} is an individual phenotypic observation; μ is the grand mean; family is the fixed effect of the ith family; genotype_i is the fixed effect of the jth genotype; genotype × family is the fixed effect of the interaction between the ith family and jth genotype; and ε_{ij} is the residual error effect assumed to be independently and identically distributed according to a normal distribution with mean zero and variance σ_ε^2, that is ~ iid $N(0, \sigma_\varepsilon^2)$. In the analysis of mature kernels from the 2020 greenhouse experiment, the variables of family and the interaction between family and genotype were removed, as only the family CR9.3–5 was analyzed. This is because the family CR9.3–41 did not produce enough seed for mature kernel analysis after the harvest of developing kernels. Studentized residuals were calculated using a Bonferroni correction with

α = 0.01 in the package 'car'. Two outliers were identified and removed: two chlorophyll *a* data points from endosperm tissue of the CR9.3–4.1 family, causing the removal of the family and the interaction term in subsequent analysis for this metabolite-tissue combination. Levene's test and ANOVA (type III) (Table S6) were performed with the package 'car' version 3.1–2 [11]. Tukey's HSD test (Table S7) was performed with the package "agricolae" [10].

RT-qPCR and RNASeq analysis of *POR* expression in double knockout mutant embryos

The four biological replicates of collected *por1/por1;por2/por2* and WT embryos from the CR9.3–5 family were used for isolating total RNA. Briefly, each replicate comprising 10 embryos underwent grinding in a 1.5 mL conical bottom microcentrifuge tube using a pellet pestle (Fisher scientific, Pittsburgh, PA, USA). The grinding process ensured that the bottom of the tube remained submerged in liquid nitrogen, maintaining low temperatures throughout the procedure. We subsampled 100 mg of ground tissue for total RNA extraction using a modified hot borate method [52], followed by DNA removal employing an Ambion TURBO DNA-free Kit (Thermo Fisher Scientific, Waltham, MA, USA). The concentration and purity of RNA samples were assessed using an Epoch 2 Microplate Spectrophotometer (Agilent, Santa Clara, CA, USA), and the integrity of RNA samples was evaluated through gel electrophoresis. One microgram of total RNA was reverse-transcribed using the PrimeScript 1st strand cDNA Synthesis Kit (Takara, Kyoto, Japan) and oligo-dT according to manufacturer's recommendations. Primer pairs were designed to amplify selected regions of *POR1*, *POR2*, and *POR3* (*Zm00001d001820*) using Primer Premier 5 [25] (Table S4). RT-qPCR was performed in 96-well plates with a QuantStudio 7 Pro Real-Time PCR system (Thermo Fisher Scientific, Waltham, MA, USA), employing Power SYBR Green PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) (Table S5). The amplification conditions were as follows: 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 60 s. Each RT-qPCR analysis was performed in quadruplicate. Relative expression levels were calculated by normalizing to the reference gene *ACTIN1* (*Zm00001d010159*) using the comparative CT method [43]. The $2^{-\Delta\Delta CT}$ method [31] was used to analyze RT-qPCR in R (version 4.4.0), with the average expression of *POR1* from WT embryos used as the control. ANOVA and Tukey HSD test were performed to compare the relative quantification of WT to double knockout mutant within and across each gene using the 'stats' package in R (version 4.4.0).

High-quality RNA from two of the four biological replicates of collected *por1/por1;por2/por2* and WT embryos was used to construct Illumina TruSeq Stranded mRNA libraries (Illumina, San Diego, CA), which were subsequently sequenced in paired-end 150 nt mode on an Illumina Novaseq 6000 (Illumina, San Diego, CA) by Psomagen (Rockville, MD). RNA-seq reads were trimmed using Cutadapt version 4 [33] with a quality cutoff score of 30 and a minimum read length of 75 nt. Trimmed RNA-seq reads were aligned to the B73 RefGen_v4 reference genome [23] using HISAT2 version 2.1.1 [24] with the following parameters: –dta-cufflinks –max-intronlen 5000 and –rna-strandness RF. Read counts per gene were calculated using the count function of HTSeq version 0.13.5 [3] with the following parameters: –format = bam, –order = pos, –stranded = reverse, –minaqual = 10, –idattr = ID, –type = gene, and –mode = union. Differential expression analysis was performed as previously described.

Results

Tocopherol synthesis is light-dependent within developing embryo tissue

POR enzyme activity requires light as a co-substrate [17] and to understand the relationship between chlorophyll synthesis and tocopherol levels in non-photosynthetic kernel tissue, we conducted a physiological experiment on developing grain by exposing the ear to light and dark treatments (Fig. S1). At 12 days after pollination (DAP), we applied both treatments to a subset of maize NAM founder lines by covering the ears with aluminum foil for the dark treatment or removing 3–4 outer husks for the light treatment. In 2018, we tested seven NAM founders with contrasting allelic effects [5] of one or both *POR* loci, while a further subset of three NAM founders was evaluated in 2019. Metabolite analysis of mature kernels revealed that in both 2018 and 2019, the dark treatment reduced total tocopherols compared to light-treated samples by a mean of 82% and control samples by 78% across genotypes (Figs. S4, S5). Similar trends were observed for individual tocopherols (α-tocopherol, δ-tocopherol, γ-tocopherol) (Figs. S4-S6). We dissected embryos from 24 DAP developing kernels (from 2018) and found both chlorophyll *a* and tocopherols were reduced by a mean of 96% and 88%, respectively, in dark-treated samples compared to light-treated samples (Fig. 2a, b). There was a strong positive correlation ($r = 0.6$) between the concentrations of chlorophyll *a* and total tocopherol (Fig. S7). In 24 DAP embryos, total tocopherol levels were 40- to 200-fold higher than chlorophyll *a* levels across all genotypes in both the control and light treatments

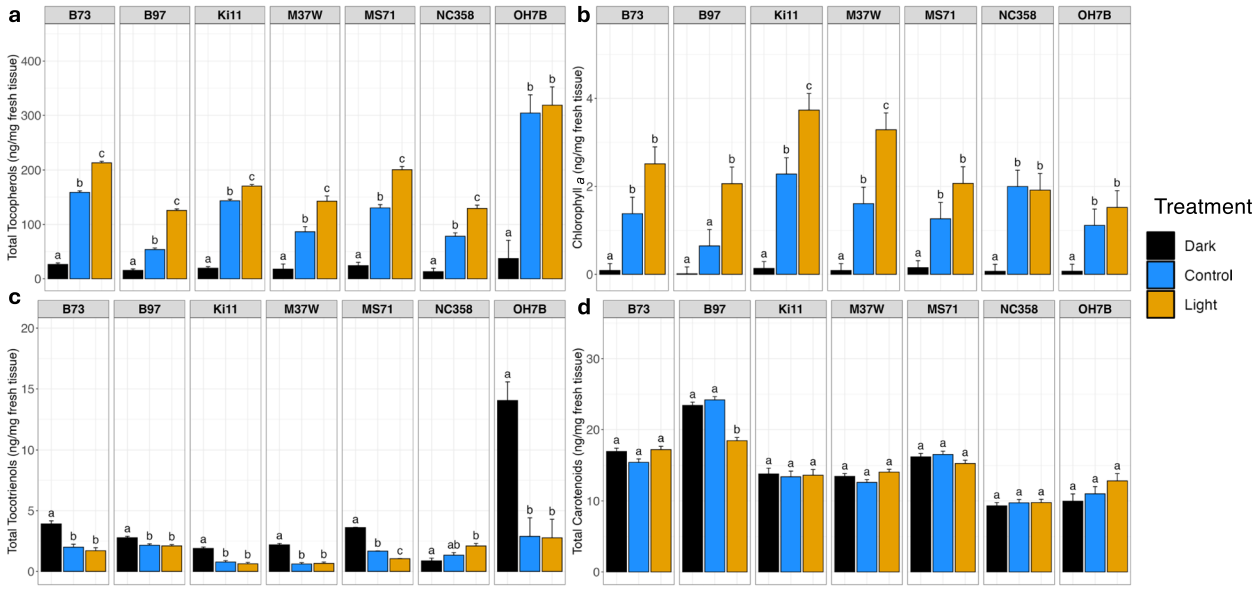

Fig. 2 Untransformed best linear unbiased estimators (BLUEs) of metabolites from developing embryo tissue. **a** Total tocopherol concentration. **b** Chlorophyll a concentration. **c** Total tocotrienol concentration. **d** Total carotenoid concentration. Different letters denote *P*-value < 0.05 in Tukey's HSD for within-genotype treatment comparisons. Error bars represent standard error estimates

(Fig. 2). While the reduction in tocopherol abundance between dark-treated and control samples was consistent across genotypes, years, and tissues, the changes in total tocopherols between light-treated and control samples were inconsistent. For example, in 2018, six of the seven genotypes had a mean increase of 54% in embryo tocopherol levels in response to light treatment compared to control treatment, but in only two lines (B97 and MS71) did the significant increase in tocopherols persist in mature kernels (Figs. 2, S4).

Tocopherols and tocotrienols rely on many of the same enzymes for their synthesis (Fig. 1), but in contrast to tocopherols, the four tocotrienol traits (α-tocotrienol, δ-tocotrienol, γ-tocotrienol, and total tocotrienol) showed no consistent significant differences among different treatments across tissue types (Figs. 2c, S3-S5). Embryo samples from the dark treatment exhibited an average increase of 156% in total tocotrienols compared to light-treated or control samples in all but one genotype (NC358). However, this trend was again not apparent in the mature kernels (Figs. S4, S5). To assess whether the physiological treatment affected an unrelated biochemical pathway, we also measured carotenoids (β-carotene, lutein, zeaxanthin, neoxanthin, violaxanthin, and total carotenoids) in embryo samples. None of the carotenoid traits showed consistent patterns of change between treatments (Figs. 2d, S8), consistent with the treatment effects being specific to tocopherol synthesis.

POR1 and *POR2* RNA levels in the embryo are not light-dependent

Substrates for tocopherols, tocotrienols, and carotenoids are derived from the same pathways: HGA from the shikimate pathway for tocopherols and tocotrienols, and for all three compound classes, GGDP from the MEP pathway. Given the impact of the dark treatment on tocopherol synthesis in both embryo and mature kernels and the lack of impact on carotenoid or tocotrienol levels, our results are consistent with the shikimate and IPP pathways being unaffected. We hypothesized that the expression of a priori genes in chlorophyll biosynthesis pathways, including *POR1* and *POR2*, and other related pathways (Table S8) could explain the changes in tocopherols resulting from light and dark treatments. To test this hypothesis, we performed RNA-seq on bulked 24 DAP embryos dissected from fresh kernels for each genotype and experimental condition combination for the seven NAM founders in the 2018 experiment. Unlike tocopherols and chlorophyll *a*, which showed light-dependent variation, principal component analysis showed that the variation in gene expression was genotype- rather than treatment-dependent (Fig. S9). We assessed gene-level differences between treatments within a genotype to determine if differences in light-dependent tocopherol synthesis could be explained by differentially expressed genes (DEGs). Surprisingly, genes encoding activities of the POR/LIL3/CHS/GGR complex were not found to be differentially expressed between treatments using DESeq2 [32] (Fig. 3).

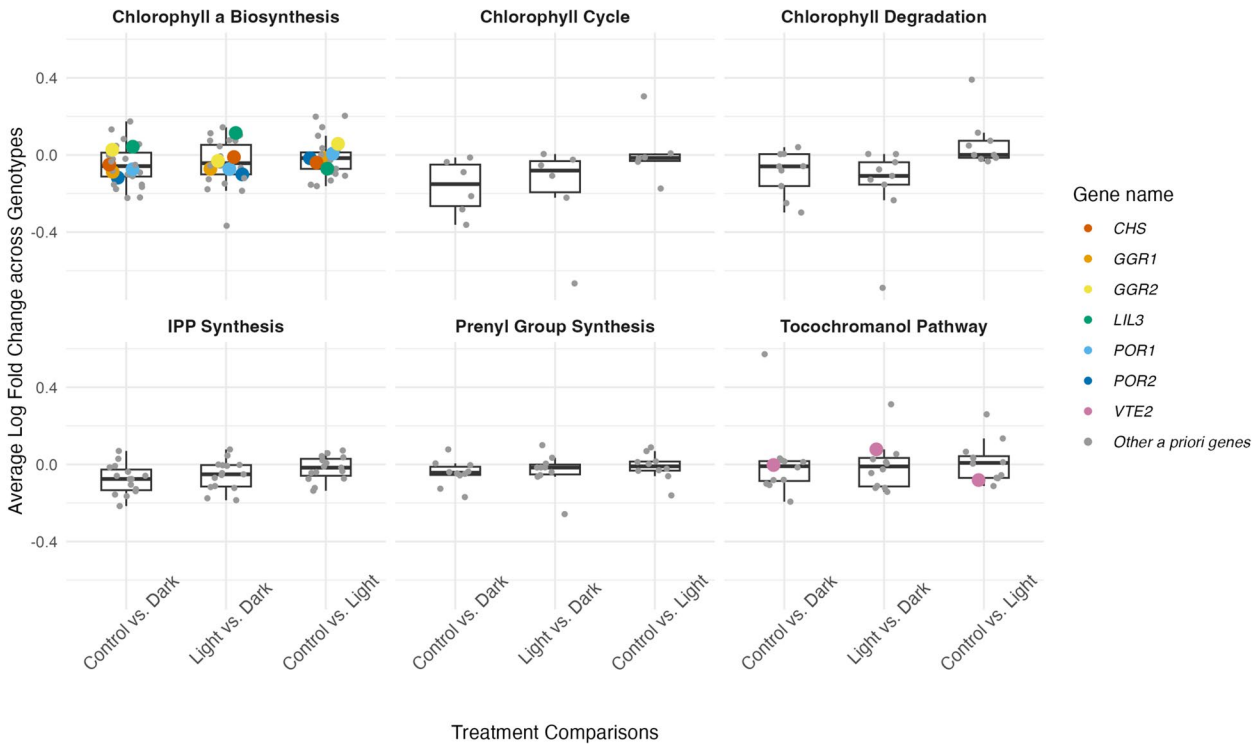

Fig. 3 Expression differences of genes within a priori pathways (Table S8) between physiological treatments in 24 DAP embryo tissue. Box plots show the average log fold change in gene expression across genotypes on the y-axis between two treatments as denoted on the x-axis. Each point represents the average log-fold change in expression between treatments for a single gene. Differences in the expression of *CHS, GGR1, GGR2, LIL3, POR1*, *POR2*, and *VTE2* between treatments are denoted by color

However, the expression of *POR* genes within treatments and across genotypes did show strong positive correlations (*r*-values between 0.76–0.94) with joint-linkage allelic effect estimates [5] (Fig. S10), indicating that the genotypic effect remained consistent across treatments. Of the 125 a priori genes assessed (Table S8), only two genes (a coproporphyrinogen oxidase (*NEC4*) from the chlorophyll *a* biosynthesis pathway; and a tyrosine aminotransferase (*NAAT1*) from the tyrosine degradation pathway) were differentially expressed in only one of three treatment pairs (dark vs. control) in only one of the seven genotypes (M37W). Overall, the RNA-seq data indicates the light-dependent chlorophyll and tocopherol changes are not due to changes in the expression of the 125 a priori genes.

Generating *por1/por2* knockout mutants by CRISPR/Cas9

We generated CRISPR/Cas9 knockouts of *POR1* and *POR2* to directly determine their contributions to chlorophyll synthesis within the embryo and any effects on tocopherol abundance. Two CRISPR lines were selected for detailed analysis, CR9.3–5 and CR9.3–41 (Figs. S11 and S12). *POR1* contained a one-base deletion in both lines, causing a frameshift at amino acid 138 and a premature stop at amino acid 188. In line CR9.3–5, *POR2* had a one-base insertion, resulting in a frameshift at amino acid 138 and premature stop at amino acid 253. In line CR9.3–41, *POR2* had a nine-base deletion that removed amino acids 138–140, along with a one-base insertion resulting in a frameshift at amino acid 203 and a premature stop at amino acid 213. All of these mutations eliminated conserved amino acids essential for substrate binding and catalysis, resulting in null mutations [6, 41, 59]. After the selection of suitable transgene-free *por1* and/or *por2* knockout BC_1F_1 individuals (Fig. S3), 380 transgene-free BC_1F_2 seedlings successfully germinated in the field and could be visually separated into three groups at the seedling stage: 308 green-leaf individuals, 44 individuals that exhibited a green-yellow leaf color, and 28 yellow-leaf individuals (Fig. 4). At maturity the green-yellow plants became indistinguishable from WT, while the yellow-leaf plants remained yellow with a small and stunted phenotype. PCR genotyping results revealed that the yellow-leaf plants were homozygous double mutants (*por1/por1;por2/por2*), the green-yellow leaf plants carried a single functional allele of *POR1* (*por1/+;por2/por2*) and the green-leaf plant group contained all other possible genotypes including plants with

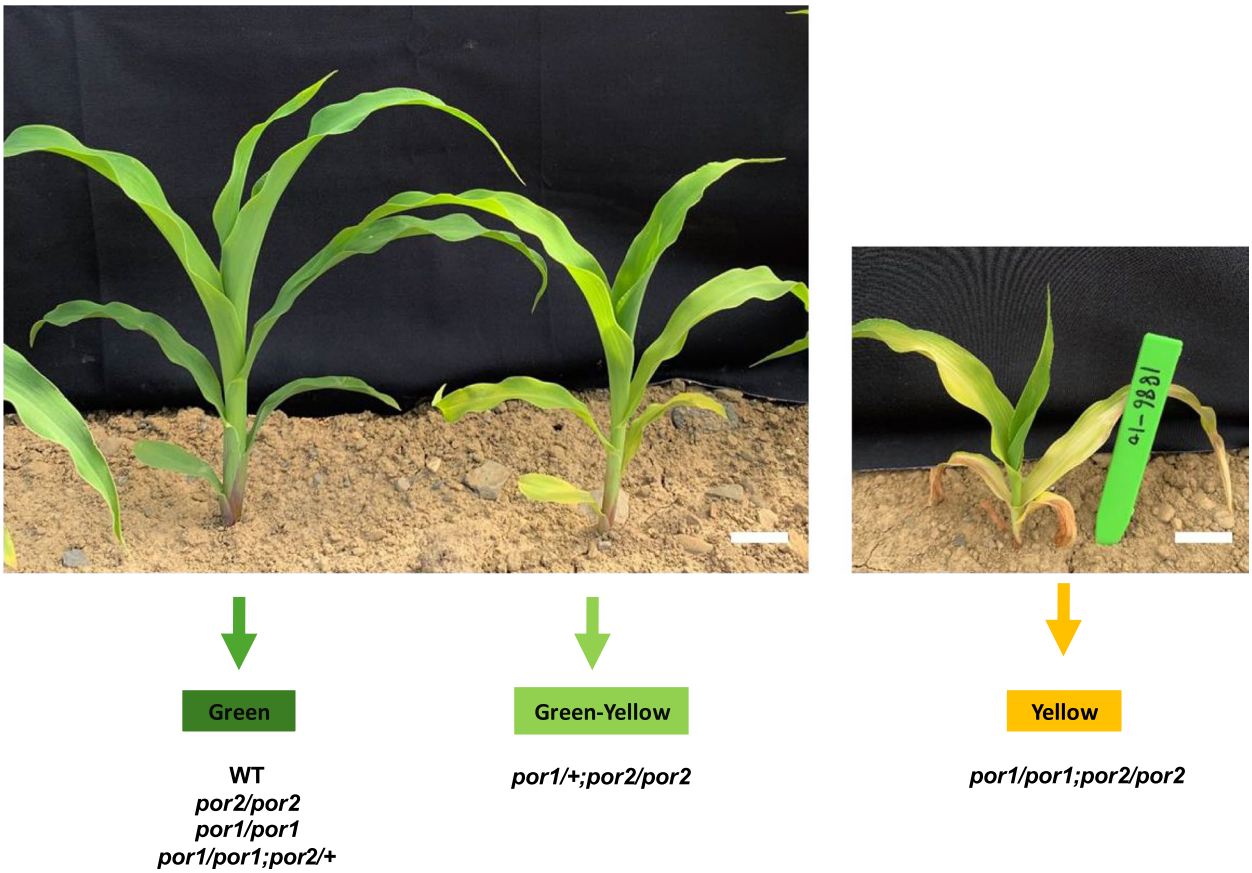

Fig. 4 Green-yellow plant color segregation was observed after emergence in BC_1F_2 progeny planted in 2019. Scale bars are 2 cm. Genotypes are grouped according to leaf color

a single functional allele of *POR2* (*por1/por1;por2/+*). This difference in leaf color suggests that the *por2* mutation had a stronger effect on maize leaf chlorophyll content during development in the B104 background, consistent with results from [29]. Since *por1* (*por1/por1*) and *por2* (*por2/por2*) homozygous single mutants were also indistinguishable from WT (Fig. 4), we speculated that the function of POR1 and POR2 were complementary in leaves and that they had dose effects on maize leaf chlorophyll content.

Metabolite analysis of *por1*/*por2* mutants

We harvested mature kernels from the field-grown, self-pollinated BC_1F_2 plants with WT, *por1* single mutant (*por1/por1*), *por2* single mutant (*por2/por2*), and *por1* or *por2* monoallelic double mutant genotypes (*por1/+;por2/por2* and *por1/por1;por2/+*, respectively) from CRISPR families CR9.3–41 and CR9.3–5 (Figs. S11, S12) and analyzed them for tocochromanols and carotenoids by HPLC. All tocochromanol traits and total carotenoids in *por1* and *por2* single mutants were similar to WT (Figs. 5a, S13). Tocopherol levels in *por1/+;por2/por2* and *por1/por1;por2/+* kernels were reduced by ~ 41% and ~ 36%, respectively, compared to WT. Total tocotrienol and carotenoid levels in the kernel were ~ 30% and ~ 36% higher, respectively, in *por1/por1;por2/+* mutants and unchanged in *por1/+;por2/por2* mutants compared to WT (Fig. 5a). Kernels of these monoallelic double mutants are still segregating for multiple genotypes at their *por1* and *por2* loci, respectively, and the wild type segregants at each locus could partially obscure their impact on tocochromanols and carotenoids. Unfortunately, the *por1 por2* double homozygous mutant plant (*por1/por1;por2/por2)* was severely stunted in the field and produced minimal seed.

To determine the consequences of the *por1/por1;por2/por2* mutant genotype on tocopherol, tocotrienol, and carotenoid levels in mature grain, we grew and harvested mature kernels from WT, *por2/por2*, *por1/+;por2/por2,* and *por1/por1;por2/por2* mutants in a greenhouse environment. Despite the reduced size of the *por1/por1;por2/por2* mutant compared to WT (Fig. S14), it yielded substantial amounts of phenotypically similar mature kernels for HPLC analysis (Fig. S15). Tocopherol levels in

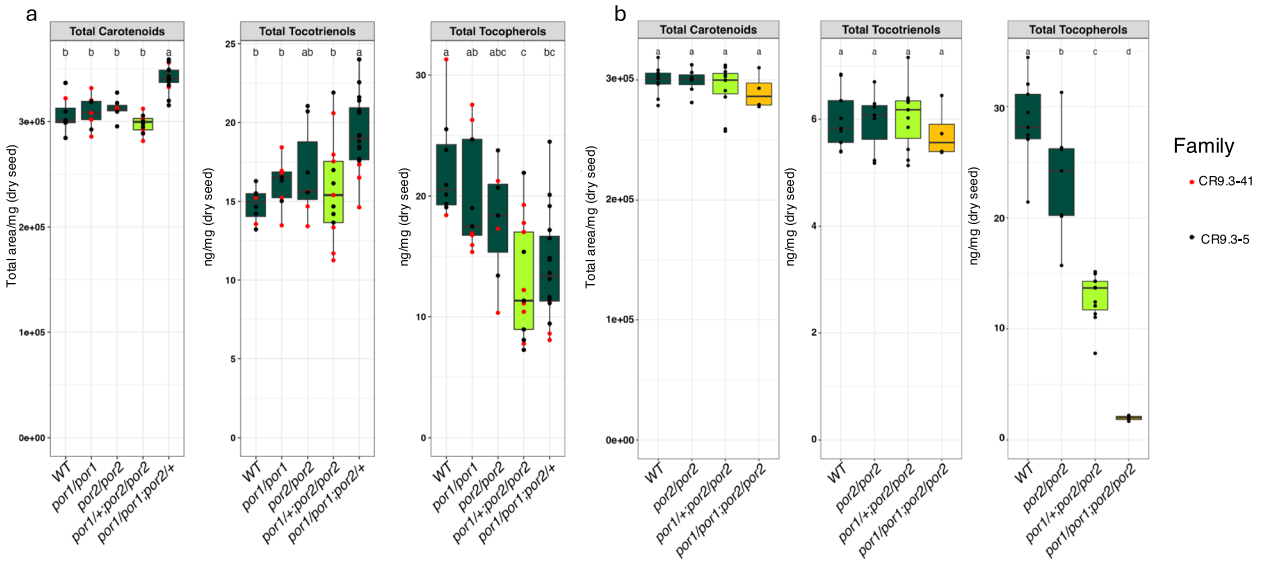

Fig. 5 Effects of *por1* and/or *por2* knockout mutations on tocochromanols and carotenoids in mature kernels. **a** The metabolomic results of 53 samples from family CR9.3–41 (indicated by red dots) and CR9.3–5 (indicated by black dots) from the 2019 field experiment. **b** The metabolomic results of 31 samples from family CR9.3–5 from the 2020 greenhouse experiment. Different letters denote *P*-value < 0.05 in Tukey's HSD for genotype comparisons. The boxplot color reflects the leaf color of the genotypes grown in the 2019 field experiment (Fig. 4)

greenhouse-grown plants generally agreed with those in field-grown genotypes (Figs. 5b, S16). Relative to the WT, the *por2* single mutant showed the smallest reduction (18%), the *por1/+;por2/por2* mutant had a ~ 55% reduction, and the *por1/por1;por2/por2* mutant was the most severe with a ~ 93% reduction in tocopherol levels (Fig. 5b). In contrast, the carotenoid and tocotrienol levels of the three mutant genotypes were indistinguishable from WT (Fig. 5b).

To assess if the stunted phenotype of the *por1/por1;por2/por2* mutant contributes to the exceedingly low tocopherol in mature kernels (Fig. 5b), we collected and genotyped developing kernels (24 DAP) from the self-pollinated ears of greenhouse-grown *por1/+;por2/por2* mutants, allowing us to generate *por1/por1;por2/por2* kernels from phenotypically WT plants (Fig. S17). We used single kernel genotyping to identify kernels carrying *por1/por1;por2/por2*, *por2/por2*, and *por1/+;por2/por2* genotypes and then performed frozen kernel dissection to collect embryo and endosperm tissues for RT-qPCR analysis of the *POR* genes and metabolite analyses. Expression of *POR1* and *POR2* in *por1/por1;por2/por2* mutant embryos was significantly decreased compared to WT, while average *POR3* expression remained unchanged between the genotypes and was only 0.2% of WT *POR1* expression. (Fig. S18). No other a priori genes besides *POR1* and *POR2* (Table S8), including *VTE* genes and *GGR* homologs, were differentially expressed between the *por1/por1;por2/por2* mutant and the WT (Fig. S19), similar to the RNAseq results from the physiological experiment (Fig. 3). In embryo tissues, the *por1/por1;por2/por2* mutant displayed a ~ 76–83% reduction in total tocopherols, along with a ~ 60–64% increase in total tocotrienols compared to the WT (Figs. 6a, S20). The *por1/por1;por2/por2* mutant also had reduced tocopherol levels compared to the *por2/por2* mutant and the *por1/+;por2/por2* mutant (~ 76–77% and ~ 64–70% respectively). Additionally, the *por1/por1;por2/por2* mutant embryos showed a > 98% reduction in chlorophyll *a* compared to the WT (Fig. 6a). Comparable levels of total carotenoids were found in double mutant embryos compared to WT. In contrast to embryo tissue, the metabolite levels in endosperm tissue were consistent across genotypes (Fig. 6b). Endosperm tissue accumulated chlorophyll *a* below our limit of detection, at least 100-fold lower than that observed in embryo tissue (Fig. 6b). By dissecting endosperm and embryo tissue from genotypically distinct kernels from the same ear, we show that tocopherol levels in maize grain are dependent on the genotype of the embryo. Overall, the functional knockout of *POR1* and *POR2* supports the reliance of tocopherol biosynthesis on chlorophyll biosynthesis within the embryo.

Discussion

The core tocopherol biosynthesis pathway is well known, but the mechanism(s) supplying the phytol group, a key precursor for tocopherol synthesis, has yet to be understood in both photosynthetic and non-photosynthetic tissues, such as maize grain. Despite maize embryos having

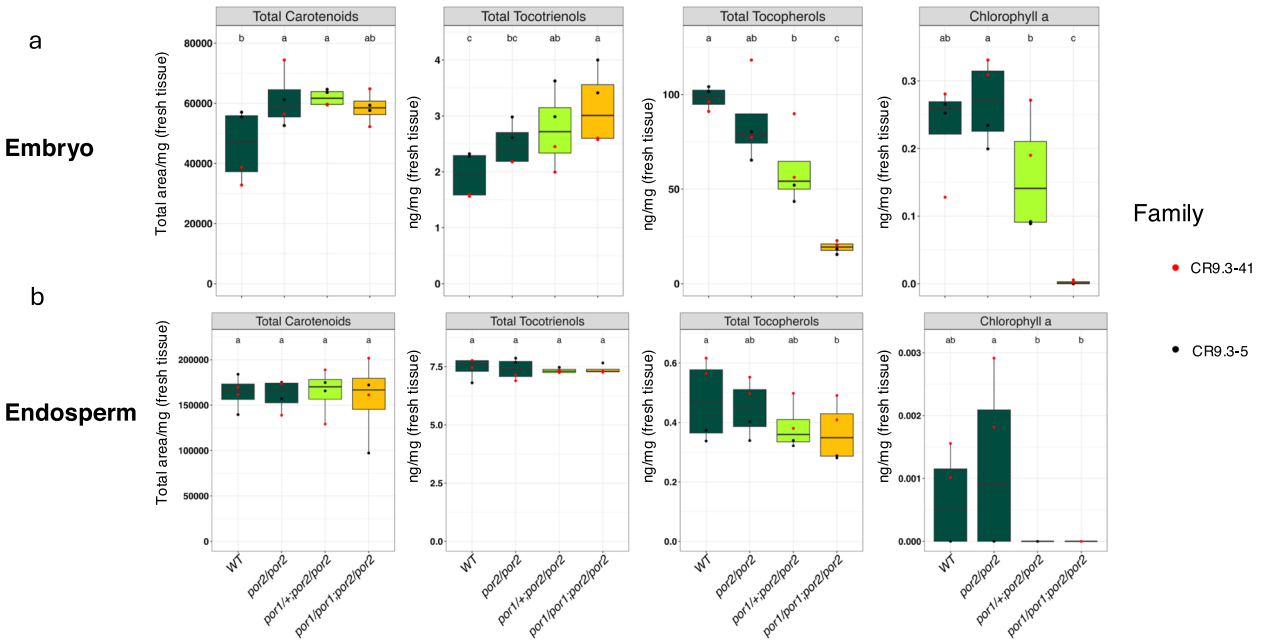

Fig. 6 Effects of *por1* and/or *por2* knockout mutations on tocochromanols, carotenoids, and chlorophyll *a* in developing kernel tissues. The metabolomic results from family CR9.3–41 (indicated by red dots) and CR9.3–5 (indicated by black dots) from 24 DAP embryo (**a**) and endosperm (**b**). Different letters denote *P*-value < 0.05 in Tukey's HSD for genotype comparisons. The boxplot color reflects the leaf color of the genotypes grown in the 2019 field experiment (Fig. 4)

extremely low chlorophyll levels (~ 500-fold lower than leaves) and tocopherol levels up to ~ 1000-fold higher than chlorophyll *a*, association studies in maize indicated that *POR1* and *POR2*, which encode a key activity in chlorophyll biosynthesis, underlie major QTL for natural variation in total tocopherol levels in mature grain [5, 55]. To understand the contribution that chlorophyll biosynthesis makes to tocopherol biosynthesis within the maize embryo, we altered POR1 and POR2 activity by two complementary approaches: physiological inhibition of POR activity by darkness and single and double CRISPR/Cas9-mediated knockouts of the two genes. This resulted in up to a 93% reduction in tocopherol content in mature kernels, indicating tocopherol biosynthesis is almost entirely dependent on chlorophyll biosynthesis within the non-photosynthetic maize embryo.

In the first approach, we took advantage of the light-dependent enzyme activity of POR1/POR2 to inhibit chlorophyll synthesis in vivo [39]. By withholding light during kernel development, we could assess the dependence of tocopherol synthesis on chlorophyll synthesis. Using a diverse subset of the NAM founders with differing *POR1* and/or *POR2* allelic effects [5], we found that the abundance of chlorophyll *a* and total tocopherols in dark-treated embryo tissue decreased by 92–99% and 87–90%, respectively, compared to light-treated embryo tissue (Fig. 2a and b). Because light levels strongly impact *POR* gene expression in vegetative tissues [47], we assessed if light deprivation impacts *POR1* and *POR2* gene expression in embryo tissue. Surprisingly, compared to control, the dark treatment did not significantly reduce the mRNA levels of *POR1, POR2*, or other a priori genes for IPP synthesis, chlorophyll metabolism, and tocopherol biosynthesis (Table S8) within embryos across all genotypes (Fig. 3). This is consistent with the impact of the dark treatment on tocopherol and chlorophyll *a* abundance being caused by translational or posttranslational impacts on POR activity, or the stability/activity of other members of the POR/CHS/GGR/LIL3 complex [18, 48, 49, 58], as well as potential interactions with VTE biosynthetic enzymes. The low tocopherol levels remaining in dark-treated grain samples could arise from multiple sources including a trace amount of light leakage into the ear, the accumulation of tocopherols or their precursors prior to initiating dark treatment at 12 DAP, and/or from the availability of small amounts of GGDP-derived PDP that might not rely on chlorophyll-derived phytol.

In the second approach, we targeted the third exon of both *POR1* and *POR2* to create homozygous single and double knockout mutants with severely truncated, non-functional enzymes. Compared to the WT, the *por1 por2* double homozygous mutant (*por1/por1;por2/por2*) had a metabolomic phenotype similar to that of dark-treated ears: a ~ 93% reduction of tocopherols in mature

kernels (Fig. 5b) and a 76–83% and 98–100% reduction in tocopherols and chlorophyll *a*, respectively, in 24 DAP developing embryos (Fig. 6a). The low levels of tocopherol in mature grain of the *por1/por1;por2/por2* double mutant could be due to the activity of POR3, the third POR homolog in maize that is mainly expressed during the first several days of grain development [7]. The similarity of reductions in grain tocopherol levels between the single *por* mutants (*por1/por1* or *por2/por2*) and between the monoallelic double mutants (*por1/+;por2/por2* or *por1/por1;por2/+*) indicate that *POR1* and *POR2* play complementary roles in tocopherol synthesis within maize grain. In contrast, leaves of *por1/por1;por2/+* mutant seedlings were visibly paler than the *por1/+;por2/por2* mutant (Fig. 4), consistent with POR2 making a greater contribution to chlorophyll biosynthesis in leaf tissue than POR1 [56]. Both the genetic and physiological inhibition of POR provide conclusive evidence that nearly all tocopherol produced in maize kernels depends on chlorophyll biosynthesis within the embryo.

Tocopherols and tocotrienols are closely related chemically and biosynthetically with only a single gene differing in their pathways (Fig. 1); tocopherols require VTE2 for the condensation of HGA and PDP, while tocotrienols require HGGT for condensation of HGA and GGDP. Unlike the stark reduction in tocopherols, *por* mutations and dark treatments increased the tocotrienol content in developing embryos, suggesting that the availability of HGA was not decreased by the enzymatic activity of POR (Figs. 2d, and 6a). Similarly, GGDP levels also appear to be unaffected, as carotenoids—which depend on GGDP for their synthesis—remain unchanged in both experiments (Figs. 2c, and 6a). The increase in embryo tocotrienols is likely due to the ability of VTE2 to utilize its less preferred substrate, GGDP, when the levels of GGDP greatly exceed that of PDP [57]. The opposite response of tocotrienol and tocopherol synthesis to light deprivation is consistent with decreased availability of chlorophyll-derived PDP specifically for tocopherol synthesis being impacted rather than a reduction in other tocochromanol pathway substrates or *VTE2* expression (Fig. 3). Given the large stoichiometric differences between chlorophyll *a* and tocopherol levels in developing maize kernels (Figs. 2a,b, and 6a), our data supports the hypothesis [5] that chlorophyll participates in a biosynthetic cycle within embryos. This biosynthetic cycle generates many molecules of phytol per molecule of chlorophyll—unlike chlorophyll degradation, an irreversible process that produces only a single molecule of phytol per molecule of chlorophyll [20].

A prior study [56] used tocopherol data from reciprocal hybrids with differing *POR2* alleles, along with in vitro feeding of phytol to isolated, developing kernels, to hypothesize that the phytol for tocopherol synthesis in embryos is generated in leaf tissue and transported to the embryo by an unknown mechanism. Our data do not support this hypothesis, instead, they indicate the source of phytol for tocopherol synthesis in embryos is the embryo itself. The physiological experiment did not affect leaves, and if phytol was transported from leaves for synthesis in embryos, one would expect no impact on embryo tocopherol synthesis. Embryo and mature kernel tocopherol levels were instead severely decreased (Figs. 2, S4 and S5). In a second independent approach, we self-pollinated the phenotypically robust *por1/+;por2/por2* mutant in the greenhouse to generate ears segregating for *por1/por1;por2/por2, por1/+;por2/por2,* and *por2/por2* genotypes, which were identified by single kernel genotyping. In this experiment, kernel genotypes on each ear experience identical maternal effects and environmental conditions during development and, if phytol for tocopherol synthesis in embryos were transported from other tissue, one would again expect no impact on embryo tocopherol levels. Instead, a large reduction in tocopherol and chlorophyll levels was observed in *por1/por1;por2/por2* embryos and mature kernels compared to *por2/por2* and *por1/+;por2/por2*. These two experiments conclusively demonstrate that the phytol produced within the embryo, rather than the transport of phytol from maternal leaf tissue, determines embryo tocopherol levels (Fig. 6a).

In summary, we can now conclude that ~ 93% of total tocopherol in mature maize kernels is dependent on chlorophyll-derived phytol synthesized within embryo tissue. Defining the contribution of chlorophyll biosynthesis to tocopherol synthesis in the maize embryo indicates that *POR1* and *POR2* are primary breeding targets for elevating the total tocopherol content in maize, and likely other cereal grains. Selecting superior *POR1* and *POR2* alleles for total tocopherols, along with other loci that control the types of tocopherols accumulated (*VTE3* and *VTE4*), should result in an optimized grain profile that has both increased tocopherol content and enhanced vitamin E activity to benefit agriculture and human health.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06267-6.

Supplementary Material 1. Fig. S1. Representative image of the physiological experiment from 2018. Fig. S2. Map of the assembled plasmids of pRGEB32Bar-*POR1/2*-double used to introduce Cas9 and gRNAs into B104. Fig. S3. Transgene PCR screen of six selected BC_1F_1 plants using the Bar-based primer bar1. Fig. S4. Untransformed best linear unbiased estimators (BLUEs) of tocochromanols from mature kernels in 2018. Fig. S5. Untransformed best linear unbiased estimators (BLUEs) of tocochromanols from mature kernels in 2019. Fig. S6. Untransformed best linear unbiased estimators (BLUEs) of tocochromanols from developing embryo tissue

in 2018. Fig. S7. Pearson's correlation plot for 24 DAP embryo metabolites. Fig. S8. Untransformed best linear unbiased estimators (BLUEs) of carotenoids from developing embryo tissue in 2018. Fig. S9. Principal component analysis (PCA) plot of the rlog of counts data from RNA-seq analysis of 24 days after pollination (DAP) embryo tissue. Fig. S10. Gene expression of *POR1* and *POR2* from 24 days after pollination (DAP) embryo tissue correlates with joint linkage-quantitative trait loci (QTL) allelic effect estimates. Fig. S11. CRISPR/Cas9 induced knockout mutations for *POR1*. Fig. S12. CRISPR/Cas9 induced knockout mutations for *POR2*. Fig. S13. Effects of *por1* and/or *por2* knockout mutations on tocochromanols in mature kernels. Fig. S14. Visual assessment of whole plants from the CRISPR/Cas9 knockout greenhouse experiment. Fig. S15. Visual assessment of ears and kernels from the CRISPR/Cas9 knockout mutants. Fig. S16. Effects of *por1* and/or *por2* knockout mutations on tocochromanols in mature kernels. Fig. S17. Selfed WT and *por1/+;por2/por2* plants from the 2020 greenhouse experiment. Fig. S18. RT-qPCR results of POR1, POR2, and POR3 from 24 days after pollination (DAP) embryos from the CR9.3-5 family. Fig. S19. Expression differences of genes within a priori pathways (Table S8) between the wildtype and the *por1/por1;por2/por2* mutant. Fig. S20. Effects of *por1* and/or *por2* knockout mutations on tocochromanols in developing embryo tissues. Table S1. *P*-values from ANOVA for tocochromanols, carotenoids, and chlorophyll a from the 24 DAP embryos (2018), and tocochromanols from mature kernels (2018 and 2019). Table S2. *P*-values from Tukey's HSD tests for tocochromanols, carotenoids, and chlorophyll a from the 24 DAP embryos (2018), and tocochromanols from mature kernels (2018 and 2019). Table S3. Best linear unbiased estimators (BLUEs) of tocochromanols, carotenoids, and chlorophyll a from the 24 DAP embryos (2018), and tocochromanols from mature kernels (2018 and 2019). Table S4. List of primer pairs used in the study. Table S5. PCR and RT-qPCR procedures used in the study. Table S6. *P*-values from ANOVA for tocochromanols and carotenoids from the 2019 field experiment and tocochromanols, chlorophyll *a*, and carotenoids from 2020 greenhouse experiment. Table S7. *P*-values from Tukey's HSD for tocochromanols and carotenoids from the 2019 field experiment and tocochromanols, chlorophyll *a*, and carotenoids from 2020 greenhouse experiment. Table S8. List of 125 a priori candidate genes. Dataset S1. Raw metabolomics data from each experiment. Methods S1. Materials and Methods.

Acknowledgements
We gratefully acknowledge Kan Wang and the Iowa State Plant Transformation Facility for generating the T0 transgenic plantlets. We also gratefully acknowledge Christine H. Diepenbrock for assisting in genotype selection for the physiological experiment. We thank current and past members of the Gore, DellaPenna, and Buell labs for their efforts in pollination, harvest, and sample preparation.

Authors' contributions
D.D.P. and M.A.G. conceptualized the project. C.R.B. and M.A.G. oversaw the informatics work. X.L, M.M-L., D.W., C.T.H., and N.K. performed the experiments. C.T.H. contributed new reagents/analytic tools. D.W., S.H, X.L., J.C.W., C.H.D., C.R.B., M.M-L. analyzed the data. S.H., D.D.P., and M.A.G. co-wrote the first draft of the manuscript, with input from all authors.

Funding
This research was supported by the National Science Foundation (IOS-1546657 to CRB, DDP, and MAG); the National Institute of Food and Agriculture; the USDA Hatch under accession numbers 1013641 and 1023660 (MAG); and Cornell University startup funds (MAG). This study was also made possible by the support of the American People provided to the Feed the Future Innovation Lab for Crop Improvement through the United States Agency for International Development (USAID). The contents are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government. Program activities are funded by the United States Agency for International Development (USAID) under Cooperative Agreement No. 7200AA-19LE-00005 (MAG).

Data availability
All raw RNA-seq data are available from the NCBI Sequence Read Archive under BioProject PRJNA643165. Supplementary datasets, figures, methods, and tables are available in the Supplementary Materials section on the journal website. All code is available on Github (https://github.com/GoreLab/POR_tocopherols).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 12 December 2024 Accepted: 17 February 2025
Published online: 13 March 2025

References

1. Albert E, Kim S, Magallanes-Lundback M, Bao Y, Deason N, Danilo B, Wu D, Li X, Wood JC, Bornowski N, Gore MA, Buell CR, DellaPenna D. Genome-wide association identifies a missing hydrolase for tocopherol synthesis in plants. Proc Natl Acad Sci. 2022;119(23):Article 23. https://doi.org/10.1073/pnas.2113488119.
2. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389–402. https://doi.org/10.1093/nar/25.17.3389.
3. Anders S, Pyl PT, Huber W. HTSeq—A Python framework to work with high-throughput sequencing data. Bioinformatics (Oxford, England). 2015;31(2):166–9. https://doi.org/10.1093/bioinformatics/btu638.
4. Collakova E, DellaPenna D. The role of homogentisate phytyltransferase and other tocopherol pathway enzymes in the regulation of tocopherol synthesis during abiotic stress. Plant Physiol. 2003;133(2):930–40. https://doi.org/10.1104/pp.103.026138.
5. Diepenbrock CH, Kandianis CB, Lipka AE, Magallanes-Lundback M, Vaillancourt B, Góngora-Castillo E, Wallace JG, Cepela J, Mesberg A, Bradbury PJ, Ilut DC, Mateos-Hernandez M, Hamilton J, Owens BF, Tiede T, Buckler ES, Rocheford T, Buell CR, Gore MA, DellaPenna D. Novel loci underlie natural variation in vitamin E levels in maize grain. Plant Cell. 2017;29(10):Article 10. https://doi.org/10.1105/tpc.17.00475.
6. Dong C-S, Zhang W-L, Wang Q, Li Y-S, Wang X, Zhang M, Liu L. Crystal structures of cyanobacterial light-dependent protochlorophyllide oxidoreductase. Proc Natl Acad Sci. 2020;117(15):8455–61. https://doi.org/10.1073/pnas.1920244117.
7. Downs GS, Bi Y-M, Colasanti J, Wu W, Chen X, Zhu T, Rothstein SJ, Lukens LN. A developmental transcriptional network for maize defines coexpression modules. Plant Physiol. 2013;161(4):1830–43. https://doi.org/10.1104/pp.112.213231.
8. Dror DK, Allen LH. Vitamin E deficiency in developing countries. Food Nutr Bull. 2011;32(2):Article 2. https://doi.org/10.1177/156482651103200206.
9. Federer WT, King F. Variations on split plot and split block experiment designs. Hoboken: John Wiley and Sons, Inc.; 2007.
10. Mendiburu F, Yaseen M. agricolae: Statistical Procedures for Agricultural Research. R package version 1.4.0. 2020. https://myaseen208.github.io/agricolae/https://cran.r-project.org/package=agricolae.
11. Fox J, Weisberg S. An R companion to applied regression (Third). Sage. 2019. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.
12. Frame BR, McMurray JM, Fonger TM, Main ML, Taylor KW, Torney FJ, Paz MM, Wang K. Improved Agrobacterium-mediated transformation of three maize inbred lines using MS salts. Plant Cell Rep. 2006;25(10):1024–34. https://doi.org/10.1007/s00299-006-0145-2.
13. Frame BR, Shou H, Chikwamba RK, Zhang Z, Xiang C, Fonger TM, Pegg SEK, Li B, Nettleton DS, Pei D, Wang K. Agrobacterium tumefaciens-mediated transformation of maize embryos using a standard binary vector system. Plant Physiol. 2002;129(1):13–22. https://doi.org/10.1104/pp.000653.

14. Gabruk M, Mysliwa-Kurdziel B. Light-dependent protochlorophyllide oxidoreductase: phylogeny, regulation, and catalytic properties. Biochemistry. 2015;54(34):5255–62. https://doi.org/10.1021/acs.biochem.5b00704.
15. Galani YJH, Orfila C, Gong YY. A review of micronutrient deficiencies and analysis of maize contribution to nutrient requirements of women and children in Eastern and Southern Africa. Crit Rev Food Sci Nutr. 2022;62(6):Article 6. https://doi.org/10.1080/10408398.2020.1844636.
16. Grams GW, Blessin CW, Inglett GE. Distribution of tocopherols within the corn kernel. J Am Oil Chem' Soc. 1970;47(9):337–9. https://doi.org/10.1007/BF02638997.
17. Griffiths WT. Reconstitution of chlorophyllide formation by isolated etioplast membranes. Biochem J. 1978;174(3):681–92. https://doi.org/10.1042/bj1740681.
18. Herbst J, Hey D, Grimm B. Chapter six - posttranslational control of tetrapyrrole biosynthesis: interacting proteins, chaperones, auxiliary factors. In: Grimm B, editor. Adv Bot Res, vol. 91. Academic Press; 2019. p. 163–94. https://doi.org/10.1016/bs.abr.2019.01.001.
19. Hey D, Rothbart M, Herbst J, Wang P, Müller J, Wittmann D, Gruhl K, Grimm B. LIL3, a light-harvesting complex protein, links terpenoid and tetrapyrrole biosynthesis in arabidopsis thaliana1[OPEN]. Plant Physiol. 2017;174(2):1037–50. https://doi.org/10.1104/pp.17.00505.
20. Hörtensteiner S, Hauenstein M, Kräutler B. Chapter Seven—Chlorophyll breakdown—Regulation, biochemistry and phyllobilins as its products. In B. Grimm (Ed.), Advances in botanical research. Academic Press. 2019;90:213–271. https://doi.org/10.1016/bs.abr.2019.03.004.
21. Horvath G, Wessjohann L, Bigirimana J, Monica H, Jansen M, Guisez Y, Caubergs R, Horemans N. Accumulation of tocopherols and tocotrienols during seed development of grape (Vitis vinifera L. cv. Albert Lavallée). Plant Physiol Biochem. 2006;44(11):724–31. https://doi.org/10.1016/j.plaphy.2006.10.010.
22. Huang J, Weinstein SJ, Yu K, Männistö S, Albanes D. Relationship between serum alpha-tocopherol and overall and cause-specific mortality. Circ Res. 2019;125(1):29–40. https://doi.org/10.1161/CIRCRESAHA.119.314944.
23. Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, Campbell MS, Stein JC, Wei X, Chin CS, Guill K, Regulski M, Kumari S, Olson A, Gent J, Schneider KL, Wolfgruber TK, May MR, Springer NM, et al. Improved maize reference genome with single-molecule technologies. Nature. 2017;546(7659):524–7. https://doi.org/10.1038/nature22971.
24. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15. https://doi.org/10.1038/s41587-019-0201-4.
25. Lalitha S. Primer Premier 5. Biotech Software & Internet Report. 2000;1(6):270–2. https://doi.org/10.1089/152791600459894.
26. Leth T, Sondergaard H. Biological activity of vitamin E compounds and natural materials by the resorption-gestation test, and chemical determination of the vitamin E activity in foods and feeds. J Nutr. 1977;107(12):2236–43. https://doi.org/10.1093/jn/107.12.2236.
27. Lin Y-P, Charng Y. Supraoptimal activity of CHLOROPHYLL DEPHYTYLASE1 results in an increase in tocopherol level in mature arabidopsis seeds. Plant Signal Behav. 2017;12(11): e1382797. https://doi.org/10.1080/15592324.2017.1382797.
28. Lipka AE, Gore MA, Magallanes-Lundback M, Mesberg A, Lin H, Tiede T, Chen C, Buell CR, Buckler ES, Rocheford T, DellaPenna D. Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. G3 Genes|Genomes|Genetics. 2013;3(8):1287–99. https://doi.org/10.1534/g3.113.006148.
29. Liu N, Du Y, Yan S, Chen W, Deng M, Xu S, Wang H, Zhan W, Huang W, Yin Y, Yang X, Zhao Q, Fernie AR, Yan J. The light and hypoxia induced gene ZmPORB1 determines tocopherol content in the maize kernel. Sci China Life Sci. 2024;67:435. https://doi.org/10.1007/s11427-023-2489-2.
30. Liu X, Hua X, Guo J, Qi D, Wang L, Liu Z, Jin Z, Chen S, Liu G. Enhanced tolerance to drought stress in transgenic tobacco plants overexpressing VTE1 for increased tocopherol production from Arabidopsis thaliana. Biotech Lett. 2008;30(7):1275–80. https://doi.org/10.1007/s10529-008-9672-y.
31. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods (San Diego, Calif). 2001;25(4):402–8. https://doi.org/10.1006/meth.2001.1262.
32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):Article 12. https://doi.org/10.1186/s13059-014-0550-8.
33. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetJournal. 2011;17(1):Article 1. https://doi.org/10.14806/ej.17.1.200.
34. Mène-Saffrané L, DellaPenna D. Biosynthesis, regulation and functions of tocochromanols in plants. Plant Physiol Biochem. 2010;48(5):301–9. https://doi.org/10.1016/j.plaphy.2009.11.004.
35. Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. Applied linear statistical models. 4th ed. Chicago, Illinois: Irwin. 1996.
36. Owens BF, Lipka AE, Magallanes-Lundback M, Tiede T, Diepenbrock CH, Kandianis CB, Kim E, Cepela J, Mateos-Hernandez M, Buell CR, Buckler ES, DellaPenna D, Gore MA, Rocheford T. A foundation for provitamin A biofortification of maize: genome-wide association and genomic prediction models of carotenoid levels. Genetics. 2014;198(4):Article 4. https://doi.org/10.1534/genetics.114.169979.
37. Pellaud S, Mène-Saffrané L. Metabolic origins and transport of vitamin E biosynthetic precursors. Front Plant Sci. 2017;8: 1959. https://doi.org/10.3389/fpls.2017.01959.
38. R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing.https://www.R-project.org/
39. Reinbothe C, Bakkouri ME, Buhr F, Muraki N, Nomata J, Kurisu G, Fujita Y, Reinbothe S. Chlorophyll biosynthesis: Spotlight on protochlorophyllide reduction. Trends Plant Sci. 2010;15(11):614–24. https://doi.org/10.1016/j.tplants.2010.07.002.
40. Romer J, Gutbrod K, Schuppener A, Melzer M, Müller-Schüssele SJ, Meyer AJ, Dörmann P. Tocopherol and phylloquinone biosynthesis in chloroplasts requires the phytol kinase VITAMIN E PATHWAY GENE5 (VTE5) and the farnesol kinase (FOLK). Plant Cell. 2024;36(4):1140–58. https://doi.org/10.1093/plcell/koad316.
41. Sameer H, Victor G, Katalin S, Henrik A. Elucidation of ligand binding and dimerization of NADPH:protochlorophyllide (Pchlide) oxidoreductase from pea (L.) by structural analysis and simulations. Proteins Struct Funct Bioinform. 2021;89(10):1300–14. https://doi.org/10.1002/prot.26151.
42. Sattler SE, Gilliland LU, Magallanes-Lundback M, Pollard M, DellaPenna D. Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. Plant Cell. 2004;16(6):1419–32. https://doi.org/10.1105/tpc.021360.
43. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc. 2008;3(6):1101–8. https://doi.org/10.1038/nprot.2008.73.
44. Schwingshackl L, Boeing H, Stelmach-Mardas M, Gottschald M, Dietrich S, Hoffmann G, Chaimani A. Dietary supplements and risk of cause-specific death, cardiovascular disease, and cancer: A systematic review and meta-analysis of primary prevention trials. Adv Nutr. 2017;8(1):27–39. https://doi.org/10.3945/an.116.013516.
45. Sen CK, Khanna S, Roy S. Tocotrienols: Vitamin E beyond tocopherols. Life Sci. 2006;78(18):2088–98. https://doi.org/10.1016/j.lfs.2005.12.001.
46. Steel RGD, Torrie JH, Dickey DA. Principles and procedures of statistics: a biometrical approach. 3rd ed. New York: McGraw Hill Book International Co.; 1997.
47. Su Q, Frick G, Armstrong G, Apel K. POR C of Arabidopsis thaliana: A third light- and NADPH-dependent protochlorophyllide oxidoreductase that is differentially regulated by light. Plant Mol Biol. 2001;47(6):805–13. https://doi.org/10.1023/A:1013699721301.
48. Tanaka R, Oster U, Kruse E, Rüdiger W, Grimm B. Reduced Activity of Geranylgeranyl Reductase Leads to Loss of Chlorophyll and Tocopherol and to Partially Geranylgeranylated Chlorophyll in Transgenic Tobacco Plants Expressing Antisense RNA for Geranylgeranyl Reductase1. Plant Physiol. 1999;120(3):695–704. https://doi.org/10.1104/pp.120.3.695.
49. Tanaka R, Rothbart M, Oka S, Takabayashi A, Takahashi K, Shibata M, Myouga F, Motohashi R, Shinozaki K, Grimm B, Tanaka A. LIL3, a light-harvesting-like protein, plays an essential role in chlorophyll and tocopherol biosynthesis. Proc Natl Acad Sci. 2010;107(38):16721–5. https://doi.org/10.1073/pnas.1004699107.
50. Tzin V, Galili G. New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. Mol Plant. 2010;3(6):956–72. https://doi.org/10.1093/mp/ssq048.
51. vom Dorp K, Hölzl G, Plohmann C, Eisenhut M, Abraham M, Weber APM, Hanson AD, Dörmann P. Remobilization of phytol from chlorophyll degradation Is essential for tocopherol synthesis and growth of Arabidopsis. Plant Cell. 2015;27(10):2846–59. https://doi.org/10.1105/tpc.15.00395.

52. Wan CY, Wilkins TA. A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (Gossypium hirsutum L.). Analyt Biochem. 1994;223(1):7–12. https://doi.org/10.1006/abio.1994.1538.
53. Wei, T., & Simko, V. (2021). R package "corrplot": Visualization of a correlation matrix. https://github.com/taiyun/corrplot.
54. Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org.
55. Wu D, Li X, Tanaka R, Wood JC, Tibbs-Cortes LE, Magallanes-Lundback M, Bornowski N, Hamilton JP, Vaillancourt B, Diepenbrock CH, Li X, Deason NT, Schoenbaum GR, Yu J, Buell CR, DellaPenna D, Gore MA. Combining GWAS and TWAS to identify candidate causal genes for tocochromanol levels in maize grain. Genetics. 2022;221(4): iyac091. https://doi.org/10.1093/genetics/iyac091.
56. Zhan W, Liu J, Pan Q, Wang H, Yan S, Li K, Deng M, Li W, Liu N, Kong Q, Fernie AR, Yan J. An allele of ZmPORB2 encoding a protochlorophyllide oxidoreductase promotes tocopherol accumulation in both leaves and kernels of maize. Plant J. 2019;100(1):Article 1. https://doi.org/10.1111/tpj.14432.
57. Zhang C, Cahoon RE, Hunter SC, Chen M, Han J, Cahoon EB. Genetic and biochemical basis for alternative routes of tocotrienol biosynthesis for enhanced vitamin E antioxidant production. Plant J. 2013;73(4):628–39. https://doi.org/10.1111/tpj.12067.
58. Zhang C, Zhang W, Ren G, Li D, Cahoon RE, Chen M, Zhou Y, Yu B, Cahoon EB. Chlorophyll synthase under epigenetic surveillance is critical for vitamin E synthesis, and altered expression affects tocopherol levels in arabidopsis. Plant Physiol. 2015;168(4):1503–11. https://doi.org/10.1104/pp.15.00594.
59. Zhang S, Heyes DJ, Feng L, Sun W, Johannissen LO, Liu H, Levy CW, Li X, Yang J, Yu X, Lin M, Hardman SJO, Hoeven R, Sakuma M, Hay S, Leys D, Rao Z, Zhou A, Cheng Q, Scrutton NS. Structural basis for enzymatic photocatalysis in chlorophyll biosynthesis. Nature. 2019;574(7780):722–5. https://doi.org/10.1038/s41586-019-1685-2.
60. Zhang W, Liu T, Ren G, Hörtensteiner S, Zhou Y, Cahoon EB, Zhang C. Chlorophyll degradation: the tocopherol biosynthesis-related phytol hydrolase in Arabidopsis seeds Is still missing. Plant Physiol. 2014;166(1):70–9. https://doi.org/10.1104/pp.114.243709.

Publisher's Note