

RESEARCH ARTICLE

# Carotenoid Content and Root Color of Cultivated Carrot: A Candidate-Gene Association Study Using an Original Broad Unstructured Population

Matthieu Jourdan<sup>1,2,3</sup>, Séverine Gagné<sup>1,2,3</sup>, Cécile Dubois-Laurent<sup>1,2,3</sup>, Mohamed Maghraoui<sup>1,2,3</sup>, Sébastien Huet<sup>1,2,3</sup>, Anita Suel<sup>1,2,3</sup>, Latifa Hamama<sup>1,2,3</sup>, Mathilde Briard<sup>1,2,3</sup>, Didier Peltier<sup>1,2,3</sup>, Emmanuel Geoffriau<sup>1,2,3\*</sup>

**1** Agrocampus Ouest, UMR1345 Institut de Recherche en Horticulture et Semences, Angers, France, **2** Université d'Angers, UMR1345 Institut de Recherche en Horticulture et Semences, Angers, France, **3** INRA, UMR1345 Institut de Recherche en Horticulture et Semences, Beaucouzé, France

\* [emmanuel.geoffriau@agrocampus-ouest.fr](mailto:emmanuel.geoffriau@agrocampus-ouest.fr)



**OPEN ACCESS**

**Citation:** Jourdan M, Gagné S, Dubois-Laurent C, Maghraoui M, Huet S, Suel A, et al. (2015) Carotenoid Content and Root Color of Cultivated Carrot: A Candidate-Gene Association Study Using an Original Broad Unstructured Population. PLoS ONE 10(1): e0116674. doi:10.1371/journal.pone.0116674

**Academic Editor:** Eric A. Johnson, University of Wisconsin, Food Research Institute, UNITED STATES

**Received:** July 23, 2014

**Accepted:** December 11, 2014

**Published:** January 23, 2015

**Copyright:** © 2015 Jourdan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was funded by the French ministry of Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Accumulated in large amounts in carrot, carotenoids are an important product quality attribute and therefore a major breeding trait. However, the knowledge of carotenoid accumulation genetic control in this root vegetable is still limited. In order to identify the genetic variants linked to this character, we performed an association mapping study with a candidate gene approach. We developed an original unstructured population with a broad genetic basis to avoid the pitfall of false positive detection due to population stratification. We genotyped 109 SNPs located in 17 candidate genes – mostly carotenoid biosynthesis genes – on 380 individuals, and tested the association with carotenoid contents and color components. Total carotenoids and  $\beta$ -carotene contents were significantly associated with genes zeaxanthin epoxidase (ZEP), phytoene desaturase (PDS) and carotenoid isomerase (CRTISO) while  $\alpha$ -carotene was associated with CRTISO and plastid terminal oxidase (PTOX) genes. Color components were associated most significantly with ZEP. Our results suggest the involvement of the couple PDS/PTOX and ZEP in carotenoid accumulation, as the result of the metabolic and catabolic activities respectively. This study brings new insights in the understanding of the carotenoid pathway in non-photosynthetic organs.

## Introduction

Carotenoid compounds play an essential role in human health, preventing disease thanks to their antioxidant capacity, but also as provitamin A precursors. As humans cannot synthesize carotenoids, they have to be provided by plant-based dietary [1]. Carrot is one of the most important vegetables in the world, and a critical source of carotenoid as a large amount is accumulated in root tissues [2]. Moreover genetic resources exhibit a large range of colors and

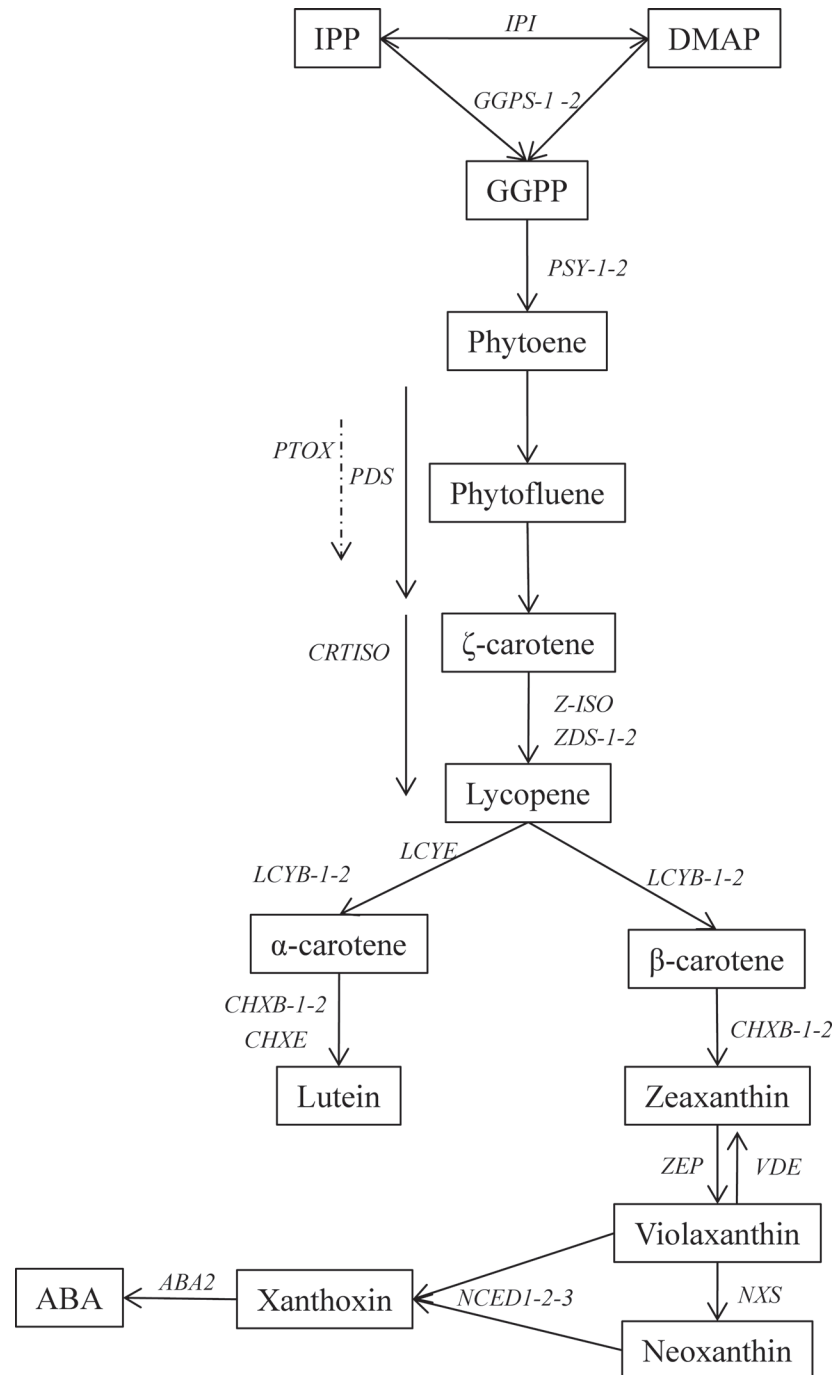
carotenoid content patterns [3], questioning the genetic control of carotenoid accumulation in carrot.

Carotenoid biosynthesis is today well established (Fig. 1) and genes encoding carotenoid enzymes have been characterized in many species [4–7]. Multiple steps in the pathway have been identified as controlling the carotenoid diversity and amount in various plant organs. Substrate availability—*isopentenyl diphosphate* and *dimethylallyl-diphosphate*—is generally considered as a limiting factor as well as the catabolic activity [4,8]. Accumulation of phytoene, controlled by the phytoene synthase and the phytoene desaturase, has emerged as a key regulatory step in the accumulation of carotenoids in various storage organs [9–13].

Many studies have shown that carotenoid biosynthetic genes are involved in the genetic control of carotenoid content (maize [14], tomato [15], wheat [12,16], pepper [17]). Depending on the species, all carotenoid biosynthetic genes may be involved in the genetic basis of carotenoid content and are therefore meaningful candidate genes [4]. In some species, engineering the pathway using biosynthetic genes is now possible for crop enhancement of the carotenoid content. Golden Rice is such an example of metabolic pathway engineering for quality enhancement [18].

However, little is known about the genetic control of carotenoid accumulation in carrot. Heritability of carotenoid content in carrot roots has been estimated by [19] and ranges from 28% to 98% depending on the compound and the investigated genetic background. Two major loci *Y* and *Y2* governing the orange intensity of xylem/phloem were identified [20]. The *Y* locus may block the synthesis of carotene and xanthophyll, whereas the *Y2* locus determines the carotene accumulation but not the xanthophyll one [21, 22]. A path analysis showed that phytoene accumulation may be one key step limiting carotenoid accumulation in white roots [9]. This was confirmed by [11], who turned a white rooted carrot in orange by overexpressing a phytoene synthase gene. Recently, a polymorphism of carotene hydroxylase *CYP97A3* controlling the  $\alpha$ -carotene content was identified [23] and the authors suggested a negative feedback regulation on *PSY* determining the carotenoid flux. Only two studies [21,24] have studied the genetic determinism of carotenoid content in carrot roots by linkage mapping, using a cross between an orange cultivated carrot and a white wild one. Moreover, almost all biosynthetic genes have been sequenced and mapped in carrot [25]. Two major QTLs governing carotenoid accumulation were localized, with some of carotenoid biosynthetic genes – zeaxanthin epoxidase, carotene hydroxylase and carotenoid dioxygenase families – mapped in the confidence interval or near these two QTLs.

As QTLs might be population-specific, association mapping has emerged in the last decade as an alternative to linkage analysis to dissect the basis of quantitative traits in plants. Such studies address the relationship between marker-based polymorphism and phenotypic variation in a diversified population. Using a diversified population may increase the resolution of such a study by using all ancestral recombination events [26]. One major interest of such a population is also the opportunity to study many alleles compared to a bi-parental cross study [27]. Association mapping targeting candidate genes has proven successful in many instances [28–31] and might bring new insights for carotenoid content as the genetic pathway has already been dissected through forward and reverse genetics in many organisms. However, one pitfall in association mapping is the lack of power when performed in structured panels. Structure can lead to an increase of false discovery rate. Indeed, false positives can be detected when phenotypic traits are correlated with underlying population structure at non causal loci [32]. *Daucus carota* L. genetic resources are known to be structured into two distinct genetic groups [33–35] according to their geographical origin. Moreover carotenoid content pattern is closely linked to these genetic groups: cultivars with high lycopene content belong mostly to one genetic group. In this case, association mapping can typically lead to false positive detection.



**Figure 1. Carotenoid biosynthetic pathway in plants and gene copy number in carrot.** IPP: isopentenyl diphosphate, DMAP: Dimethylallyl-diphosphate, IPI: Isopentenyl pyrophosphate isomerase, GGPS: Geranylgeranyl pyrophosphate synthase, PSY: Phytoene synthase, PDS: Phytoene desaturase, PTOX: Plastid terminal oxidase, CRTISO: Carotene isomerase, ZDS: ζ-carotene isomerase, LCYE: ε-lycopene cyclase, LCYB: β-lycopene cyclase, CHXB: β-carotene hydroxylase, CHXE: ε-carotene hydroxylase, ZEP: zeaxanthin epoxydase, VDE: violaxanthin de-epoxydase, NXS: Neoxanthin synthase, NCED: nine-cis-epoxycarotenoid dioxygenase, ABA: abscisic acid. Adapted from [25] and [58].

doi:10.1371/journal.pone.0116674.g001

Population stratification can be estimated with different ways, and added as a covariate in association models, limiting the detection of false positive. The first proposed model [36] – called Q model – estimates the population stratification with the Bayesian model, now implemented in Structure software [37,38]. Then a principal component analysis has been proposed to correct for the structure [39]. The use of both Kinship matrix (K) and population structure in a unified mixed model approach to account for relatedness between individuals has been proposed by [27].

In order to overcome the structure bias, we have created a specific unstructured population with a broad genetic basis to perform an association mapping study for carrot root carotenoid content.

The aim of this study was to investigate the implication of biosynthetic genes in carrot root carotenoid content and related color traits using a broad unstructured population in a candidate-gene association approach. This work will offer new insights in the global understanding of the carotenoid biosynthetic pathway in carrot. This will allow us to identify favorable alleles with associated markers usable in marker-assisted selection (MAS) for product quality enhancement.

## Materials and Methods

### Plant material

The discovery population consisted of 380 individuals from the third generation of intercrossing of an initial panel of 67 cultivars, each represented by 6 individuals. This panel represents a large diversity of cultivated carrot: three white (Europe and Middle-east), eight yellow (Europe, Central Asia, Asia), two red (Asia), 45 orange (Europe, South & North America, Australia, Madagascar, central Asia, Asia) and eight purple ones (Europe & Middle East). Pollinators were introduced at maximum blooming to avoid genetic drift and seeds were harvested from each plant in a balanced way as suggested by [40]. At each generation, 150 seeds were randomly chosen and sown for the next one. The third intercrossing generation was sown in the field in Agrocampus-Ouest (Angers, France) and grown following standard practices. Roots were harvested at 97 days after sowing. At harvesting, 380 individuals were randomly chosen.

### Phenotypic data

#### Color evaluation by spectrophotometry

Since color is a quality attribute and may be considered as an indirect measure of carotenoid content [41–42], color was evaluated with a CM2600d Minolta (Japan) spectrophotometer equipped with a 5 mm measuring area. The illuminant used was D65 and calibration was done with a white standard and specular component of light was excluded. Two measures were done for epidermis and secondary phloem and one for secondary xylem.

CIELAB color space coordinates ( $L^*$ ,  $a^*$ ,  $b^*$ ,  $C^*$  and  $h$ ) were recorded. Data represent the mean of measures per tissue type.

Then roots extremities were cut off and roots were immediately ground and frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . Both carotenoids and DNA were extracted from the same root sample. Dry matter was determined after drying approximately 10 grams sample at  $55^{\circ}\text{C}$  for 7 days.

#### Carotenoid quantification by HPLC

The procedure was adapted from [43]. Extraction was done on approximately 500 mg of crushed frozen material to which 50  $\mu\text{L}$  of  $\beta$ -apo-8'-carotenal at 0.1 g/L was first added as an

internal standard. Samples were mixed with 7 mL MgCO<sub>3</sub> 0.57%, 3,5-di-*tert*-butyl-4-hydroxytoluene (BHT) 0.1% in methanol, then vortexed, and mixed with 7 mL of 0.1% BHT-containing chloroform. After 15 min incubation in darkness, 7 mL of ultrapure water were added, and samples were centrifuged at 236 g for 10 min. One milliliter from the lower layer was concentrated under vacuum evaporation, and the dry extract was dissolved in 200  $\mu$ L of acetonitrile/dichloromethane (50:50, v/v) containing 0.1% BHT. Samples were kept at 4°C and protected from direct light during the whole procedure. Extraction was carried out in duplicates.

The analyses for carotenoid quantification were done on a Shimadzu (Shimadzu Corporation, Kyoto, Japan) HPLC equipped with a ternary pumps (LC-10AT VP), a thermostated autosampler (SIL-10AD VP), a photodiode array detector (SPD-M10A VP), a controller (SCL-10A VP), an on-line degasser (Degasys DG-1310), and a temperature controller (Crococil). Data acquisition and processing were done using the LC workstation Class-VP (Shimadzu). The procedure was adapted from [44]. Carotenoids were separated along an YMC C30 (YMC, Japan) column (150  $\times$  4.6 mm; 3  $\mu$ m) kept at 22°C. The mobile phases were MeOH/ACN/H<sub>2</sub>O 84/14/2 (v/v/v) as eluent A, and methylene chloride as eluent B. The elution program had the following proportions of solvent A: 0–5 min, 95–70%; 5–25 min, 70–45%; 25–30 min, 45–10%; 30–35 min, 10–95%; 35–42 min, 95%. The flow rate was 0.9 mL/min. The injection volume of filtered sample (0.45  $\mu$ m PTFE membrane filter) was 20  $\mu$ L. The detection was monitored from 200 nm to 800 nm.

Carotenoid compounds were identified according to their elution order and UV-visible spectrum in comparison with their authentic standards (analysed individually and in combination in the conditions used for samples), and with data from the literature [45] when unavailable. Quantification was done at 296 nm (phytoene), 348 nm (phytofluene), 450 nm (lutein,  $\beta$ -apo-8'-carotenal,  $\alpha$ -carotene,  $\beta$ -carotene), and 472 nm (lycopene), based on internal calibration using  $\beta$ -apo-8'-carotenal,  $\beta$ -carotene calibration curve and extraction yield. Data represent the mean of two assays per individual and are expressed as  $\beta$ -carotene equivalents in mg per 100gr of dry mater (DM).

## Genotypic data

### SNP discovery, haplotype-tagging SNP selection and genotyping

Assuming results from previous studies, we choose carotenoid biosynthetic genes as potential good candidate genes to explain the observed variation in root carotenoid content.

Polymorphic sequences for *UPI*, *PDS*, *LCYE*, *LCYBI*, *CHXE*, *ZEP* and *CRTISO* genes were obtained from previous works [33,35].

In order to identify polymorphism in other candidate genes, 48 lines representing a large diversity were sequenced for *GGPS2*, *ZDS1*, *PTOX*, *PSY1*, *PSY2*, *NCED1-2-3*, *LCYB2* and *ABA2* gene fragments. Primers were designed based on publicly available databases (S1 Table). PCR reactions, cycling conditions and sequencing protocol were identical as in [35]. All sequences were aligned by using Geneious software and haplotypes were inferred with DnaSP.

The sequence for a marker associated to the *Y2* locus responsible for carotenoid accumulation was obtained from [46].

### SNP Genotyping

DNA was isolated and purified with a modified CTAB protocol [47] and DNA concentration was adjusted to 15 ng/ $\mu$ L. SNP genotyping was carried out by KASPar Assay (KBioscience – LGC Genomics). This technology is based on a property competitive allele-specific PCR

system. The allele detection is based on a FRET quencher cassette which allows bi-allelic discrimination of known SNPs and InDels.

A total of 470 SNPs were found over all the 12,351 bp sequences from 17 genes. We identified 169 haplotypes over the 17 genes. Haplotype-tagging SNPs (HtSNPs) were chosen to maximize the number of haplotypes. We were able to design primers for 109 SNPs out of the 120 predicted (context sequences are provided in [S2 Table](#)). Finally, 93 SNPs were used after removing SNPs with Minor Allele Frequency (MAF) lower than 5% and missing data higher than 20%.

For each gene, haplotypes were reconstructed using PHASE implemented in DNASP [48]. LD within and between candidate genes was assessed with  $r^2$  generated by TASSEL [49].

### SSR Genotyping

For SSR markers, 15 primers used in [35] were chosen regarding to their genome coverage and reproducibility [50]. PCR reactions, capillary electrophoresis and fragment sizing were identical as in [35].

### Population structure and relatedness

Both SSR and SNP datasets were used to investigate population stratification and relatedness between individuals as they can lead to false positive detection during association analysis.

Population stratification was first investigated on SSR dataset with the Bayesian model-based STRUCTURE software [37] which is used to infer distinct populations and to assign individuals to the identified populations. The model allowed admixture and allele correlated frequencies and was run with a burnin period of  $10^5$  and a run length of  $10^6$  iterations. Ten independent runs were performed for each putative cluster number (K). The range of possible K tested was from one to ten. Evanno's method [51] was used to estimate the most probable K number.

Population stratification was also studied by PCA analysis by using SNP dataset on TASSEL software [49].

In order to study the relatedness between individuals, two Kinship matrixes were calculated by using TASSEL software based on SSR (K-SSR) markers and SNP (K-SNP) markers respectively.

### Association tests

Association analysis was performed by using TASSEL software [49]. Single polymorphism with MAF less than 5% and/or more than 20% of missing data were removed from the analysis. Marker-Trait associations were calculated using six models to evaluate the effects of population stratification and kinship: first a naive model without correction, then models correcting for population stratification estimated by STRUCTURE (Q) and by PCA (P), and finally mixed linear models taking account for relatedness between individuals estimated with SNP (K-SNP) and SSR (K-SSR) and the population structure (Q+K and P+K).

Due to multiple testing, SNPs were declared as significantly associated with a threshold of  $5.3 \times 10^{-4}$  *p value* as corrected with the standard Bonferonni procedure. The amount of variation explained by a SNP ( $r^2$ ) was calculated for each significant association using a simple general linear model. Haplotypes were declared significantly associated with an arbitrary threshold of 0.05 *p value* due to the small number of tested genes.

The model pertinence was evaluated by plotting *p value* in a cumulative way as described by [52]. A uniform distribution of *p values* indicates an ideal model.

**Table 1. Carotenoid content variation in the unstructured population.**

	phytofluene	phytoene	lutein	$\alpha$ -carotene	$\beta$ -carotene	Total
Mean	1.436	1.134	3.624	5.900	16.388	28.831
SD	1.822	1.503	2.428	9.739	20.165	33.228
Min	0.000	0.000	0.949	0.014	0.083	1.253
Max	11.440	9.706	16.615	88.076	136.531	254.248

Traits are described by mean, standard deviation (SD), minimum (min) and maximum (max). Carotenoid contents are expressed in mg/100g of dry material. n = 380 individuals.

doi:10.1371/journal.pone.0116674.t001

## Results

### Phenotypic variation for carotenoid and color

A first visual observation of all 380 individuals showed no white or red roots. The whole population was composed with a gradient of yellow to orange roots. Due to the very low number of individuals exhibiting lycopene and limited content variability for this compound, association tests for lycopene content were not performed.

On average,  $\beta$ -carotene represented almost half of total carotenoid content and  $\alpha$ -carotene represented about the third of  $\beta$ -carotene content (Table 1). Lutein content was relatively low as well as precursor compounds like phytoene and phytofluene. The unstructured population exhibited a large variation for all carotenoid compounds as shown by the high standard deviation.

Color components were measured for epidermis, secondary phloem and secondary xylem. Phenotypic variations were quite similar between tissues (Table 2).

Table 3 shows the correlation between carotenoid compounds. All carotenoid compounds except lutein were highly correlated with  $\beta$ -carotene with a highly significant  $r^2$  between 0.78 and 0.99. On the contrary, lutein was not correlated with any other compound ( $r^2$  between 0.22 and 0.49). As secondary phloem represents the largest part of the roots, correlations between color components of this tissue and carotenoid content were investigated (Table 3). Except for lutein, all carotenoid compounds were mainly correlated with  $a^*$  and  $h$  ( $r^2$  between 0.70 and 0.77 or 0.6 and 0.69 respectively). Lutein was not correlated with any other trait.

### Population Structure

All 15 SSR markers were polymorphic and 132 alleles were identified with a mean of 8.8 alleles per locus.

Population stratification was first investigated with STRUCTURE based on SSR markers. LnP (K) plot (Fig. 2A) did not reach a plateau as expected in the presence of structure in the

**Table 2. Variation for color components for root epidermis, secondary phloem and secondary xylem.**

Traits	Epidermis		Secondary Phloem		Secondary Xylem	
	Mean	SD	Mean	SD	Mean	SD
L*	57.46	6.92	60.23	6.73	62.00	7.913
a*	15.85	9.65	14.61	9.83	11.98	9.28
b*	36.06	7.53	39.57	6.76	38.48	7.74
C*	40.62	7.35	43.38	7.32	41.24	8.32
h	65.90	15.33	70.24	13.09	73.19	12.48

Traits are described by mean and standard deviation (SD). n = 380 individuals.

doi:10.1371/journal.pone.0116674.t002

**Table 3. Pearson correlation between root content of carotenoid compounds and color components of the secondary phloem.**

	phytofluene	phytoene	α-carotene	β-carotene	lutein	total	L*	a*	b*	C*
phytoene	0.99***									
α-carotene	0.80***	0.78***								
β-carotene	0.90***	0.87***	0.88***							
lutein	0.49***	0.46***	0.22***	0.41***						
total	0.92***	0.89***	0.93***	0.99***	0.44***					
L*	0.45***	0.43***	0.39***	0.46***	0.06	0.44***				
a*	0.72***	0.70***	0.70***	0.77***	0.21***	0.77***	0.58***			
b*	0.07	0.06	0.06	0.08***	0.25***	0.09	0.59***	0.09		
C*	0.41***	0.4***	0.42***	0.46***	0.28***	0.47***	0.19***	0.57***	0.86***	
h	0.66***	0.64***	0.60***	0.69***	0.16**	0.67***	0.8***	0.90***	0.27***	0.23***

(Significance: \*\*\* < 0.0001; \*\* < 0.001)

doi:10.1371/journal.pone.0116674.t003

sample. The number of genetic groups was also investigated with the Evanno’s method [51]. As shown on Fig. 2B, the most probable K was 2, 3 or 5. However, as shown in Fig. 2 (C, D, E) all individuals were admixed and none of them was clearly assigned to one group. At least the proportion of samples assigned to each group was roughly symmetric (~ 1/K). All these elements showed an absence of structure in the population. This conclusion is reinforced by the principal component analysis performed with SNP data (Fig. 3), in which no group was clearly defined.

### Linkage Disequilibrium

No intergenic LD between genes was observed. All  $r^2$  between genes were less than 0.2 between SNPs located in different genes. But high or low LD was observed between SNPs within genes depending on the considered gene. *ZDS1*, *LCYB1*, *IPI*, *LCYE* and *ZEP* genes showed a relatively high LD between SNPs as observed by [33]. On the contrary, other genes such as *NCED* family or *PTOX* exhibited no intragenic LD (S1 Fig.).

### Association analysis

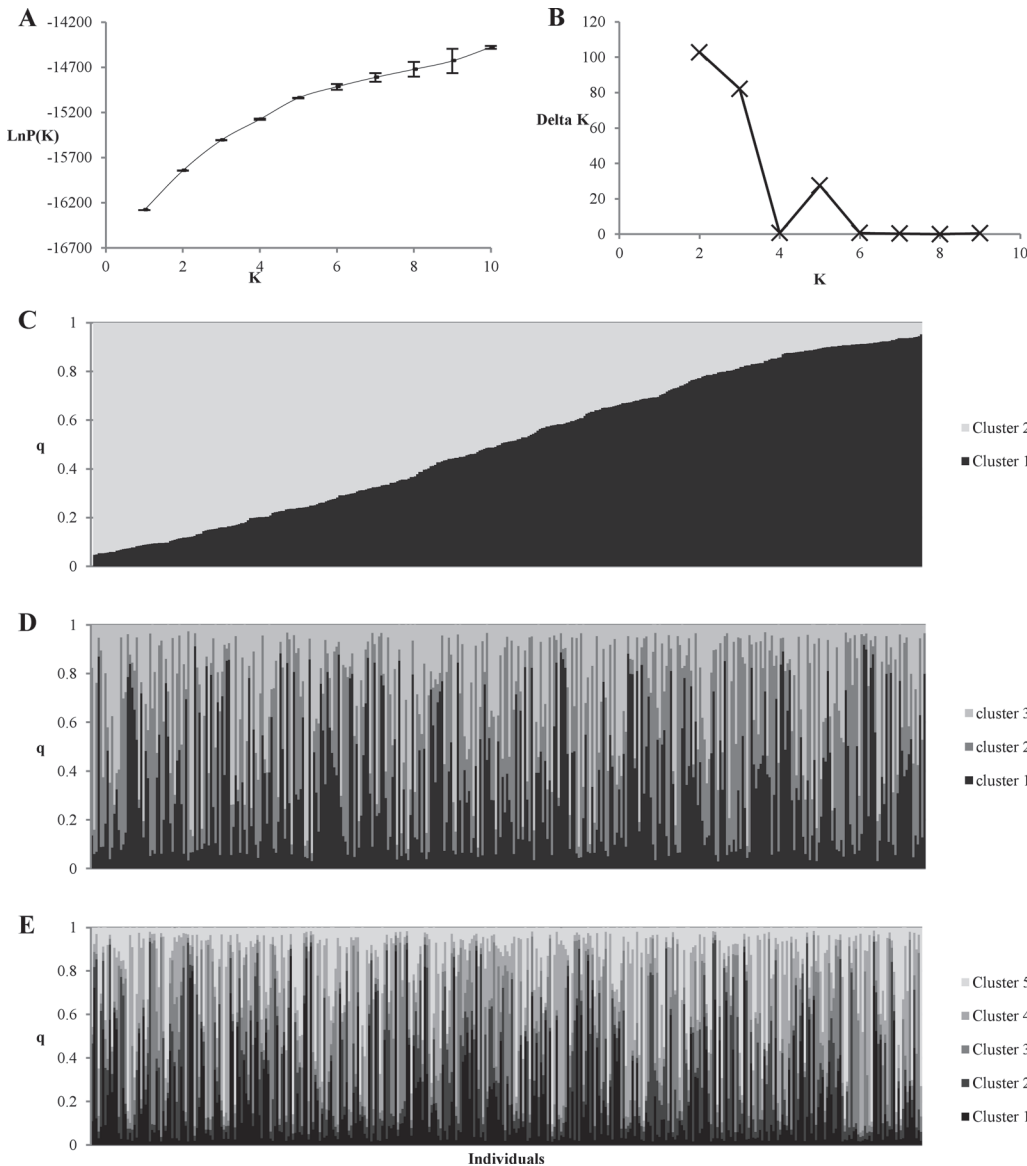
#### Model choice

As Structure did not converge and PCA did not show any structure in our population, models with Q and PCA covariables are not presented. To assess the goodness of each model (naive, K-SNP or K-SSR), cumulative *p* value plots are presented in Fig. 4 for carotenoid traits and SNPs based associations. For all traits, the naive model did not perform well and an excess of low *p* values was found. At least, the K-SSR model performed better than the naive one for SNP-trait associations as well as for haplotype-trait associations. The K-SNP model showed a uniform distribution of *p* values and therefore minimized the chance of spurious association. The presented association results are thus all based on the K-SNP model. Results were similar for color components traits: the K-SNP model performed better than any other models.

#### Association results based on SNPs

With the K-SNP model, 93 SNPs were tested against 21 traits. Among the 1953 marker-trait pairs, 23 significant associations were found with a Bonferroni corrected threshold of  $5.3 \times 10^{-5}$  (Fig. 5). Two SNPs (*ZEP*-117 and *ZEP*-361) in the zeaxanthin epoxydase gene were associated with total carotenoids ( $R^2_{ZEP-117} = 0.21$ ), β-carotene ( $R^2_{ZEP-117} = 0.22$ ), phytoene ( $R^2_{ZEP-117} = 0.22$ ) and phytofluene ( $R^2_{ZEP-117} = 0.23$ ) content. These results were similar when



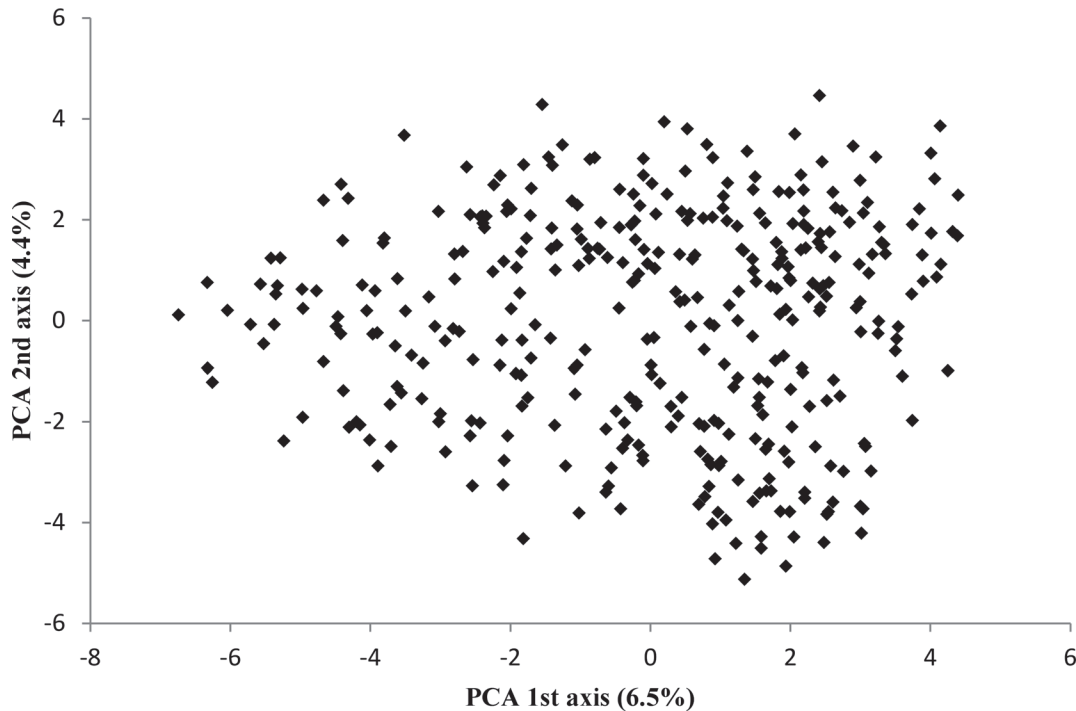


**Figure 2. Population stratification according to STRUCTURE based on 15 SSR markers.** Plots of (A) Delta K and (B) the log likelihood. Each individual is represented by a single vertical box broken in colored segments according to the number of assumed populations: 2 (C), 3 (D), 5 (E). Vertical box length is proportional to each cluster assignment probability (y-axis).

doi:10.1371/journal.pone.0116674.g002

carotenoid traits were expressed on a fresh matter basis (data not shown). These two SNPs were also associated with color components a\* and h for all three tested tissues (Table 4). These two SNPs were in high LD ( $r^2 = 1$ ) and therefore redundant. These SNPs were located in a non-coding region. Fig. 6 shows the distribution of carotenoid content for each allele of the ZEP-117 SNP (Results were similar for the ZEP-361 SNP, data not shown). For all associated compounds, C:C, C:T and T:T genotype means were significantly different from each other ( $p < 0.05$ , Kruskal-Wallis test). This reveals a typical dominant action for this locus.

One polymorphism in the carotenoid isomerase gene (*CRTISO*) was associated with total carotenoids,  $\beta$ -carotene and  $\alpha$ -carotene. At least one SNP in the plastid terminal oxidase (*PTOX*) – a cofactor of the phytoene desaturase – was associated with  $\alpha$ -carotene (Fig. 5). No association was detected between the Y2 related marker and carotenoid content.



**Figure 3. Principal component analysis of population based on SNP markers.** Numbers in parentheses refer to the proportion of explained variance.

doi:10.1371/journal.pone.0116674.g003

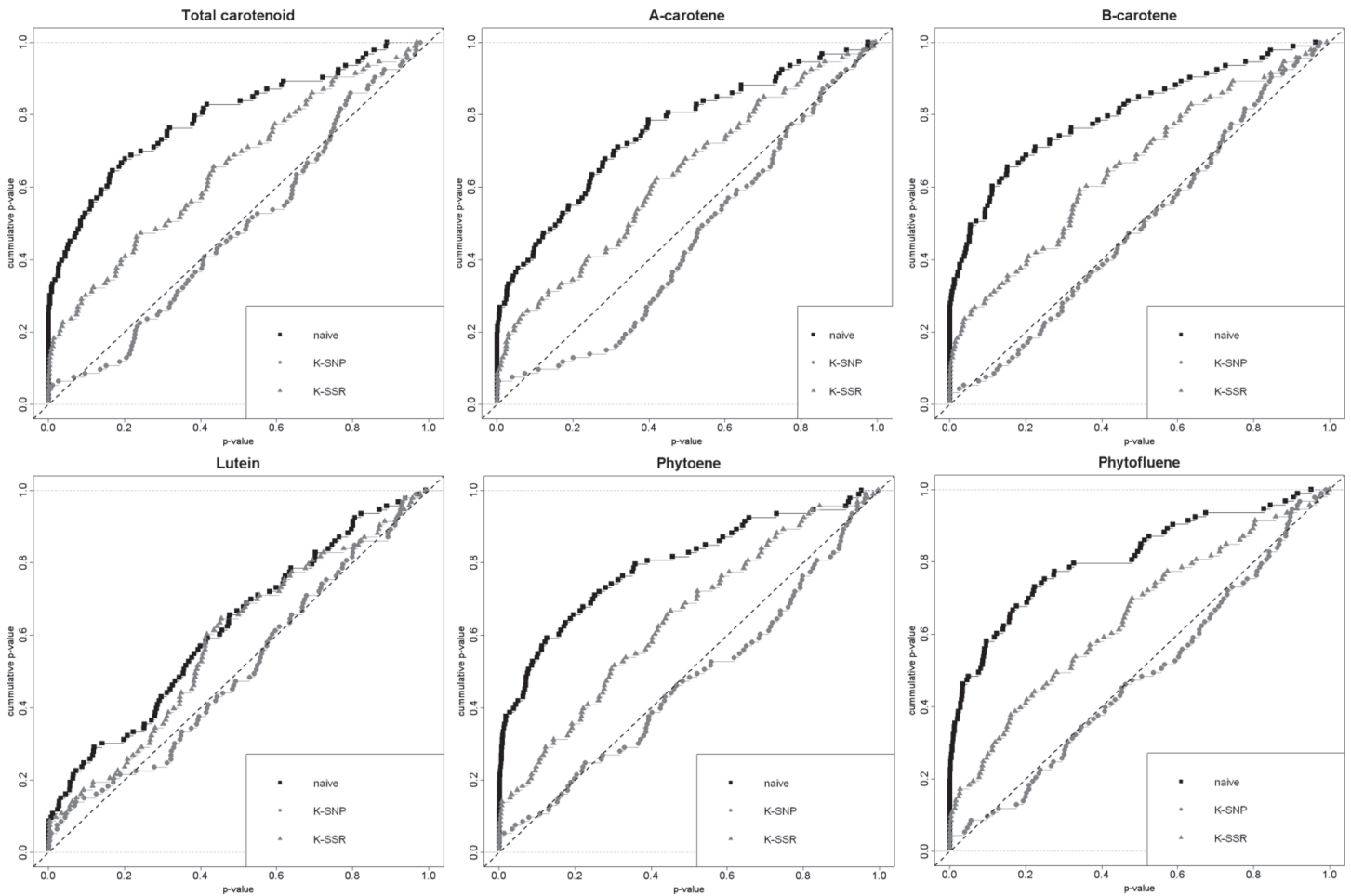
### Association results based on haplotypes

When testing haplotypes against carotenoid content, we detected 32 associations with a *p* value lower than 0.05 (Table 5). Among these 32 significant associations, 27 were associated with three genes. Total carotenoids,  $\beta$ -carotene, phytoene and phytofluene content were associated with *ZEP* gene. This gene was also associated with color components for the three investigated tissues. Phytoene desaturase gene was associated with lutein, phytoene, phytofluene and total carotenoid content and with the color saturation  $C^*$  of inner root. The plastid terminal oxidase gene (*PTOX*) was associated with  $\alpha$ -carotene, phytoene, phytofluene and total carotenoid content. It was also associated with the color components  $b^*$  for both epidermis and secondary phloem and  $C^*$  for epidermis. At last *PSY2*, *ZDS1* and *NCED1* were associated with color components and *PSY1* with lutein content.

## Discussion

### Unstructured population and consequences for association study

Carrot genetic resources often exhibit a strong stratification [33–35] which can lead to false positive detection when association mapping is performed. Here we conducted a candidate gene association study on root carotenoid content by using an unstructured population. Both SNPs and SSRs confirmed the absence of population stratification in this population. The same set of SSR markers were used in [35] and revealed a strong population stratification on a diversified panel of carrot accessions. So neither Q matrix nor PCA component were used in association mapping model. But the naive model showed an excess of low *p* values and was therefore not fitting. As all individuals descended from the same parents, relatedness between individuals was relatively high and association model had to be corrected for relatedness. A kinship matrix was still needed.



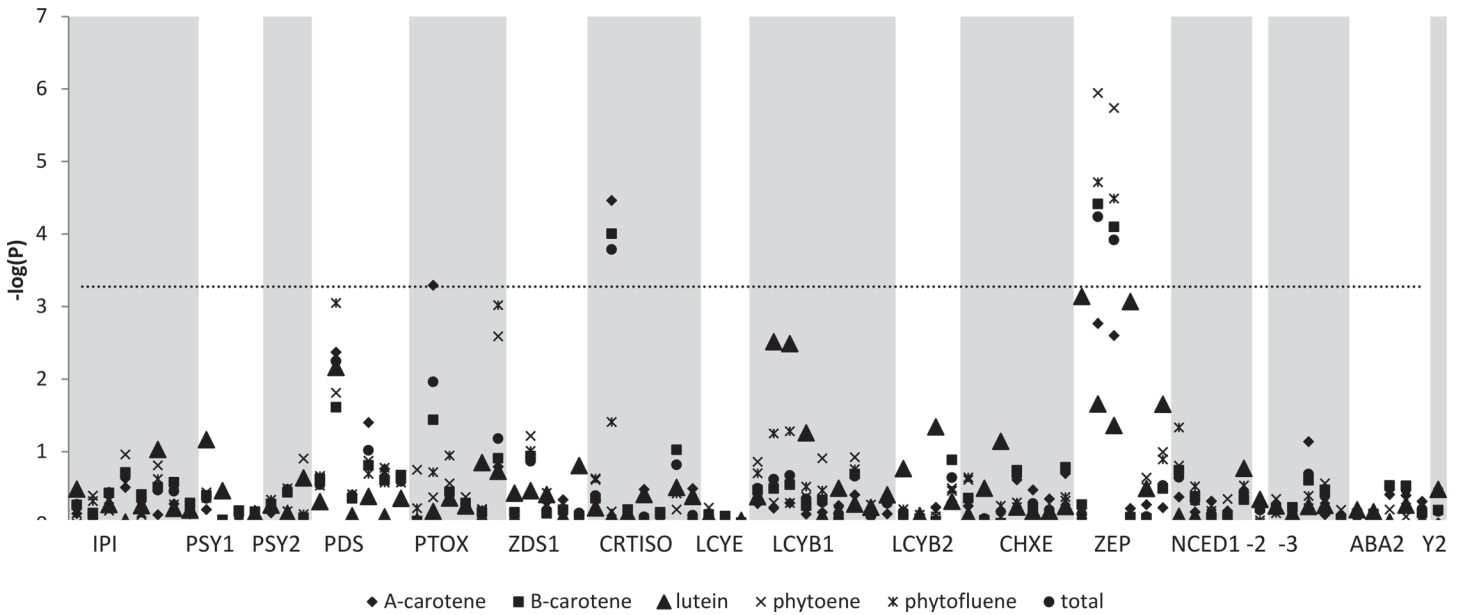
**Figure 4. Comparison of various association mapping models for carotenoid content.** Naive (black line) = Generalized Linear Model without correction. K-SSR (blue line) = Mixed Linear Model with Kinship evaluated from SSR data. K-SNP (red line) = Mixed Linear Model with Kinship evaluated from SNP data.

doi:10.1371/journal.pone.0116674.g004

Auzanneau et al. [53] already demonstrated in ray grass the interest of using a panmictic population in association mapping studies to overcome the population stratification bias. The choice of parents appeared to be essential. In our original set of cultivars set, red rooted carrot was under represented and we were not able to observe red roots after three generations. As population stratification was disrupted after three intercross generations, the choice of the parents should be based more on the phenotypical variation representativity than on original stratification. In order to reach a large variation for the targeted traits after several intercross generations, the first generation must exhibit a large range of phenotypic variation but also an equilibrated representation of phenotypic patterns. Indeed, as mentioned by [53], detected associations may be different depending on population parents.

### Major role of the catabolic gene zeaxanthin epoxydase

Zeaxanthin epoxydase gene polymorphism was associated with total carotenoid,  $\beta$ -carotene, phytoene and phytofluene content. Moreover, associations with color components  $a^*$  and  $h$  were also detected. As the two polymorphisms associated with traits were in a non-coding region, we were unable to detect the causal polymorphism of the phenotypic variation. But as



**Figure 5. Manhattan plots of the K-SNP model for carotenoid content.** Grey horizontal line indicates the significance threshold. Genes are displayed in the pathway order.

doi:10.1371/journal.pone.0116674.g005

this gene exhibited a moderate LD as shown in [S1 Fig.](#) and in [\[33\]](#), the causal polymorphism may be located somewhere else in the LD block.

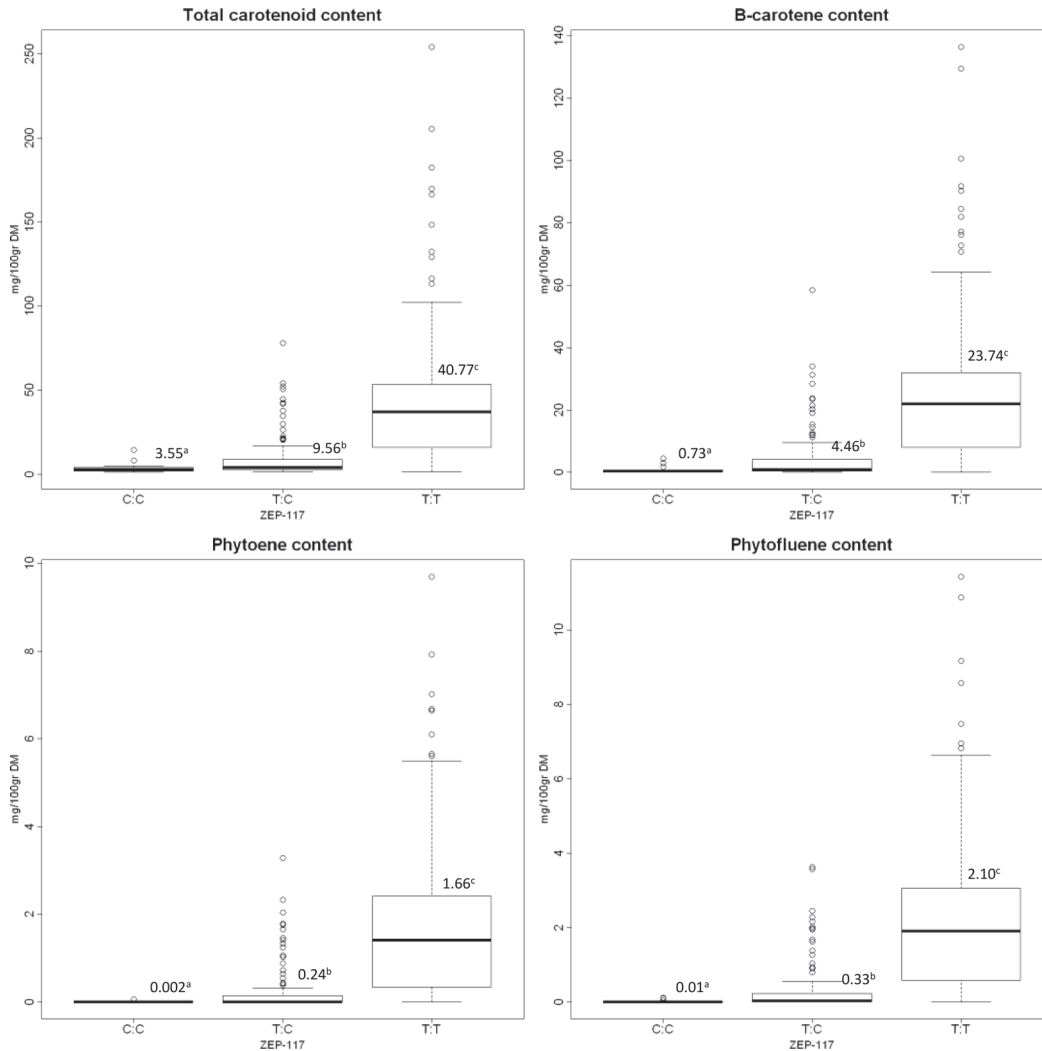
*ZEP* is one of the major steps in the carotenoid pathway. Nevertheless, carotenoid accumulation in various organs is known to be the result of biosynthesis, degradation and storage [\[4,8\]](#). An impaired function of the zeaxanthin epoxydase protein may result in an accumulation of  $\beta$ -carotene. As  $\beta$ -carotene, phytoene and phytofluene were highly and positively correlated, a high level of  $\beta$ -carotene was often associated with a high level of precursors phytoene and phytofluene. High level of phytoene and phytofluene may also be the result of a large  $\beta$ -carotene accumulation due to a reduced degradation. It also seems that the zeaxanthin epoxydase gene may drive the biosynthesis pathway towards the  $\beta$ -branch.

In a previous study, [\[21\]](#) performed a QTL detection in a biparental cross between a wild and an orange cultivated carrot. One major QTL was detected for  $\alpha$ - and  $\beta$ -carotene, phytoene,

**Table 4. Significant SNP-trait associations according to the K-SNP model for color components.**

SNP	Tissue	Trait	<i>p</i> value
ZEP-117	Epidermis	a*	2.70E-06
	Epidermis	h	1.41E-05
	Secondary Phloem	a*	1.79E-06
	Secondary Phloem	h	4.00E-06
	Secondary Xylem	a*	3.32E-06
	Secondary Xylem	h	3.64E-05
ZEP-361	Epidermis	a*	1.46E-05
	Secondary Phloem	a*	8.11E-06
	Secondary Phloem	h	2.88E-05
	Secondary Xylem	a*	5.87E-06
	Secondary Xylem	h	3.62E-05

doi:10.1371/journal.pone.0116674.t004



**Figure 6. Boxplots of the ZEP-117 alleles for total carotenoids, β-carotene, phytoene and phytofluene content.** Mean and group according to Kruskal-Wallis test are given for each boxplot.

doi:10.1371/journal.pone.0116674.g006

zeta-carotene, phytoene and total carotenoid content but not for lutein content. As the marker Y for the Y2 locus related to carotenoid accumulation was mapped in the confidence interval of this QTL and the STS (Sequence-Tagged Site) marker ZEP was mapped outside of the confidence interval, the authors concluded that the role of ZEP was not obvious to explain the Y2 locus. According to our results, we suggest that ZEP may one probable candidate gene underlying the Y2 locus.

Previous results on potato tubers also showed an association between ZEP and flesh color [54]. A specific allele only present in orange colored flesh genotypes was identified. This allele showed a reduced level of expression probably due to a large retrotransposon in the first intron. Similar observations were done in maize. Carotenoid accumulation in kernel was inversely associated with ZEP transcript levels [55]. However, no significant variation of ZEP transcript abundance between cultivars was found during carrot root development [43]. This suggests that a putative transcription difference would be at the allele level, which needs to be confirmed by a further study on ZEP alleles.

In tomato, the mutation *high pigment 3* (*hp3*) occurred in the ZEP gene, leading to a 30% increase of carotenoid accumulation in the mature fruit [56]. These results, with related effects

**Table 5. Significant haplotype-trait associations according to the K-SNP model for carotenoid content and color components.**

Gene	Tissue	Trait	<i>p</i> value
<i>PSY1</i>	whole root	lutein	0.019
<i>PSY2</i>	Secondary Phloem	L*	0.030
		b*	0.037
<i>ZDS1</i>	Epidermis	h	0.027
<i>PDS</i>	whole root	lutein	0
		phytoene	0.016
		phytofluene	0.014
	Secondary Phloem	C*	0.002
	Secondary Xylem	C*	0.015
<i>PTOX</i>	whole root	α-carotene	0.002
		phytoene	0.008
		phytofluene	0.017
		total	0.016
	Epidermis	b*	0.014
		C*	0.028
	Secondary Phloem	b*	0.009
	<i>ZEP</i>	whole root	β-carotene
phytoene			0.049
phytofluene			0.050
total			0.042
Epidermis		L*	0.028
		a*	0.004
		b*	0.032
		C*	0.023
		h	0.005
Secondary Phloem		L*	0.003
		a*	0.004
		b*	0.046
		h	0.001
Secondary Xylem	a*	0.004	
	h	0.016	
<i>NCED1</i>	Secondary Xylem	L*	0.033

doi:10.1371/journal.pone.0116674.t005

on ABA content and carotenoid storage capacity, reinforce our conclusion that *ZEP* is a major candidate gene governing carotenoid content in carrot roots.

### Limiting role of the synthesis phytoene desaturase gene

We identified *PDS* and his cofactor *PTOX* [57] as associated with total carotenoid, phytoene, phytofluene, α-carotene and lutein content. This suggests that these genes are involved in the global carotenoid accumulation. An early regulation of the pathway may explain the large number of associations detected for these genes. These results are consistent with previous

results which identified phytoene accumulation as a major regulatory step in the carotenoid pathway in plants [4,6].

Moreover a putative signature of selection for *PDS* after domestication of carrot was shown by [58]. This may be in relation with an early control of the metabolic pathway. *PTOX* has been identified as playing a major role in carotenoid accumulation. *Arabidopsis* mutant IMMUTANS [59,60] and tomato mutant *ghost* [61] exhibit an impaired function of *PTOX* leading to a deficient phytoene desaturation and an accumulation of the carotenoid precursor phytoene.

Just et al. [21] also identified *CHXE*, *NCED2* and *PDS* as potential candidate genes linked to another major carotenoid QTL, proposed as the *Y* locus. But they could not conclude on the effect of each gene in this QTL region. As neither SNPs nor haplotypes from *CHXE* and *NCED2* were associated with carotenoid traits, our results suggest that *PDS* may be one probable candidate explaining the *Y* locus.

Unfortunately, *PTOX* is not yet mapped into the carrot genome, which would be needed to better explain its role. A recent study suggests a complicated role for the *PTOX* which is also involved in chloroplast biogenesis and in photosystem II photoprotection [62].

## Towards a particular mechanism driving the metabolic flux through the $\alpha$ -branch to lutein

The orientation of the pathway towards the  $\alpha$ -branch or the  $\beta$ -branch results in different pattern of carotenoid content. However, the underlying mechanism remains unclear in carrot. The metabolic node is known to be a major regulation step in the pathway [4]. For example, variation in *LCYE* in maize explained 58% of the variation in the two branches [63]. This was also observed in *Arabidopsis thaliana* [64], *Brassica napus* [65] and *Solanum tuberosum* [66]. But we did not detect any association for  $\alpha$ -carotene and lutein content with the  $\epsilon$ -lycopene cyclase gene. This suggests a more complex regulation orientating the pathway to one of the branch, controlled by a genetic factor different from the tested carotenoid biosynthetic genes. Actually, Arango et al. [23] have shown the specific role of carotene hydroxylase *CYP97A3* gene in the  $\alpha$ -carotene level.

The study of correlation between all traits showed the independence of lutein content. All carotenoid compounds except lutein were correlated together. This also suggests a particular mechanism in lutein accumulation from  $\alpha$ -carotene.

## Conclusion

For the first time, we performed an association analysis on *Daucus carota* L. Moreover we developed an original unstructured population with a broad genetic basis, limiting the risk of spurious association. However, as all individuals descended from the same parents, relatedness estimated with kinship matrixes had to be included in the association model.

We identified several SNPs and genes associated with carotenoid content and color components. Our results bring evidence that zeaxanthin epoxydase and phytoene desaturase are candidate genes involved in carotenoid accumulation of non-photosynthetic organs. Our study brings new insight into the carotenoid pathway functioning by stressing out two major steps in carotenoid metabolism and catabolism in a storage organ. Functional validation and dissection of the regulation of *ZEP* expression may clarify the mechanisms involved in carotenoid accumulation. Clarification of the involvement of the *PTOX* has also to be investigated. A mechanism explaining both the accumulation of xanthophylls and the pathway orientation towards the  $\alpha$ -branch as well as lutein accumulation still remains to be specified. However, the genes

identified in this study as associated with color components and carotenoid content may be useful in marker-assisted selection for carotenoid content enhancement in a breeding program.

## Supporting Information

**S1 Table.** Primers used to amplify candidate genes fragments.  
(DOC)

**S2 Table.** Context sequences of genotyped SNPs.  
(XLSX)

**S1 Fig.** Linkage disequilibrium between SNPs in all genes.  $r^2$  and LD significance are represented at the top and bottom of the matrix, respectively.  
(TIF)

## Acknowledgments

We would like to thank Françoise Gros of the ‘Plateforme de séquençage et de Génotypage’ (IFR26, Nantes, France) for sequencing PCR products. We wish to thank Charles Poncet of the ‘Plateforme Gentyane’ (UMR1095 GDEC, Clermont-Ferrand, France) for microsatellite markers genotyping and Valérie Le Clerc for her critical review of this article. We are grateful to our colleagues from INEM (Remi Gardet and Jacky Granger, Agrocampus Ouest, Angers) involved in field experiments.

## Author Contributions

Conceived and designed the experiments: MJ SG DP EG. Performed the experiments: MJ CDL MM SH AS EG. Analyzed the data: MJ SG DP EG. Contributed reagents/materials/analysis tools: SG CDL LH. Wrote the paper: MJ SG MB DP EG.

## References

1. Rao AV, Rao LG (2007) Carotenoids and human health. *Pharmacological Res* 55:207–216. doi: [10.1016/j.phrs.2007.01.012](https://doi.org/10.1016/j.phrs.2007.01.012) PMID: [17349800](https://pubmed.ncbi.nlm.nih.gov/17349800/)
2. Surles RL, Weng N, Simon PW, Tanumihardjo SA (2004) Carotenoid profiles and consumer sensory evaluation of specialty carrots (*Daucus carota* L.) of various colors. *J Agric Food Chem* 52:3417–3421. doi: [10.1021/jf035472m](https://doi.org/10.1021/jf035472m) PMID: [15161208](https://pubmed.ncbi.nlm.nih.gov/15161208/)
3. Baranski R, Allender C, Klimek-Chodacka M (2012) Towards better tasting and more nutritious carrots: Carotenoid and sugar content variation in carrot genetic resources. *Food Res Int* 47:182–187.
4. Cazzonelli CI, Pogson BJ (2010) Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci* 15:266–274. doi: [10.1016/j.tplants.2010.02.003](https://doi.org/10.1016/j.tplants.2010.02.003) PMID: [20303820](https://pubmed.ncbi.nlm.nih.gov/20303820/)
5. Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. *Curr Opin Plant Biol* 4:210–218. doi: [10.1016/S1369-5266\(00\)00163-1](https://doi.org/10.1016/S1369-5266(00)00163-1) PMID: [11312131](https://pubmed.ncbi.nlm.nih.gov/11312131/)
6. Howitt CA, Pogson BJ (2006) Carotenoid accumulation and function in seeds and non-green tissues. *Plant Cell Env* 29:435–445. doi: [10.1111/j.1365-3040.2005.01492.x](https://doi.org/10.1111/j.1365-3040.2005.01492.x) PMID: [17080597](https://pubmed.ncbi.nlm.nih.gov/17080597/)
7. Sandmann G, Römer S, Fraser PD (2006) Understanding carotenoid metabolism as a necessity for genetic engineering of crop plants. *Metab Eng* 8:291–302. doi: [10.1016/j.ymben.2006.01.005](https://doi.org/10.1016/j.ymben.2006.01.005) PMID: [16621640](https://pubmed.ncbi.nlm.nih.gov/16621640/)
8. Lu S, Li L (2008) Carotenoid metabolism: biosynthesis, regulation, and beyond. *J Integr Plant Biol* 50:778–785. doi: [10.1111/j.1744-7909.2008.00708.x](https://doi.org/10.1111/j.1744-7909.2008.00708.x) PMID: [18713388](https://pubmed.ncbi.nlm.nih.gov/18713388/)
9. Antonio C, Santos F, Senalik D, Simon PW (2005) Path analysis suggests phytoene accumulation is the key step limiting the carotenoid pathway in white carrot roots. *Genetics Mol Biol* 28:287–293.
10. Huh JH, Kang BC, Nahm SH, Kim S, Ha KS, et al. (2001) A candidate gene approach identified phytoene synthase as the locus for mature fruit color in red pepper (*Capsicum* spp.). *Theor Appl Genet* 102:524–530.



11. Maass D, Arango J, Wüst F, Beyer P, Welsch R (2009) Carotenoid crystal formation in *Arabidopsis* and carrot roots caused by increased phytoene synthase protein levels. PLoS One 4:e6373. doi: [10.1371/journal.pone.0006373](https://doi.org/10.1371/journal.pone.0006373) PMID: [19636414](https://pubmed.ncbi.nlm.nih.gov/19636414/)
12. Ravel C, Dardevet M, Leenhardt F, Bordes J, Joseph JL, et al. (2012) Improving the yellow pigment content of bread wheat flour by selecting the three homoeologous copies of Psy1. Mol Breed 31:87–99.
13. Rodriguez-Villalon A, Gas E, Rodriguez-Concepcion M (2009) Phytoene synthase activity controls the biosynthesis of carotenoids and the supply of their metabolic precursors in dark-grown *Arabidopsis* seedlings. Plant J 60:424–435. doi: [10.1111/j.1365-3113X.2009.03966.x](https://doi.org/10.1111/j.1365-3113X.2009.03966.x) PMID: [19594711](https://pubmed.ncbi.nlm.nih.gov/19594711/)
14. Chandler K, Lipka AE, Owens BF, Li H, Buckler ES, et al. (2013) Genetic analysis of visually scored orange kernel color in maize. Crop Sci 53:189–200.
15. Paran I, van der Knaap E (2007) Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. J Exp Bot 58:3841–3852. doi: [10.1093/jxb/erm257](https://doi.org/10.1093/jxb/erm257) PMID: [18037678](https://pubmed.ncbi.nlm.nih.gov/18037678/)
16. Roncallo PF, Cervigni GL, Jensen C, Miranda R, Carrera AD, et al. (2012) QTL analysis of main and epistatic effects for flour color traits in durum wheat. Euphytica 185:77–92.
17. Brand A, Borovsky Y, Meir S, Rogachev I, Aharoni A, et al. (2012) Pc8.1, a major QTL for pigment content in pepper fruit, is associated with variation in plastid compartment size. Planta 235:579–588. doi: [10.1007/s00425-011-1530-9](https://doi.org/10.1007/s00425-011-1530-9) PMID: [21987007](https://pubmed.ncbi.nlm.nih.gov/21987007/)
18. Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, et al. (2000) Engineering the provitamin A ( $\beta$ -carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science 287:303–305. doi: [10.1126/science.287.5451.303](https://doi.org/10.1126/science.287.5451.303) PMID: [10634784](https://pubmed.ncbi.nlm.nih.gov/10634784/)
19. Santos CAF, Simon PW (2006) Heritabilities and minimum gene number estimates of carrot carotenoids. Euphytica 151:79–86.
20. Buishand JG, Gabelman WH (1979) Investigations on the inheritance of color and carotenoid content in phloem and xylem of carrot roots (*Daucus carota* L.). Euphytica 28:611–632.
21. Just BJ, Santos CAF, Yandell BS, Simon PW (2009) Major QTL for carrot color are positionally associated with carotenoid biosynthetic genes and interact epistatically in a domesticated  $\times$  wild carrot cross. Theor Appl Genet 119:1155–1169. doi: [10.1007/s00122-009-1117-z](https://doi.org/10.1007/s00122-009-1117-z) PMID: [19657616](https://pubmed.ncbi.nlm.nih.gov/19657616/)
22. Simon PW (1996) Inheritance and expression of purple and yellow storage root color in carrot. J Hered 87:63–66.
23. Arango J, Jourdan M, Geoffriau E, Beyer P, Welsch R (2014) Carotene hydroxylase activity determines the levels of both  $\alpha$ -carotene and total carotenoids in orange carrots. Plant Cell. PMID: [24858934](https://pubmed.ncbi.nlm.nih.gov/24858934/)
24. Santos CAF, Simon PW (2002) QTL analyses reveal clustered loci for accumulation of major provitamin A carotenes and lycopene in carrot roots. Mol Genet Genomics 268:122–129. doi: [10.1007/s00438-002-0735-9](https://doi.org/10.1007/s00438-002-0735-9) PMID: [12242507](https://pubmed.ncbi.nlm.nih.gov/12242507/)
25. Just BJ, Santos CAF, Fonseca MEN, Boiteux LS, Oloizia BB, et al. (2007) Carotenoid biosynthesis structural genes in carrot (*Daucus carota*): isolation, sequence-characterization, single nucleotide polymorphism (SNP) markers and genome mapping. Theor Appl Genet 114:693–704. doi: [10.1007/s00122-006-0469-x](https://doi.org/10.1007/s00122-006-0469-x) PMID: [17186217](https://pubmed.ncbi.nlm.nih.gov/17186217/)
26. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, et al. (2009) Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21:2194–2202. doi: [10.1105/tpc.109.068437](https://doi.org/10.1105/tpc.109.068437) PMID: [19654263](https://pubmed.ncbi.nlm.nih.gov/19654263/)
27. Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208. doi: [10.1038/ng1702](https://doi.org/10.1038/ng1702) PMID: [16380716](https://pubmed.ncbi.nlm.nih.gov/16380716/)
28. Ehrenreich IM, Hanzawa Y, Chou L, Roe JL, Kover PX, et al. (2009) Candidate gene association mapping of *Arabidopsis* flowering time. Genetics 183:325–335. doi: [10.1534/genetics.109.105189](https://doi.org/10.1534/genetics.109.105189) PMID: [19581446](https://pubmed.ncbi.nlm.nih.gov/19581446/)
29. González-Martínez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB (2007) Association genetics in *Pinus taeda* L. I. Wood property traits. Genetics 175:399–409. doi: [10.1534/genetics.106.061127](https://doi.org/10.1534/genetics.106.061127) PMID: [17110498](https://pubmed.ncbi.nlm.nih.gov/17110498/)
30. Weber A, Clark RM, Vaughn L, Sánchez-Gonzalez JDJ, Yu J, et al. (2007) Major regulatory genes in maize contribute to standing variation in teosinte (*Zea mays* ssp. *parviglumis*). Genetics 177:2349–2359. doi: [10.1534/genetics.107.080424](https://doi.org/10.1534/genetics.107.080424) PMID: [17947410](https://pubmed.ncbi.nlm.nih.gov/17947410/)
31. Cardoso S, Lau W, Dias JE, Fevereiro P, Maniatis N (2012) A candidate-gene association study for berry colour and anthocyanin content in *Vitis vinifera* L. PLoS One 7(9): e46021 doi: [10.1371/journal.pone.0046021](https://doi.org/10.1371/journal.pone.0046021) PMID: [23029369](https://pubmed.ncbi.nlm.nih.gov/23029369/)
32. Nordborg M, Weigel D (2008) Next-generation genetics in plants. Nature 456:720–3. doi: [10.1038/nature07629](https://doi.org/10.1038/nature07629) PMID: [19079047](https://pubmed.ncbi.nlm.nih.gov/19079047/)

33. Clotault J, Geoffriau E, Lionneton E, Briard M, Peltier D (2010) Carotenoid biosynthesis genes provide evidence of geographical subdivision and extensive linkage disequilibrium in the carrot. *Theor Appl Genet* 121:659–672. doi: [10.1007/s00122-010-1338-1](https://doi.org/10.1007/s00122-010-1338-1) PMID: [20411232](https://pubmed.ncbi.nlm.nih.gov/20411232/)
34. Baranski R, Maksylewicz-Kaul A, Nothnagel T, Cavagnaro PF, Simon PW, et al. (2011) Genetic diversity of carrot (*Daucus carota* L.) cultivars revealed by analysis of SSR loci. *Genet Res Crop Evol* 59:163–170.
35. Soufflet-Freslon V, Jourdan M, Clotault J, Huet S, Briard M, et al. (2013) Functional gene polymorphism to reveal species history: the case of the *CRTISO* gene in cultivated carrots. *PloS One* 8(8):e70801 doi: [10.1371/journal.pone.0070801](https://doi.org/10.1371/journal.pone.0070801) PMID: [23940644](https://pubmed.ncbi.nlm.nih.gov/23940644/)
36. Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, et al. (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet* 28:286–289. doi: [10.1038/90135](https://doi.org/10.1038/90135) PMID: [11431702](https://pubmed.ncbi.nlm.nih.gov/11431702/)
37. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587. PMID: [12930761](https://pubmed.ncbi.nlm.nih.gov/12930761/)
38. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959. PMID: [10835412](https://pubmed.ncbi.nlm.nih.gov/10835412/)
39. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909. doi: [10.1038/ng1847](https://doi.org/10.1038/ng1847) PMID: [16862161](https://pubmed.ncbi.nlm.nih.gov/16862161/)
40. Le Clerc V, Briard M, Granger J, Delettre J (2003) Genebank biodiversity assessments regarding optimal sample size and seed harvesting techniques for the regeneration of carrot accessions. *Biodivers Conserv* 12:2227–2236.
41. Gómez R, Pardo JE, Navarro F, Varón R (1998) Colour differences in paprika pepper varieties (*Capsicum annum* L.) cultivated in a greenhouse and in the open air. *J Sci Food Agric* 77(2): 268–272.
42. Singha S, Baugher TA, Townsend EC, D'Souza MC (1991) Anthocyanin distribution in 'Delicious' apples and the relationship between anthocyanin concentration and chromaticity values. *J Am Soc Hortic Sci* 116 (3 ): 497–499.
43. Clotault J, Peltier D, Berruyer R, Thomas M, Briard M, et al. (2008) Expression of carotenoid biosynthesis genes during carrot root development. *J Exp Bot* 59:3563–3573. doi: [10.1093/jxb/ern210](https://doi.org/10.1093/jxb/ern210) PMID: [18757491](https://pubmed.ncbi.nlm.nih.gov/18757491/)
44. Inbaraj BS, Chien JT, Chen BH (2006) Improved high performance liquid chromatographic method for determination of carotenoids in the microalga *Chlorella pyrenoidosa*. *J Chromatogr A* 1102:193–199. doi: [10.1016/j.chroma.2005.10.055](https://doi.org/10.1016/j.chroma.2005.10.055) PMID: [16298378](https://pubmed.ncbi.nlm.nih.gov/16298378/)
45. De Rosso VV, Mercadante AZ (2007) Identification and quantification of carotenoids, by HPLC-PDA-MS/MS, from Amazonian fruits. *J Agric Food Chem* 55:5062–5072. doi: [10.1021/jf0705421](https://doi.org/10.1021/jf0705421) PMID: [17530774](https://pubmed.ncbi.nlm.nih.gov/17530774/)
46. Bradeen JM, Simon PW (1998) Conversion of an AFLP fragment linked to the carrot Y2 locus to a simple, codominant, PCR-based marker form. *Theor Appl Genet*, 97:960–967.
47. Briard M, Le Clerc V, Grzebelus D, Senalik D, Simon P (2000) Modified protocols for rapid carrot genomic DNA extraction and AFLP™ analysis using silver stain or radioisotopes. *Plant Mol Biol Report* 18:235–241.
48. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452. doi: [10.1093/bioinformatics/btp187](https://doi.org/10.1093/bioinformatics/btp187) PMID: [19346325](https://pubmed.ncbi.nlm.nih.gov/19346325/)
49. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, et al. (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635. doi: [10.1093/bioinformatics/btm308](https://doi.org/10.1093/bioinformatics/btm308) PMID: [17586829](https://pubmed.ncbi.nlm.nih.gov/17586829/)
50. Cavagnaro PF, Chung SM, Manin S, Yildiz M, Ali A, et al. (2011) Microsatellite isolation and marker development in carrot—genomic distribution, linkage mapping, genetic diversity analysis and marker transferability across *Apiaceae*. *BMC Genomics* 12:386. doi: [10.1186/1471-2164-12-386](https://doi.org/10.1186/1471-2164-12-386) PMID: [21806822](https://pubmed.ncbi.nlm.nih.gov/21806822/)
51. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. doi: [10.1111/j.1365-294X.2005.02553.x](https://doi.org/10.1111/j.1365-294X.2005.02553.x) PMID: [15969739](https://pubmed.ncbi.nlm.nih.gov/15969739/)
52. Kang HM, Zaitlen N, Wade CM, Kirby A, Heckerman D, et al. (2008) Efficient control of population structure in model organism association mapping. *Genetics* 178:1709–1723. doi: [10.1534/genetics.107.080101](https://doi.org/10.1534/genetics.107.080101) PMID: [18385116](https://pubmed.ncbi.nlm.nih.gov/18385116/)
53. Auzanneau J, Huyghe C, Escobar-Gutiérrez AJ, Julier B, Gastal F, et al. (2011) Association study between the gibberellic acid insensitive gene and leaf length in a *Lolium perenne* L. synthetic variety. *BMC Plant Biol* 11:183. doi: [10.1186/1471-2229-11-183](https://doi.org/10.1186/1471-2229-11-183) PMID: [22204490](https://pubmed.ncbi.nlm.nih.gov/22204490/)

54. Wolters AM, Uitdewilligen J, Kloosterman B, Hutten RCB, Visser RGF, et al. (2010) Identification of alleles of carotenoid pathway genes important for zeaxanthin accumulation in potato tubers. *Plant Mol Biol* 73:659–671. doi: [10.1007/s11103-010-9647-y](https://doi.org/10.1007/s11103-010-9647-y) PMID: [20490894](https://pubmed.ncbi.nlm.nih.gov/20490894/)
55. Vallabhaneni R, Wurtzel ET (2009) Timing and biosynthetic potential for carotenoid accumulation in genetically diverse germplasm of maize. *Plant Physiol* 150:562–572. doi: [10.1104/pp.109.137042](https://doi.org/10.1104/pp.109.137042) PMID: [19346441](https://pubmed.ncbi.nlm.nih.gov/19346441/)
56. Galpaz N, Wang Q, Menda N, Zamir D, Hirschberg J (2008) Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. *Plant J*, 53(5):717–730. doi: [10.1111/j.1365-313X.2007.03362.x](https://doi.org/10.1111/j.1365-313X.2007.03362.x) PMID: [17988221](https://pubmed.ncbi.nlm.nih.gov/17988221/)
57. Shahbazi M, Gilbert M, Labouré AM, Kuntz M (2007) Dual role of the plastid terminal oxidase in tomato. *Plant Physiol* 145:691–702. doi: [10.1104/pp.107.106336](https://doi.org/10.1104/pp.107.106336) PMID: [17873087](https://pubmed.ncbi.nlm.nih.gov/17873087/)
58. Clotault J, Peltier D, Soufflet-Freslon V, Briard M, Geoffriau E (2012) Differential selection on carotenoid biosynthesis genes as a function of gene position in the metabolic pathway: a study on the carrot and dicots. *PLoS One* 7:e38724. doi: [10.1371/journal.pone.0038724](https://doi.org/10.1371/journal.pone.0038724) PMID: [22737218](https://pubmed.ncbi.nlm.nih.gov/22737218/)
59. Carol P, Stevenson D, Bisanz C, Breitenbach J, Sandmann G, et al. (1999) Mutations in the *Arabidopsis* gene IMMUTANS cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. *Plant Cell* 11:57–68. doi: [10.2307/3870838](https://doi.org/10.2307/3870838) PMID: [9878632](https://pubmed.ncbi.nlm.nih.gov/9878632/)
60. Wu D, Wright D, Wetzl C, Voytas DF, Rodermeil S (1999) The IMMUTANS variegation locus of *Arabidopsis* defines a mitochondrial alternative oxidase homolog that functions during early chloroplast biogenesis. *Plant Cell* 11:43–55. doi: [10.2307/3870837](https://doi.org/10.2307/3870837) PMID: [9878631](https://pubmed.ncbi.nlm.nih.gov/9878631/)
61. Josse EM, Simkin AJ, Gaffé J, Labouré AM, Kuntz M, et al. (2000) A plastid terminal oxidase associated with carotenoid desaturation during chromoplast differentiation. *Plant Physiol* 123:1427–36. doi: [10.1104/pp.123.4.1427](https://doi.org/10.1104/pp.123.4.1427) PMID: [10938359](https://pubmed.ncbi.nlm.nih.gov/10938359/)
62. McDonald AE, Ivanov AG, Bode R, Maxwell DP, Rodermeil SR, et al. (2011) Flexibility in photosynthetic electron transport: the physiological role of plastoquinol terminal oxidase (PTOX). *Acta Biochim Biophys* 1807:954–967. doi: [10.1016/j.bbabo.2010.10.024](https://doi.org/10.1016/j.bbabo.2010.10.024) PMID: [21056542](https://pubmed.ncbi.nlm.nih.gov/21056542/)
63. Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, et al. (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319:330–333. doi: [10.1126/science.1150255](https://doi.org/10.1126/science.1150255) PMID: [18202289](https://pubmed.ncbi.nlm.nih.gov/18202289/)
64. Pogson BJ, Rissler HM (2000) Genetic manipulation of carotenoid biosynthesis and photoprotection. *Philos Trans R Soc Lond B Biol Sci* 355:1395–1403. doi: [10.1098/rstb.2000.0701](https://doi.org/10.1098/rstb.2000.0701) PMID: [11127994](https://pubmed.ncbi.nlm.nih.gov/11127994/)
65. Yu B, Lydiate DJ, Young LW, Schäfer UA, Hannoufa A (2008) Enhancing the carotenoid content of *Brassica napus* seeds by downregulating lycopene epsilon cyclase. *Transgenic Res* 17:573–585. doi: [10.1007/s11248-007-9131-x](https://doi.org/10.1007/s11248-007-9131-x) PMID: [17851775](https://pubmed.ncbi.nlm.nih.gov/17851775/)
66. Diretto G, Tavazza R, Welsch R, Pizzichini D, Mourgues F, et al. (2006) Metabolic engineering of potato tuber carotenoids through tuber-specific silencing of lycopene epsilon cyclase. *BMC Plant Biol* 6:13. doi: [10.1186/1471-2229-6-13](https://doi.org/10.1186/1471-2229-6-13) PMID: [16800876](https://pubmed.ncbi.nlm.nih.gov/16800876/)