



## Path analysis suggests phytoene accumulation is the key step limiting the carotenoid pathway in white carrot roots

Carlos Antonio Fernandes Santos<sup>1</sup>, Douglas Senalik<sup>2</sup> and Philipp Wallace Simon<sup>2</sup>

<sup>1</sup>*Empresa Brasileira de Pesquisa Agropecuária, Semi-Árido, Petrolina, PE, Brazil.*

<sup>2</sup>*University of Wisconsin, Agricultural Research Service, Department of Horticulture, Vegetable Crops Research Unit, Madison, WI, USA.*

### Abstract

Two F<sub>2</sub> carrot (*Daucus carota* L.) populations (orange rooted Brasília x very dark orange rooted High Carotene Mass - HCM cross and the dark orange rooted cultivated variety B493 x white rooted wild carrot Queen Anne's Lace - QAL cross) with very unrelated genetic backgrounds were used to investigate intrinsic factors limiting carotenoid accumulation in carrots by applying phenotypic correlation and path analysis to study the relationships between major root carotenes, root color and several other morphological traits. Most of the correlations between traits were close and agreed in sign between the two populations. Root weight had a moderate to highly significant positive correlation with leaf length, root length and top and middle root diameter. Although phenotypic correlations failed to identify the order of the substrates and products in the carotenoid pathway the correct order of substrates and products (phytoene → zeta-carotene → lycopene) was identified in the causal diagram of beta-carotene for the Brasília x HCM population. Path analysis of beta-carotene synthesis in the B493 x QAL population suggested that selection for root carotenes had little effect on plant morphological traits. Causal model of beta-carotene and lycopene in the B493 x QAL population suggested that phytoene synthesis is the key step limiting the carotenoid pathway in white carrots. Path analysis, first presented by Sewall Wright to study quantitative traits, appears to be a powerful statistical approach for the identification of key compounds in complex pathways.

*Key words:* *Daucus carota*, correlation, path diagram, genetic engineering.

Received: July 21, 2004; Accepted: November 25, 2004.

### Introduction

Vitamin A deficiency is not only widespread in developing countries but is also found in developed countries in poor urban populations and among the elderly, heavy drinkers and smokers (Giuliano *et al.* 2000). Consumption of horticultural crops provides more than 70% of vitamin A for the world population (Simpson, 1983), with carrots accounting for 30% of the total vitamin A precursor in countries such as the United States of America (Simon, 1992).

The carotenoid biosynthetic pathway is a well established biochemical pathway which has been studied in many plants (Cunningham and Gantt, 1998; Sandmann, 1998), fungi and microorganisms (Sandmann, 1998; Armstrong, 1994). The formation of the colorless carotene phytoene from two molecules of geranylgeranyl diphosphate (GGDP) or geranylgeranyl pyrophosphate

(GGPP) is the first committed step in the carotenoid pathway. Phytoene undergoes a series of four desaturation reactions that result in the sequential formation of phytofluene, zeta-carotene ( $\zeta$ -carotene), neurosporene and then the red-colored lycopene. A single gene product, lycopene beta-cyclase ( $\beta$ -cyclase) (LCYB), catalyzes the formation of the bicyclic  $\beta$ -carotene (with two  $\beta$  rings) from the linear symmetrical lycopene in plants and cyanobacteria (Cunningham and Gantt, 1998) as demonstrated in studies with *Erwinia herbicola* and tomatoes (Sandmann, 1998). In the case of alpha-carotene ( $\alpha$ -carotene), with one  $\beta$  and one epsilon ( $\epsilon$ ) ring, two different enzymes, LCYB and lycopene  $\epsilon$ -cyclase (LCYE), are involved (Sandmann, 1998). It has not been determined whether the route to  $\alpha$ -carotene biosynthesis is only via the  $\epsilon$  ring first or if it can also proceed via the  $\beta$  ring first, and carotenoids with two  $\epsilon$  rings being unusual in plants (Cunningham and Gantt, 1998).

A high correlation between two variables can result from the effect of a third or group of variables, because the total correlation simply measures mutual association with-

out regard to causation (Bhatt, 1973). Wright (1921) presented an approach to analyze networks of causes and effects by breaking down the correlation into direct and indirect components to produce what he called 'path coefficients'. According to Li (1956), the separation of a correlation coefficient into various components is analogous to the analysis of variance. Path analysis has been applied in many fields, such as population genetics, social and economics science, evolution and plant and animal breeding (Lynch and Walsh, 1998).

The general rules and features of path analysis are: 1) the analysis is based on a unidirectional forward-flowing cause and effect diagram; 2) a double-headed arrow denotes correlation and a single-headed arrow denotes a path coefficient ( $p_{yx}$ ); 3) only dependent variables should have a residual term; 4)  $p_{yx}$  may be greater or less than one with positive or negative values; and 5) the sum of all direct and indirect effects exactly equals the correlation coefficient (Li, 1975; Lynch and Walsh, 1998; Hatcher, 1994).

Relationships between compounds in the carotenoid pathway and some vegetative traits in two unrelated carrot (*Daucus carota* L.) crosses (Brasilia x High Carotene Mass - HCM cross and the B493 x Queen Anne's Lace - QAL cross) were studied by applying path analysis. It was envisaged that the information gained from this type of analysis would not only shed light on the mechanisms by which selection for altered levels of carotenoids may cause changes in morphological traits but would help in the identification of the major steps limiting carotenoid accumulation. The information produced by such an analysis could also provide insights into the evolution of the carotenoid pathway, and identify rate-limiting steps in this pathway where focused selection and genetic transformation may further increase carotenoid content. This is the first report applying path analysis to the study of the interrelationships between products in a biosynthetic pathway.

## Material and Methods

### Plant material and character measurement

Two  $F_2$  carrot populations with very unrelated genetic backgrounds were derived from single  $F_1$  plants resulting from crosses between Brasilia x HCM and B493 x QAL. The Brasilia variety is a typical orange carrot developed in Brazil for production in warmer production areas (Hamerschmidt, 1993) and has a typical carotene content ranging from 50 to 90  $\mu\text{gg}^{-1}$  while the HCM variety, developed from a cross between Asian and European germplasm (Simon *et al.* 1989), is a very dark orange with an average carotene content of 460 to 499  $\mu\text{gg}^{-1}$ . B493 is a dark orange inbred carrot derived from European germplasm with a carotene content of 180 to 210  $\mu\text{gg}^{-1}$  (Simon *et al.* 1990) while QAL is a white (carotenoid-free) wild carrot (*D. carota* var. *carota*) which is widely distributed in temperate regions of eastern North and South America and from the Atlantic

coast of Eastern Europe to western China (Rubatzky *et al.* 1999), the QAL parent plant used in this study being from Madison, WI, USA. The population sizes used in our study were 62 for the Brasilia (orange) x HCM (very dark orange) cross and 83 for the B493 (orange) x QAL (white) cross.

Phytoene (PHY),  $\zeta$ -carotene (ZET), lycopene (LIC),  $\beta$ -carotene (BET) and  $\alpha$ -carotene (ALP) were extracted as described by Simon and Wolff (1987) and quantified by high-performance liquid chromatography (HPLC) using system B as described by Khachik *et al.* (1992) with detection provided by a Waters 996 Photodiode Array Detector (Waters Associates, Milford, MA, USA). Leaf length (LL), root length (RL), root weight (RW), top and middle root diameter (TRD and MRD, respectively) and total dissolved solids (TDS) were obtained as described by Stommel and Simon (1989), Rubatzky *et al.* (1999) and Simon (2000). Root color (RC) was based on visual evaluation of root cross sections using one scale (1 = very pale orange, 2 = pale orange, 3 = orange, 4 = dark orange, and 5 = very dark orange) for the Brasilia x HCM cross and another scale (7 = white, 8 = yellow, 9 = pale orange and 10 = orange) for the B493 x QAL cross. Note that for root length, weight and color only the main tap-root was considered.

### Phenotypic correlation

The twelve characters were paired in all possible combinations and the values of each pair were summed to estimate the covariance according the formula  $Cov_{XY} = (\hat{\sigma}_{X+Y}^2 - \hat{\sigma}_X^2 - \hat{\sigma}_Y^2) / 2$ . Phenotypic correlations were estimated according to the formula  $r_{XY} = (Cov_{XY}) / (\hat{\sigma}_X^2 \hat{\sigma}_Y^2)^{1/2}$ . All correlations were tested by the Student's t-test at the 1% and 5% probability level for  $n-2$  degrees of freedom (SAS 1989).

### Path analysis

A network of interrelationships between measured characters was established using a causal path diagram with lycopene as the dependent variable and  $\zeta$ -carotene and phytoene as primary explanatory variables. Leaf length, root length, root weight, top and middle root diameter, total dissolved solids and root color were analyzed as secondary explanatory variables for lycopene. The causal diagram of  $\beta$ -carotene considered lycopene,  $\zeta$ -carotene and phytoene as primary explanatory variables and leaf length, root length, root weight, top and middle root diameter, total dissolved solids and root color as secondary explanatory variables.

The system of equations derived from the causal diagram was solved using normal equations in matrix terms:  $X'X\hat{\beta}$  (Li, 1975; Cruz and Regazzi, 1997). The ordinary least squares solution was provided by solving for  $\hat{\beta} = (X'X)^{-1}X'Y$ , where  $X'X$  is a non-singular matrix of phenotypic correlations between independent variables,  $\hat{\beta}$  is a vector of path coefficients to be estimated and  $X'Y$  is a

vector of phenotypic correlation between dependent and independent variables.

All estimations were obtained with the Genes software (Cruz, 1998) and using the Calis procedure (SAS, 1989). Multicollinearity tests in the  $X'X$  matrices were performed as described by Chatterjee and Price (1991) using the Genes software (Cruz, 1998).

## Results

For the B493 (dark orange) x QAL (white)  $F_2$  population the following range of carotenoid pathway metabolite values ( $\mu\text{gg}^{-1}$ ) were found: phytoene, 0 to 406 (mean = 27);  $\zeta$ -carotene, 0 to 177 (mean = 8); lycopene, 0 to 345 (mean = 10),  $\beta$ -carotene, 0 to 266 (mean = 20); and  $\alpha$ -carotene, 0 to 197, (mean = 6). For the Brasilia (orange) x HCM (very dark orange)  $F_2$  population the values ( $\mu\text{gg}^{-1}$ ) were: phytoene, 189 to 1249 (mean = 491);  $\zeta$ -carotene, 114 to 344 (mean = 144); lycopene, 0 to 77 (mean = 39);  $\beta$ -carotene, 177 to 912 (mean = 515); and  $\alpha$ -carotene, 116 to 1011 (mean = 489).

## Correlation analysis

The overall correlations between vegetative characters were similar and agreed in sign for both populations, except that there were some differences as regards total dissolved solids (Table 1). In both populations root weight correlated positively and significantly with leaf length, root length, top-root diameter and middle-root diameter. The root length and the diameter at the top of the root had the highest positive correlation with root weight (Table 1). Similar results have been reported by Natarajan and Arumugam (1980) and Pariari *et al.* (1992). Increased root

weight is known to be an important character as regards increased total carrot yield (Krarup and Mosnaim, 1980).

Total dissolved solids was not only negatively correlated with root weight but also with all other vegetative characters evaluated in this study (Table 1). A negative correlation between total dissolved solids and leaf length, root weight and root yield has also been reported by Randhir *et al.* (1992) in a study involving 40 carrot populations. In our study, the correlation between total dissolved solids and root weight was negatively significant for the B493 x QAL (cultivated x wild) population and non-significant for the Brasilia x HCM (cultivated x cultivated) population, which may indicate break up of linkage between gene blocks controlling yield in cultivated carrots with the result that although yield increases total dissolved solids remain more or less constant. Many carrot breeding programs select for high levels of total dissolved solids because selection for this trait can be effective in improving sweetness and flavor (Stommel and Simon, 1989).

Root color correlations with all characters except  $\alpha$ -carotene were non-significant for the Brasilia x HCM population but for the B493 x QAL population root color was positively and significantly correlated with all major carotenes (Table 1). These results suggest that when there is a scale of easily graded root color this trait is a very efficient selection parameter for the improvement of carotene content, but since we also observed continuous variation our results also supports the previously observed (Buishand and Gabelman, 1979) fact that it is difficult to divide orange into discrete intensity classes. Emsweller *et al.* (1935) reported a correlation of 0.83 between carrot root color intensity and average carotene content, while in our study the

**Table 1** - Phenotypic correlations between some morphological traits and major carotene content in two different  $F_2$  carrot populations. Upper diagonal = Brasilia x HCM, lower diagonal = B493 x QAL.

Trait	LL	RL	RW	TRD	MRD	TDS	RC	ZET	ALP	BET	PHY	LIC
LL		0.47**	0.53**	0.55**	0.54**	-0.03	0.20	0.12	0.50**	0.28*	0.07	0.40**
RL	0.49**		0.71**	0.51**	0.52**	-0.10	0.01	0.23	0.25*	0.34*	0.09	0.39**
RW	0.56**	0.66**		0.72**	0.70**	-0.17	-0.05	0.37**	0.38**	0.52**	0.17	0.45**
TRD	0.39**	0.33**	0.61**		0.90**	-0.19	-0.12	0.33*	0.42**	0.55**	0.11	0.52**
MRD	0.46**	0.49**	0.75**	0.59**		-0.25*	-0.01	0.38**	0.47**	0.60**	0.12	0.58**
TDS	-0.32*	-0.13	-0.36**	-0.24*	-0.22		0.17	0.22	0.07	0.06	0.28*	0.01
RC	-0.10	-0.10	-0.11	-0.11	-0.02	0.04		0.05	0.28*	0.24	0.13	0.15
ZET	0.00	0.01	0.08	-0.03	0.06	-0.04	0.48**		0.34*	0.68**	0.59**	0.40**
ALP	-0.13	-0.24	-0.16	-0.23	-0.26	0.11	0.42**	0.43**		0.69**	0.46**	0.72**
BET	-0.08	-0.10	-0.01	-0.09	0.01	-0.04	0.62**	0.84**	0.51**		52**	0.72**
PHY	-0.06	-0.13	-0.02	-0.13	-0.07	-0.03	0.64**	0.89**	0.64**	0.92**		0.34*
LIC	-0.12	-0.27*	-0.17	-0.27	0.30**	0.07	0.48**	0.36**	0.93**	0.51**	0.64**	

<sup>1</sup>Key: LL = leaf length; RL = root length; RW = root weight; TRD and MRD = top and middle root diameter; TDS = total dissolved solids; RC = root color; ZET =  $\zeta$ -carotene; ALP =  $\alpha$ -carotene; BET =  $\beta$ -carotene; PHY = phytoene; and LIC = lycopene.

\*significant by the t-test (n-2 degrees of freedom) at 1%.

\*\*significant by the t-test (n-2 degrees of freedom) at 5%.

correlation between root color and carotene content was 0.28 for the Brasilia x HCM population and 0.78 for the B493 x QAL population. For both of the carrot populations studied by us root color correlations with all the vegetative characters were non-significant suggesting that selection for root color does not have any impact on carrot crop production.

The correlations between  $\zeta$ -carotene,  $\alpha$ -carotene,  $\beta$ -carotene, phytoene and lycopene were significant and positive in both populations (Table 1). The correlation between  $\beta$ -carotene and  $\alpha$ -carotene also indicated that when  $\alpha$ -carotene increased  $\beta$ -carotene also increased and vice versa. Products did not always have their highest correlation with the immediate precursor of the carotenoid pathway (Table 1), *e.g.* for the Brasilia x HCM population, phytoene had the highest correlation with  $\zeta$ -carotene while the highest correlation for  $\zeta$ -carotene was with  $\beta$ -carotene. The highest correlations concordant with substrate and product of the carotenoid pathway were observed in the B493 x QAL population for phytoene x  $\zeta$ -carotene,  $\zeta$ -carotene x  $\beta$ -carotene and lycopene x  $\alpha$ -carotene, while in the Brasilia x HCM population the correlations were  $\zeta$ -carotene x  $\beta$ -carotene, lycopene x  $\beta$ -carotene and lycopene x  $\alpha$ -carotene. These results indicate that the least biochemical steps between substrate and product did not always result in the highest correlation in the carotenoid pathway.

**Causal diagrams, path Student's t-test and sample size**

An indication of the appropriateness of causal diagrams in explaining variation in major carotenoid levels is given by the coefficient of determination ( $R^2$ ) and path significance values. According to Hatcher (1994) it is generally agreed that when  $R^2$  is greater than 60% a relatively large percentage of the variance can be explained by a causal model. In our study,  $R^2 > 71\%$  for the  $\beta$ -carotene dependent variable, indicating that this model explained a considerable portion of the total variance of this dependent variable in both populations. The lycopene  $R^2$  value was 62% for the B493 x QAL population. It had been expected that an explanation accounting for  $R^2$  values  $> 90\%$  would be found since carotenoid accumulation is a direct result of substrate transformation steps in the carotenoid pathway and no alternative pathway has been reported (Cunningham and Gantt, 1998).

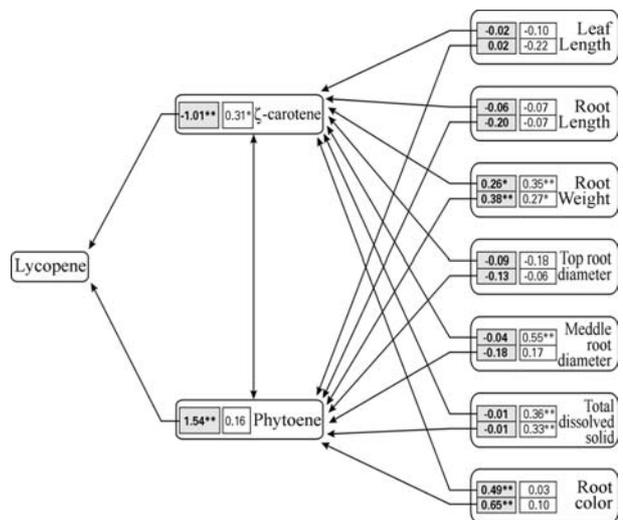
The paths estimated by  $\hat{\beta} = (X'X)^{-1}X'Y$  had similar values to the estimates obtained with the Calis procedure. Approximated t-tests were provided by the Calis procedure and were estimated by dividing the path values with the standard error of the path (Hatcher, 1994). There is no reference in the plant breeding and plant genetics literature which tests the path value (Dewey and Lu, 1959; Li 1975; Samonte *et al.* 1998), so our estimates are probably the first. These tests, derived from path analysis used in social sci-

ences and economics (Hatcher, 1994), will be useful to provide the most likely conclusion for interpreting and applying path analysis. The variance inflation factors (VIF) were less than 10 which according to Chatterjee and Price (1991) indicate an absence of multicollinearity. The ratio between maximum and minimum eigenvalues was greater than 15, indicating the presence of weak collinearity. The overall multicollinearity tests demonstrated absent or weak collinearity effects in the explanatory variable matrix. These values suggest that inferences can reliably be drawn about the fitted causal diagrams. Causal diagrams explaining lycopene and  $\beta$ -carotene variation demonstrated relationships between the carotenoids and morphological traits.

**Causal diagram explaining lycopene variation**

The estimated effect of  $\zeta$ -carotene on lycopene accumulation was a path coefficient ( $\hat{p}$ ) = -1.01 in the B493 x QAL population and  $\hat{p}$  = 0.31 in the Brasilia x HCM population (Figure 1),  $\zeta$ -carotene being the precursor of lycopene in the carotenoid pathway and high positive values were thus expected. The indirect effect of  $\zeta$ -carotene via phytoene was 1.37 (multiplying the correlation value of 0.89 by the phytoene path value of 1.54) in the B493 x QAL population, indicating the important role of  $\zeta$ -carotene in determining lycopene content. These results suggest that  $\zeta$ -carotene is efficiently used by the  $\zeta$ -carotene desaturase involved in the conversion of  $\zeta$ -carotene to the red lycopene.

The direct effects of phytoene on lycopene accumulation were  $\hat{p}$  = 1.54 in the B493 x QAL population and  $\hat{p}$  = 0.16 in the Brasilia x HCM population (Figure 1).



**Figure 1** - Causal diagrams reflect the interrelationships between lycopene and the primary explanatory variables ( $\zeta$ -carotene and phytoene) and the secondary explanatory variables (leaf length, root length, root weight, top and middle root diameter, total dissolved solids and root color) for two  $F_2$  carrot populations (B493 x QAL and Brasilia x HCM). The path coefficient ( $\hat{p}$ ) for cells with a black background are for the B493 x QAL population while a white background shows the values for the Brasilia x HCM population. \*significant by the t-test (n-2 degrees of freedom) at 1%. \*\* significant by the t-test (n-2 degrees of freedom) at 5%.

Phytoene is the first product of the carotenoid pathway and it was expected that  $\zeta$ -carotene, the next product of the pathway and immediate precursor of lycopene, should have the highest direct effect on lycopene accumulation. Our results suggest that lack of phytoene synthesis is the step which limits the carotenoid pathway in white carrot roots and that once formed  $\zeta$ -carotene is efficiently transformed the next metabolite, lycopene, once the pathway is activated by the accumulation of phytoene in orange carrot roots.

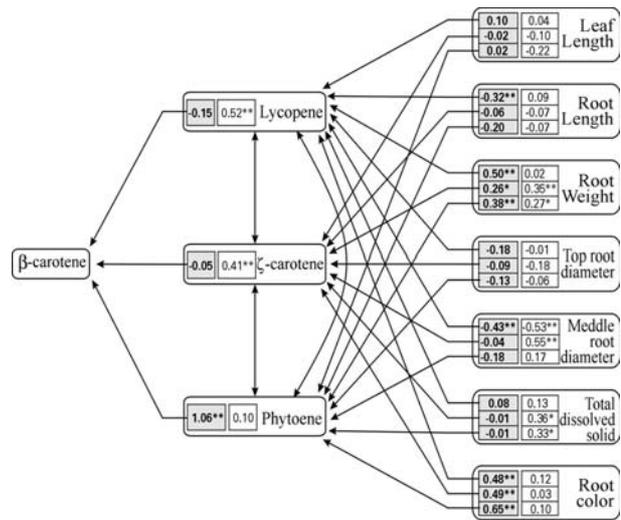
The same lycopene causal model was more efficient in explaining the variation of lycopene in the B493 x QAL population ( $R^2 = 62\%$ ) than in the Brasilia x HCM population ( $R^2 = 18\%$ ). A higher coefficient of determination was expected for the Brasilia x HCM population because lycopene is a direct product of  $\zeta$ -carotene in the carotenoid pathway. In light of the fact that lycopene was always the least plentiful carotene measured in this cross, these results suggest that phytoene is efficiently cyclized to  $\alpha$ - and  $\beta$ -carotene so that the variation in lycopene content we observed was less clearly attributable to the variables we measured.

Significant effects of vegetative traits on phytoene content were observed for root weight and root color in the B493 x QAL population and between root weight and total dissolved solids for the Brasilia x HCM population (Figure 1). In the  $\zeta$ -carotene explanatory diagram, significant direct effects were observed between root weight and root color for the B493 x QAL population and between root weight, middle root diameter and total dissolved solids for the Brasilia x HCM population. These results suggest that carotene accumulation could be increased by indirect selection for root weight and dark orange color in the B493 x QAL (orange x white) population and root weigh, middle root diameter and total dissolved solids in the Brasilia x HCM (orange x dark orange) population.

The coefficients of determination for explanatory variation of phytoene were 46% in the B493 x QAL population and 15% in the Brasilia x HCM population while the explanations for total  $\zeta$ -carotene variation were 26% in the B493 x QAL population and 32% in the Brasilia x HCM population. Once again, the determination of the model should be close to 100 since  $\zeta$ -carotene accumulation is a direct result of substrate transformation of phytoene and no alternative pathway has been reported.

### Causal diagram explaining $\beta$ -carotene variation

The  $\beta$ -carotene causal diagram (Figure 2) identified the correct order phytoene ( $\hat{p} = 0.10$ )  $\rightarrow$   $\zeta$ -carotene ( $\hat{p} = 0.41$ )  $\rightarrow$  lycopene ( $\hat{p} = 0.52$ ) of substrates and products for the Brasilia x HCM population. Lycopene, the substrate for  $\beta$ -carotene, had the highest significant  $\hat{p}$  value of all the variables. The coefficient of determination for this model was around 71%, which is considered to be a reasonable value.



**Figure 2** - Causal diagrams reflect the interrelationships between  $\beta$ -carotene and the primary explanatory variables (lycopene,  $\zeta$ -carotene and phytoene) and the secondary explanatory variables (leaf length, root length, root weight, top and middle root diameter, total dissolved solids and root color) for two  $F_2$  carrot populations (B493 x QAL and Brasilia x HCM). The path coefficient ( $\hat{p}$ ) for cells with a black background are for the B493 x QAL population while a white background shows the values for the Brasilia x HCM  $F_2$  population. significant by the t-test (n-2 degrees of freedom) at 1%. \*\* significant by the t-test (n-2 degrees of freedom) at 5%.

In the B493 x QAL population no significant  $\hat{p}$  values were observed for  $\zeta$ -carotene ( $\hat{p} = -0.05$ ) and lycopene ( $\hat{p} = -0.15$ ), the highest  $\hat{p}$  value being  $\hat{p} = 1.06$  for phytoene (Figure 2). There are two possible explanations for these results: 1) misidentification of lycopene and  $\zeta$ -carotene in the HPLC chromatograms and 2) phytoene is a key substrate necessary for  $\beta$ -carotene production in white carrots. The second hypothesis is the most likely because it has been shown (Ye *et al.*, 2000) that  $\beta$ -carotene can be produced in transformed rice with other carotenoid biosynthetic enzyme when beta lycopene cyclase is not present. The  $\zeta$ -carotene retention time and absorption spectrum were consistent with previous studies, so it appears that misidentification did not occur. Via phytoene, both  $\zeta$ -carotene and lycopene had large indirect effects on  $\beta$ -carotene production, supporting the hypothesis that phytoene is the key substrate blocking the carotenoid pathway. The path diagram for lycopene accumulation in the  $\beta$ -carotene model only used significant effects for the vegetative characters root length, root weight, middle root diameter and RCO (Figure 2), the coefficient of determination for lycopene in this causal diagram being 39% which indicates that lycopene and  $\beta$ -carotene content could be increased by selecting for these vegetative traits.

### Discussion

Path analysis was first applied to genetics by Sewall Wright (1921), one of the founders of quantitative genetics,

to explain variation in guinea pigs. In this type of analysis a network of causes and effects is seen as a series of steps in a path with a coefficient assigned to each step to quantify interrelationships (Wright, 1921). The first application of path analysis to plant breeding occurred in 1959 when Dewey and Lu (1959) analyzed crested wheatgrass seed production. Path analysis has also been applied widely in sociology, economics and psychology (Lynch and Walsh, 1998; Hatcher, 1994), with most of the recent improvements in path analysis having occurred in these disciplines.

This work described in our paper is the first report applying path analysis to the dissection of a biosynthetic pathway. The identification of key steps in a given pathway could help identify genetic transformation events which could be used to manipulate key early biosynthetic steps which limit the accumulation of final products.

The results obtained with path analyses on the carrot carotenoid pathway support the hypothesis that phytoene is a key substrate necessary for lycopene and  $\beta$ -carotene production in white carrots and that when phytoene is produced it is very efficiently converted through desaturation and cyclization to  $\alpha$ - and  $\beta$ -carotene. These results agree with the findings of Ye *et al.* (2000) who showed that in the biosynthesis of  $\beta$ -carotene in rice transformed with genes coding carotenoid biosynthetic enzymes lycopene cyclized in the absence of beta lycopene cyclase, suggesting that some 'endogenous' lycopene beta cyclase activity was already present in the transformed rice plants. Based on the fact that in our experiments phytoene had the highest positive  $\hat{p}$  values in the casual diagrams for lycopene and  $\beta$ -carotene accumulation in the B493 x QAL (orange x white) F<sub>2</sub> carrot population we hypothesize that phytoene accumulation is the key step blocking carotenoid production in white carrot roots.

Although it is well known that carotenoids protect chlorophylls from photo-oxidation and are essential light-harvesting pigments and photoreceptors, it is difficult to explain their function in storage roots. Transformation of white carrot with a root-specific phytoene synthase may provide a test for the hypothesis that phytoene synthase present in the orange background is the key enzyme blocking the carotenoid pathway in the white carrot. Assuming that this hypothesis is correct, path analysis, as originally presented by Sewall Wright (1921) to study quantitative traits, appears to be a powerful statistical approach to identify key components of complex pathways and has the potential to be applied to the elucidation of other biochemical pathways and engineer their manipulation.

## Acknowledgments

The authors are thankful to José Cletis Bezerra (Embrapa, Petrolina, PE, Brazil) for his skillful help with the drawing the path diagram.

## References

- Armstrong GA (1994) Eubacteria show their true colors: Genetics of carotenoid pigment biosynthesis from microbes to plants. *J Bacteriol* 176:4795-4802.
- Bhatt GM (1973) Significance of path coefficient analysis determining the nature of character association. *Euphytica* 22:338-343.
- Buishand JG and Gabelman WH (1979) Investigations of color and carotenoid content in phloem and xylem of carrot roots (*Daucus carota* L.). *Euphytica* 28:611-632.
- Chatterjee S and Price B (1991) Regression analysis by example. 2nd edition. John Wiley and Sons, New York, 278 pp.
- Cruz CD and Regazzi AJ (1997) Modelos Biométricos Aplicados ao Melhoramento Genético. Editora UFV, Vicososa, 390 pp.
- Cruz CD (1998) Genes - Software for experimental statistics in genetics. *Genet Mol Biol* 21:135-138.
- Cunningham FX and Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. *Annu Rev Plant Physiol Plant Mol Biol* 49:557-583.
- Dewey DR and Lu KH (1959) A correlation and path-coefficient analysis of components of crested wheatgrass seed production. *Agron J* 51:515-518.
- Emsweller SL, Burrell PC and Borthwick HA (1935) Studies on the inheritance of color in carrots. *J Am Soc Hortic Sci* 33:508-511.
- Giuliano G, Aquilani R and Dharmapuri S (2000) Metabolic engineering of plant carotenoids. *Trends Plant Sci* 5:406-409.
- Hamerschmidt I (1993) Produção de hortaliças e assistência técnica no Brasil. *Hortic bras* 11:156-157.
- Hatcher L (1994) A step-by-step approach to using the SAS system for factor analysis and structural equation modeling. SAS Institute Inc., Cary, NC, 588 pp.
- Khachik F, Beecher GR, Goli MB and Lusby WR (1992) Separation and quantification of carotenoids in foods. *Meth Enzymol* 213:347-359.
- Krurup A and Mosnaim SR (1980) Some parameters for evaluating quality in ten carrot (*Daucus carota* L.) cultivars grown in Valdivia. *Agroicultura de Sur* 8:5-9.
- Li CC (1956) The concept of path coefficient and its impact on population genetics. *Biometrics* 12:190-210.
- Li CC (1975) Path Analysis - A Primer. The Boxwood Press, Pacific Grove, 346 pp.
- Lynch M and Walsh B. (1998) Genetics and Analysis of Quantitative Traits. Sinauer Associates, Inc. Publishers, Sunderland, 980 pp.
- Natarajan S and Arumugam R (1980) Association analysis of yield and its components in carrot (*Daucus carota* L.). *Madras Agricultural Journal* 67:594-597.
- Pariari A, Maity TK and Gayen P (1992) Association of physical characters with yield and quality in carrot. *Annals of Agricultural Research* 13:96-98.
- Randhir S, Sukhija BS, Hundal JS and Singh R (1992) Correlation and path coefficient studies in carrot. *Journal of Research for the Punjab Agricultural University* 26:581-584.
- Rubatzky VE, Quiros CF and Simon PW (1999) Carrots and Related Vegetable Umbelliferae. CABI Publishing, Wallingford, 294 pp.
- Sandmann G (1998) Carotenoid biosynthesis in microorganisms and plants. *Eur J Biochem* 223:7-24.

- Samonte SOPB, Wilson LT and McClung AM (1998) Path analysis of yield and yield-related traits of fifteen diverse rice genotypes. *Crop Sci* 38:1130-1136.
- SAS Institute Inc. (1989) SAS/STAT User's Guide. Ver 6, 4th edition. SAS Institute Inc. Cary, NC, 889 pp.
- Simon PW (1992) Genetic improvement of vegetable carotene content. In: Bills DD and Kung S (eds) *Biotechnology and Nutrition*. Butterworth-Heinemann, Boston, pp 291-300.
- Simon PW and Wolff XY (1987) Carotenes in typical and dark orange carrots. *J Agric Food Chem* 35:1017-1022.
- Simon PW (2000) Domestication, historical development, and modern breeding of carrot. *Plant Breed Rev* 19:157-190.
- Simon PW, Peterson CE and Gabelman WH (1990) B493 and B9304, carrot inbreds for use in breeding, genetics, and tissue culture. *HortScience* 25:815.
- Simon PW, Wolff XY, Peterson CE, Kammerlohr DS, Rubatzky VE, Strandberg JO, Bassett MJ and White JM (1989) High Carotene Mass carrot population. *HortScience* 24:174.
- Simpson KL (1983) Relative value of carotenoids as precursors of vitamin. *A Proc Nutr Soc* 42:7-17.
- Stommel JR and Simon PW (1989) Phenotypic recurrent selection and heritability estimates for total dissolved solids and sugar type in carrot. *J Am Soc Hort Sci* 114:695-699.
- Wright S (1921) Correlation and causation. *J Agric Res* 20:557-585.
- Ye X, Salim AB, Kloti A, Jing Z, Lucca P, Beyer P and Potrykus I (2000) Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* 287:303-305.

*Associate Editor: Everaldo Gonçalves de Barros*