Theor Appl Genet (2004) 109:440–449
DOI 10.1007/s00122-004-1651-7

ORIGINAL PAPER

B. Quilot · M. Génard · J. Kervella · F. Lescourret

Analysis of genotypic variation in fruit flesh total sugar content via an ecophysiological model applied to peach

Received: 11 August 2003 / Accepted: 23 February 2004 / Published online: 17 April 2004

Abstract A simulation model of the evolution of total sugar content (C_{TS}) in fruit was developed in order to describe the within- and between-genotype variation of C_{TS} observed in a peach (*Prunus persica* (L.) Batsch) breeding population. The parameter k defines the ratio of carbon used for synthesizing compounds other than sugars for each genotype. Model input variables are dry flesh growth rate and fresh flesh mass of fruit. We estimated k for 137 peach and nectarine genotypes derived from a clone of a wild peach (*Prunus davidiana*) by three generations of crosses with commercial nectarine varieties. We tested the predictive quality of the model on independent datasets. Despite an underestimation of the observed C_{TS}, the correlation between observations and predictions was suitable (0.72). Spearman correlation coefficients between 2001 and 2002 for model input variables and parameter k were higher than for C_{TS}. None of the three components assimilation supply to the fruit, metabolism, or dilution, seemed to have a greater relative effect on C_{TS} variation than the others. Indeed, C_{TS} variation seemed to result from the balance between the three components. The interest of this approach, which consists of dissecting traits into components via an ecophysiological model, for breeding strategy and for sugar accumulation studies are discussed.

Communicated by H.C. Becker

B. Quilot (✉) · M. Génard · F. Lescourret
Institut National de la Recherche Agronomique (INRA),
Plantes et Systèmes de Culture Horticoles,
Domaine St Paul, Site Agroparc, 84914 Avignon Cedex 9, France
e-mail: quilot@avignon.inra.fr
Tel.: +33-432-722473
Fax: +33-432-722432

J. Kervella
Institut National de la Recherche Agronomique (INRA),
Génétique et Amélioration des Fruits et Légumes,
Domaine St Paul, Site Agroparc, 84914 Avignon Cedex 9, France

Introduction

Fruit sweetness depends on total sugar content (C_{TS}) in the flesh (Leonard et al. 1953; Robertson et al. 1992). Wide variations in C_{TS} have been reported between genotypes of different level of selection in different species, such as *Cucumis melo* (Stepansky et al. 1999), *Prunus armeniaca* (Gurrieri et al. 2001) and *P. persica* (Moriguchi et al. 1990). Improvement of C_{TS} through selection has been restricted (Dirlewanger et al. 1999; Saliba-Colombani et al. 2001; Hashizume et al. 2003). For many fruit species, few QTL controlling gustative fruit quality have been mapped (Abbott et al. 1998; Quarta et al. 1998; Causse et al. 2001) and genes controlling QTL for gustative fruit quality often remain unknown (Saliba-Colombani et al. 2001; Etienne et al. 2002). Various reasons can be evoked to explain this. First, few studies have been completed on gustative fruit quality, since conversely to size, firmness and appearance, gustative fruit quality has been of major economic interest only recently. Secondly, genetic variation in gustative fruit quality traits between cultivars is limited. Indeed for some species, cultivars display a narrow genetic base (Ladizinsky 1985; Reynders and Monet 1987; Byrne 2002). To overcome the lack of genetic variation interspecific transgression can be exploited. Alleles with favorable effects on many traits have been found in related species for improving tomato (de Vicente and Tanksley 1993; Tanksley et al. 1996; Fulton et al. 1997; Bernacchi et al. 1998), and rice (Xiao et al. 1998), despite a generally low agronomic level. Third, sugar content is under the influence of environmental factors. These environmental factors are yet to be identified and taken into account, and better knowledge about the way they influence phenotype is needed.

For total sugar content, variation between trees, between fruits of the same tree, and between years are not negligible in comparison with the variation between genotypes. Sugar content varies throughout fruit development according to the supply of carbohydrates to the fruit, changes in fruit metabolism, and dilution caused by

increase in fruit volume. These three processes are more or less affected by environmental factors. Microclimatic gradients (Corelli-Grappadelli and Coston 1991; Marini et al. 1991), leaf area around the fruit (Kliewer and Weaver 1971; Génard 1992) and the vigor of fruit-bearing shoots (Génard and Bruchou 1992) are frequently mentioned for causing within-plant variation. These factors have a major effect on the supply of carbohydrates to the fruit so that a strong positive correlation appears between the sugar content and the size of fruit within the same plant (Génard et al. 1991). Concerning metabolic transformations, the main enzymatic reactions responsible for sugar synthesis and transformation have been identified for several fleshy fruits (Black et al. 1987; Hubbard et al. 1991; Ho 1996), but little is known about their control by environmental factors. The effect of dilution has only been studied through the effect of irrigation on fruit quality. Sugar content usually decreases in response to enhanced irrigation (Li et al. 1989; Crisosto et al. 1994).

To deal with the major influence of environmental factors on total sugar content, it is helpful to consider dry and fresh fruit growth. These two ecophysiological variables influence total sugar content through assimilate supply and dilution and furthermore, integrate the main effects of environmental factors. A valuable way to describe total sugar content via these two variables is to use a modeling approach. Indeed, ecophysiological modeling makes it possible to describe the elaboration of a trait under environmental control and plant regulation principles. Recently, ecophysiological models have been of great interest in many disciplines (Shorter et al. 1991; Hammer et al. 1996; Boote et al. 1996) and future opportunities for their use are numerous (Boote et al. 2001; Tardieu 2003).

The aim of this study was to propose a model to describe the variation of total sugar content (C_{TS}) in the fruit flesh within a breeding population. The model had to be simple enough to be compatible with the quantitative genetic experimental constraints. Numerous genotypes had to be studied and a limited number of observations could be made for each of them. Consequently, we developed a simple model with only one genotype-dependent parameter, simulating C_{TS} evolution during the main stage of fruit enlargement. The model was used to simulate the genetic variation of C_{TS} for genotypes belonging to progenies derived from an interspecific cross (*persica* × *davidiana*). The model was calibrated and tested on distinct databases. The model was used to determine the relative contribution of three components: (1) assimilate supply to the fruit, (2) metabolism, and (3) dilution caused by change in fruit volume, on the variation of C_{TS}. Results of this study were considered from both the practical and fundamental viewpoints in order to discuss the interest of such an approach for selection and biological understanding.

Materials and methods

Description of the model

The model predicts the evolution of total sugar content in the flesh during fruit growth, based on the relative contributions of assimilate supply, metabolic transformations and fruit volume on sugar content. It is a simplified form of the SUGAR model developed by Génard and Souty (1996), that predicts the partitioning of carbon into sucrose, sorbitol, glucose and fructose in the flesh. It was not possible to use the SUGAR model because it requires too many observations to be calibrated for each of the numerous genotypes in a breeding population.

The model is based on carbon balance. Carbon arrives into the fruit as sugars, via the phloem. In the flesh, part of this flow of carbon is used as substrates for respiratory pathways. The remaining carbon is used partly for sugar synthesis and partly for synthesis of other carbohydrate compounds (e.g., starch, acids, structural carbohydrates, and proteins).

Accordingly, the model is defined by the following differential equation:

$$\frac{dM_{TS}}{dt} = c_{fl}\frac{dM_{dry}}{dt} - kM_{TS} \quad (1)$$

where M_{TS} is the total amount of carbon (g) in the fruit flesh as sugars, c_{fl} is the carbon concentration of the mesocarp (g C per gram of dry mass) that is assumed to be constant during the final stage of growth. dM_{dry}/dt is the dry flesh growth rate (g day^{-1}) and k (day^{-1}) is the relative rate of consumption of carbon as sugars in the fruit flesh for synthesis of compounds other than sugars. For simplification, k was considered independent of environmental factors.

The total sugar content, C_{TS} [g(100 g_{FM})$^{-1}$], is computed as:

$$C_{TS} = \frac{100M_{TS}}{\sigma_{TS}M_{fresh}} \quad (2)$$

where σ_{TS} is the mean carbon content of sugars (g C/g sugars) and M_{fresh} the flesh fresh mass (g).

Differentiation of Eq. 2 leads to:

$$\frac{dC_{TS}}{dt} = \frac{100}{\sigma_{TS}M_{fresh}}\frac{dM_{TS}}{dt} - \frac{100M_{TS}}{\sigma_{TS}M_{fresh}^2}\frac{dM_{fresh}}{dt} \quad (3)$$

We used the approach of Génard et al. (2003) to isolate the three components causing changes in the total sugar content: assimilate supply, metabolic transformation of carbon into compounds other than sugars, and dilution attributable to the change in fruit volume resulting from water uptake. Equation 3 was combined with Eqs. 1 and 2:

$$\frac{dC_{TS}}{dt} = \frac{100c_{fl}}{\sigma_{TS}M_{fresh}}\frac{dM_{dry}}{dt} - kC_{TS} - \frac{C_{TS}}{M_{fresh}}\frac{dM_{fresh}}{dt} \quad (4)$$

At maturity, total sugar content in the flesh is expressed as the integral of Eq. 4:

$$C_{TS} - C_{TS}^{ini} = S - M - D \quad (5)$$

with

$$S = \int_{ini}^{maturity} \frac{100c_{fl}}{\sigma_{TS}M_{fresh}}\frac{dM_{dry}}{dt}, \quad M = \int_{ini}^{maturity} kC_{TS}$$

and

$$D = \int_{ini}^{maturity} \frac{C_{TS}}{M_{fresh}}\frac{dM_{fresh}}{dt}$$

where C_{TS}^{ini} is the initial total sugar content at the beginning of the monitored period. The three integrated components of Eq. 5 will

then be called *S* for assimilate supply, *M* for metabolic transformations and *D* for dilution in the following text.

Plant material

The breeding population was derived from P1908 as follows (Pascal et al. 1998). In the first generation, P1908 was crossed with *P. persica* cv. 'Summergrand' (SG). An F1 progeny was obtained. In the second generation, one F1 hybrid resistant to powdery mildew was back-crossed to SG to produce a BC1 progeny. In the third generation, BC1 individuals were used to pollinate *P. persica* cv. 'Zéphyr' (ZE) to derive the breeding population (BC2). SG and ZE are yellow and white nectarine cultivars respectively, with large tasty fruit. In addition to providing novel genetic variation to a species that displays a narrow genetic base, the use of a wild species as a progenitor should enhance the variation of the studied traits and highlight the processes responsible for the main variation observed.

The study was conducted in two orchards, in the St Paul and the Garrigues sites, at the INRA Avignon Research Center (France). Genotypes were planted in the St Paul orchard in a completely randomized design with one tree per genotype, except for eight genotypes and the three parents (SG, ZE and P1908) that were planted with more replicates. Trees were 3 years old in 2001. Trees of all the genotypes were also available in a collection orchard in the Garrigues site. All genotypes in both sites were grafted on GF305 seedling rootstocks and were grown under optimal conditions of irrigation, fertilization and pest control.

Experiments on the breeding population

This study was carried out in St Paul in 2001 on 100 genotypes of the BC2 population, SG and ZE (BC201 database), and in 2002 on 139 genotypes of the BC2 population, SG, ZE and P1908 (BC202 database). Eighty-seven genotypes of BC2 were common to both years.

A very light-loading level treatment consisting of only five fruits per tree was applied to each tree, to place all fruits in non-limiting source conditions, i.e., in potential fruit growth conditions. We monitored five fruits per tree and per genotype, measuring fruit cheek diameter once a week from the end of May (about 590 degree-days after full bloom) to fruit maturity. The five fruits per tree were harvested at maturity. Fruits were considered ripe on the trees when they no longer grew, softened, and were easily picked.

The fresh flesh mass (M_{fresh}) was determined immediately after harvest. Fruit flesh was cut in small pieces. The dry flesh mass (M_{dry}) was determined using flesh pieces dried to constant weight at 70°C. For three of the five monitored fruits, some flesh pieces were immediately held in the freezer (−80°C) until sugar analysis.

To compute fruit dry mass and fresh flesh growth on the monitored fruits, relationships between fruit cheek diameter and fruit dry and fresh masses were needed for each genotype. To establish these relationships, we recorded fruit cheek diameter and fresh mass, then subjected flesh pieces to a temperature of 70°C for 72 h, and then measured the dry mass content. These measurements were carried out in 2001 and 2002: (1) at maturity for the five fruits of each tree, (2) at thinning on the fruits removed and (3) throughout the fruit growth period on fruits sampled from the trees in the Garrigues site. M_{fresh} and M_{dry} were calculated for each fruit monitored, at each date of diameter measurement. Since the model requires daily input values of M_{fresh} and dM_{dry}/dt, M_{fresh} and M_{dry} were extrapolated to daily data by local regression (Chambers and Hastie 1992) and dM_{dry}/dt calculated by derivation of daily M_{dry}.

Experiments on genotypes with tree replications

We also performed harvests during fruit growth in order to follow sugar accumulation in the flesh on a few replicated genotypes: 10 BC2 genotypes (five at the St Paul site, five at the Garrigues site) and SG in 2001 (kin01 database) and eight BC2 genotypes, SG, ZE and P1908 in 2002 at the St Paul site (kin02 database). For these experiments, trees sustained thinning as for commercial fruit production.

Fruits were sampled every 3–8 days from the end of May to fruit maturity. Three fruits per genotype were harvested at each sampling date for M_{fresh}, M_{dry} and sugar content measurements, as described above.

M_{fresh} and M_{dry} measurements at each sampling date were used for each genotype to compute, by local regression, a mean input daily value of M_{fresh} and M_{dry} by genotype. dM_{dry}/dt was then calculated by derivation of daily M_{dry}.

Chemical analysis

Fruit flesh samples, stored in the freezer, were immersed in liquid nitrogen and then immediately powdered in a stainless steel Dangoumeau grinder for 2 min. Five grams of the powder was mixed for 5 min with 20 ml of ultra-pure water. The mixture was centrifuged at 15,000 *g* for 15 min at 4°C. The supernatant was recovered and immediately filtered through a Waters C18 cartridge (Waters) to eliminate any interfering apolar residues and through a 0.45 μm Sep-Pak filter (Jasco France) to eliminate large particles. The extract, stored at −80°C (sealed tube), was then ready for sugar measurement by HPLC. More details on the procedure can be found in Gomez et al. (2002). Sample sugar contents were established using external standards and expressed in $g(100\ g_{FM})^{-1}$.

Calibration, test and sensitivity analysis of the model

The model was first calibrated for each of the studied genotypes. This consisted of estimating the *k* value that minimized the sum of squared differences between the three predicted and observed C_{TS} values at maturity for each genotype. In the case of non-linear models, the use of an iterative procedure is usually required to estimate the parameters (Huet et al. 1996). The nls function (Splus software, MathSoft Inc., Cambridge, Mass., USA), based on the Gauss-Newton algorithm, was used. The method is described by Chambers and Hastie (1992). We used the BC202 and BC201 databases separately to calibrate the model in order to test the stability of the *k* values from one year to another for the 87 common genotypes.

The predictive quality of the calibrated model was then tested on data independent from those used for the model calibration. The *k* value estimated from the BC202 database and input variables from the BC201 database were used to perform model simulations. Since inputs were available for each monitored fruit, the C_{TS} of each monitored fruit were predicted. C_{TS} model predictions at maturity were then compared to observations from the BC201 database. The same procedure was used to test the predictive quality of the model under different conditions of fruit growth. Model predictions were compared to observations from the kin01 and kin02 databases at maturity and during fruit growth. In this case, only mean fruit input values were available at each date for each genotype. Consequently, a mean C_{TS} per genotype was predicted throughout growth.

Goodness-of-fit criteria were computed to evaluate (1) the goodness-of-fit of the model for each genotype on the basis of data used for the estimation of *k*, and (2) the predictive quality of the model, for independent data (see above). The criterion adopted was the root mean squared error (RMSE), a common criterion used to quantify the mean difference between simulation and measurement in case of non-linear models (Kobayashi and Us Salam 2000), here defined as:

$$\mathrm{RMSE}(i) = \sqrt{\frac{\sum_{j=1}^{N} (x_{ij} - y_{ij})^2}{N}}$$

where x_{ij} is the observed *j* from genotype *i*, y_{ij} the corresponding simulation result, and *N* the total number of observed data from

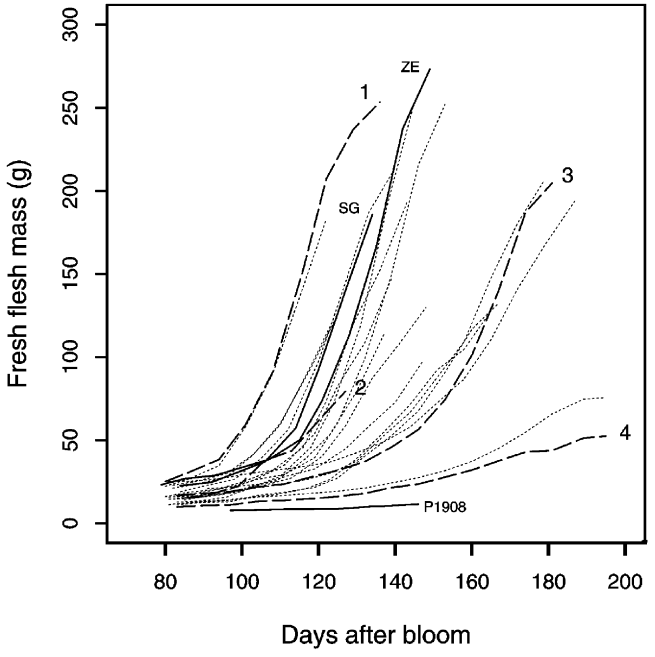

Fig. 1 Seasonal variation in fresh flesh mass (g) for 20 BC2 genotypes with contrasting growth patterns and the three parents. Parent patterns are in *solid lines* and indicated by P1908, SG and ZE. Highly contrasting growth patterns of four BC2 genotypes are in *dashed bold lines* and are indicated by 1, 2, 3 and 4

genotype *i*. The smaller the RMSE in comparison with the measurements, the better the goodness-of-fit, which can be represented through the relative RMSE (RRMSE):

$$\mathrm{RRMSE}(i) = \mathrm{RMSE}(i)/\bar{Y}_i,$$

where $\bar{Y}_i = \sum_{i=1}^{N} (x_{ij})/N$ is the mean of all observed values from genotype *i*.

The goodness-of-fit of the model for the population is computed by averaging the RRMSE values of its genotypes.

Simulations were performed to study the sensitivity of the model response to variation of both the *k* value and the daily input variables, dM_{dry}/dt and M_{fresh}. The sensitivity criterion was the difference between total sugar content [$g(100\ g_{FM})^{-1}$] for high (C_H) and low (C_L) values of *k* and input variables, expressed as a percentage of total sugar content at harvest for the default (C_o) *k* value and input variables: $100\times(C_H - C_L)/(C_o)$. Default values of *k* ($0.045\ day^{-1}$) and of daily input variables were those of the 'Summergrand' cultivar. High and low values of *k* and input variables were set to ±20% of those of the 'Summergrand' cultivar.

Similarly, we analyzed the model response to combined variations of growth pattern and *k* values. Four highly contrasting growth patterns: (1) short and high, (2) short and low, (3) long and high and (4) long and low growth (Fig. 1), observed in the BC2 population were used as model inputs. *k* values were fixed to the extreme and mean *k* values observed in this population.

All data analyses were performed with the Splus language (Splus software, MathSoft Inc., Cambridge, Mass., USA). Spearman's rank correlations between variables were calculated using the "COR" procedure of Splus.

Results

Phenotypic variation in total sugar content and input variables

The experimental measurements revealed considerable variation of C_{TS} at maturity between genotypes (Table 1). There were large differences in C_{TS} between the three parents. Some genotypes of the population showed greater levels of C_{TS} than the parents. Moreover, genotypes of the BC2 with the lowest mean content in 2001 or 2002 showed a higher level than P1908 in 2002. Distributions of the mean contents of the population were significantly different in 2001 and 2002 but the extreme levels of mean contents were comparable (Table 1). The Spearman correlation coefficient between the mean contents observed in 2001 and 2002 for the 87 genotypes common to both years was slight (0.31, Spearman's test *p* value = 0.0045).

Large differences between genotypes were also observed for the fresh and dry flesh masses. These differences concerned both the levels of fresh and dry flesh masses reached at maturity (Table 1) and the kinetics of accumulation of fresh and dry matter (Fig. 1). SG and ZE exhibited very high fruit growth during the period monitored, whereas P1908 fruits had enlarged before the beginning of the period monitored and reached very small masses. Many genotypes of the BC2 population carried fruits larger than those of SG at maturity. The mean fruit mass of some even reached the mean mass of ZE fruits. For mean dry and fresh flesh masses, the Spearman correlation coefficients between 2001 and 2002 reached 0.43 and 0.49 respectively for the 87 genotypes common to both years. The variations in rate and duration of growth within the BC2 population were higher than those be-

Table 1 Mean, standard deviation (between parentheses), minimal and maximal values of total sugar content in the flesh (C_{TS}), and of fresh and dry flesh mass at maturity for the BC2 population and the three parents (P1908, SG and ZE) in 2001 and 2002. For each genotype, data from two to five fruits were averaged. One hundred and 139 genotypes of the BC2 population were studied in 2001 and 2002, respectively

	C_{TS} [$g(100\ g_{FM})^{-1}$]		Fresh flesh mass (g)		Dry flesh mass (g)	
	2001	2002	2001	2002	2001	2002
BC2						
Min	6.02	5.63	24.61	52.38	4.99	7.35
Mean	10.89 (2.95)	9.23 (1.86)	87.90 (34.54)	139.84 (49.13)	15.48 (5.86)	21.07 (8.33)
Max	20.36	16.73	185.1	267.43	33.00	52.06
P1908						
Mean	–	2.46 (0.55)	–	10.43 (0.27)	–	1.48 (0. 16)
SG						
Mean	7.54 (0.32)	8.209 (0.26)	68.54 (3.20)	179.28 (22.03)	9.75 (0.48)	25.56 (4.72)
ZE						
Mean	10.34 (1.82)	10.63 (1.27)	95.32 (14.11)	282.24 (55.04)	12.97 (2.32)	43.32 (8.29)

Table 2 Mean, standard deviation (between parentheses), minimal and maximal values of k and the standard error of estimation, estimated from the BC202 database, for the BC2 population and the three parents (P1908, SG and ZE)

	k (day^{-1})	Standard error (day^{-1})
BC2		
Min	0.00319	0.00017
Mean	0.03294 (0.0149)	0.00667 (0.00662)
Max	0.07673	0.04192
P1908		
Mean	0.12099	0.04224
SG		
Mean	0.04523	0.00753
ZE		
Mean	0.01819	0.00402

tween the parents. The harvest dates ranged from the end of June to the end of September, depending on the genotype. All the combinations were present in the population from short and high growth to long and slow growth (Fig. 1). Spearman correlation coefficients between 2001 and 2002 reached 0.67 and 0.95, for mean growth rate of dry flesh (g day^{-1}) and mean growth duration (days), respectively.

Calibration and sensitivity analysis of the model

The k values estimated from the BC202 database ranged from 0.0032 to 0.0767 day^{-1} in the BC2 population (Table 2). The k values of SG and ZE were intermediate but the k value for P1980 was much higher than all the other estimated k values. The observed genotypic variations of C_{TS} were reproduced well by the model (Fig. 2a). The model also reproduced the variation between the fruits of the same genotype (Fig. 2b). The values of the goodness-of-fit criteria (RRMSE), calculated on individual fruit data, ranged from 0.004 to 0.492 (Table 3) depending on the genotype.

In the sensitivity analysis, the model appeared to be very sensitive to both the k parameter and the input variables. A 40% variation in the k value resulted in a 17% variation in C_{TS} at maturity. In the same way, a 40% variation in the input variables, dM_{dry}/dt and M_{fresh}, resulted in 40 and 41% variations in the model output, respectively.

Combined effect of fruit growth pattern and k value on simulated total sugar content at maturity

We used the model to illustrate the effect of fruit growth pattern and k value on predicted C_{TS}. Twelve combinations of growth pattern and k value were compared (Table 2). For a given growth pattern, the higher the k value, the smaller the predicted C_{TS} at maturity. However, strong interactions between k values and growth patterns were observed (Table 4). For the low k value, C_{TS} simulations were highest in the case of short and high growth.

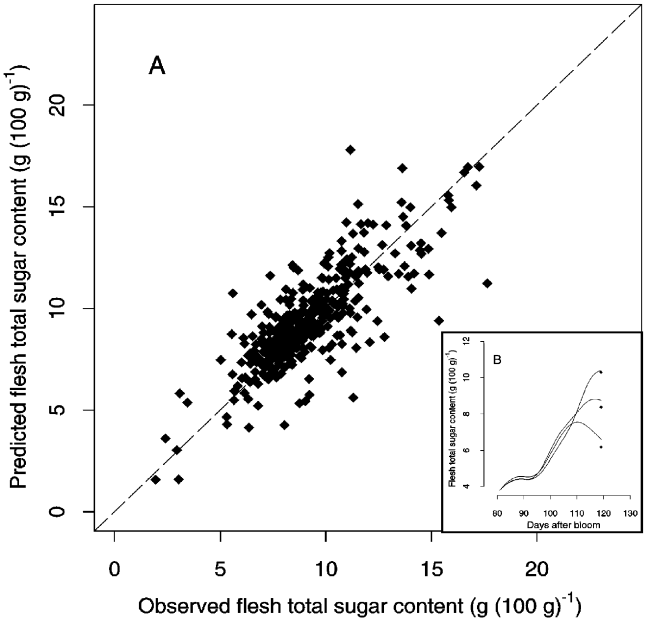

Fig. 2a–b Predicted values of flesh total sugar content [$g(100\ g_{FM})^{-1}$] for each BC2 genotype are plotted against the observed values (**a**). Three repetitions for each of the 139 genotypes are plotted. Within-plant variation in flesh total sugar content [$g(100\ g_{FM})^{-1}$] are presented for one BC2 genotype, *lines* representing the predicted seasonal variation in flesh total sugar content for three fruits of one tree and *points* corresponding to the observations at maturity (**b**)

Table 3 Estimated values of the relative mean squared error (RRMSE) for evaluating the calibration of the model and its predictive quality. Mean, standard deviation (between parentheses), minimal and maximal values of RRMSE are presented for the BC2 population and the three parents (P1908, SG and ZE)

	Calibration (maturity)		Test (during fruit growth)	
Database	BC202	BC201	kin01	kin02
Number of BC2 genotypes	139	87	9	8
BC2				
Min	0.004	0.036	0.120	0.190
Mean	0.121 (0.093)	0.240 (0.142)	0.397 (0.304)	0.282 (0.088)
Max	0.492	0.734	1.114	0.466
P1908				
Mean	0.449	–	–	0.596
SG				
Mean	0.093	0.194	0.249	0.185
ZE				
Mean	0.141	0.127	–	0.569

The rank of growth pattern was 1>4>3≥2. On the contrary, for the mean k value, the rank was 1>2>4≥3. Again, for the high k value, the rank was completely modified: 2>1>3≥4.

Table 4 Response of total sugar contents in the flesh [g(100 g_{FM})$^{-1}$] and C_{TS} to variations in the *k* values in the range of the *k* values observed within the BC2 population, and to variations in growth pattern. *k* values were taken as equal to the minimal, mean and maximal values estimated for the BC2 population (Table 2). Input variables were those of four BC2 genotypes with highly contrasting growth patterns, represented in Fig. 3

	Short growth		Long growth	
	High	Low	High	Low
Genotype	1	2	3	4
k value				
Minimal	19.29	12.39	12.88	16.09
Mean	10.39	8.67	5.60	5.61
Maximal	4.60	5.85	2.56	2.15

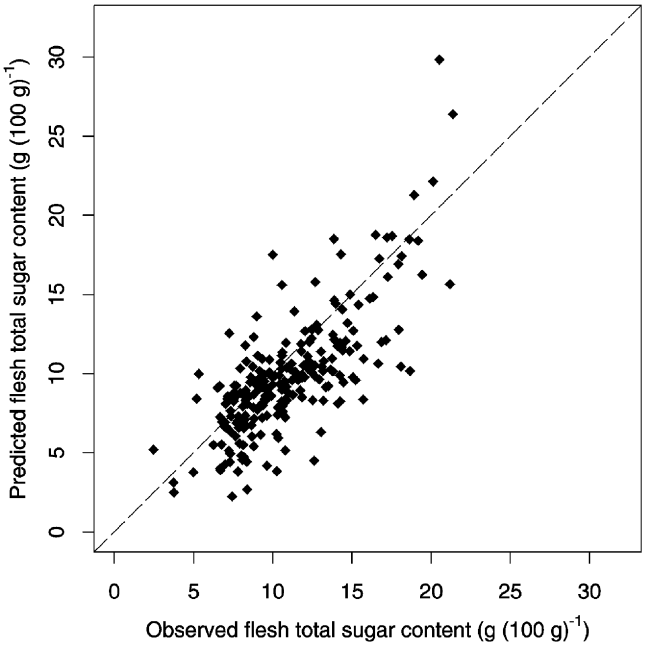

Fig. 3 Test of the predictive quality of the model at maturity, calibrated from the BC202 database, on the BC201 database. The predicted values of flesh total sugar content [g(100 g_{FM})$^{-1}$] are plotted against the corresponding observed values. Three repetitions for each of the 87 genotypes common to both years are plotted

Test of the model

The model calibrated with the BC202 database was used to predict C_{TS} at maturity for the 87 genotypes of BC201, SG and ZE present in the BC202 database. The model predictions ranked the genotypes accurately according to an average C_{TS} per genotype. The correlation between observations and predictions reached 0.72, although the predictions were often slightly biased towards underestimated values (Fig. 3). The RRMSE values were acceptable for most genotypes. For the population, they reached a maximal value of 0.73 (Table 3). They were larger than 0.50 for only four genotypes and lower than 0.2 for 39 genotypes.

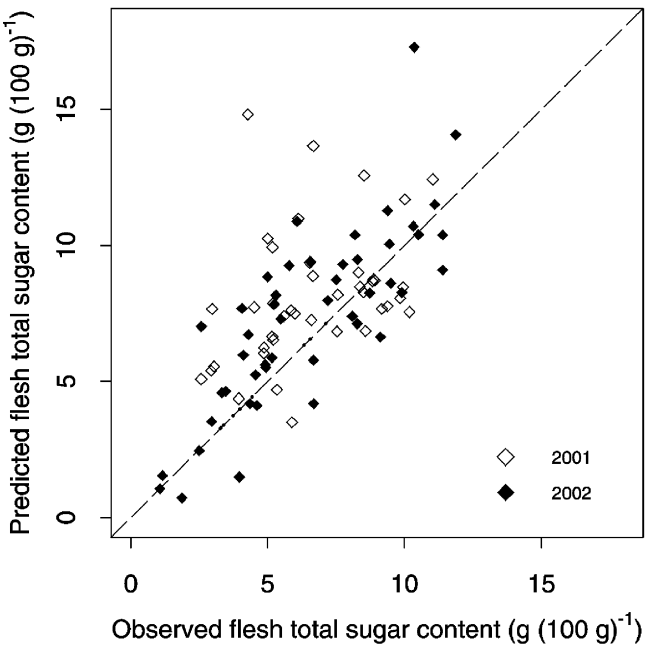

Fig. 4 Test of the predictive quality of the model during fruit growth, calibrated from the BC202 database, on kin01 and kin02 databases. Predicted values of flesh total sugar content [g(100 g_{FM})$^{-1}$] at each date of measurement are plotted against the corresponding observed values, for ten BC2 genotypes and SG in 2001 and eight BC2 genotypes, SG, ZE and P1908 in 2002

The model was also tested on the kin01 and kin02 databases for predictions of mean C_{TS} during fruit growth. Deviations between observations and predictions were comparable between the dates of measurements and were similar in 2001 and 2002 (Fig. 4). The RRMSE values were comparable to those computed for maturity predictions with the BC201 database (Table 3).

Stability of the *k* values and correlations between *k* values and model input variables

The Spearman correlation coefficient between the *k* values estimated from either the BC201 database or the BC02 database for the 87 genotypes in common was correct (0.45) and highly significant (Spearman's test *P* value = 2.8e-05). *k* values from the BC201 database were smaller than those estimated from the BC202 database.

No rank correlation was found between *k* values and the mean fresh flesh mass at maturity from either 2001 (–0.09, Spearman's test *P* value = 0.35) or 2002 (–0.07, Spearman 's test *P* value = 0.40) databases. The Spearman correlation coefficient between the *k* values and the mean fruit growth duration was significantly negative for both years (Fig. 5). Variations in *k* values were larger for the early-maturing genotypes than for the late-maturing ones.

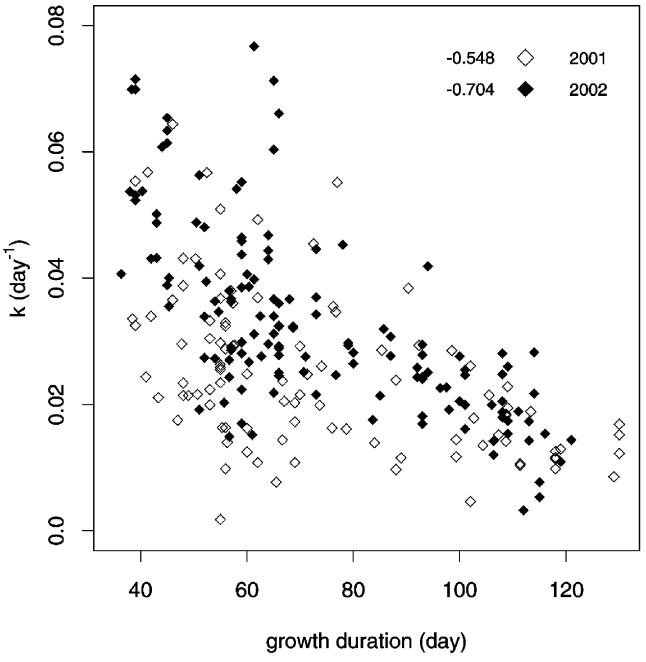

Fig. 5 *k* values estimated from the BC201 and BC202 databases plotted against the corresponding mean fruit growth duration. The Spearman correlation coefficients are indicated for both years

Analysis of total sugar contents through assimilate supply, metabolism and dilution

Variation in total sugar content at maturity, simulated by the model calibrated with the BC202 database, were analyzed through the relative importance of *S* (assimilate supply), *M* (metabolic transformations) and *D* (dilution) (Eq. 5). The 139 BC2 genotypes displayed large variations in the values of *S*, *M* and *D* and the maximal values of these three components for the BC2 genotypes were larger than those of the parents (Fig. 6). *S* and *D* were of the same order of magnitude whereas *M* values were approximately ten times smaller, and of the same order of magnitude as the total sugar content accumulated during the period monitored (C_{TS}–C_{TS}^{ini}).

No close link was found between $C_{TS} - C_{TS}^{ini}$ and any of the three components in particular. Conversely, *S* and *D* appeared to be strongly linked (Fig. 7a). We found that ($C_{TS} - C_{TS}^{ini}$), (*S-D*) and *M* were of the same order of magnitude. Moreover, within the population the observed variation of (*S-D*) was comparable to the variations of *M* and of ($C_{TS} - C_{TS}^{ini}$) (Fig. 7b, c). Indeed, considering the *S* component as reference, *D* represented 72% of *S* on an average. The remaining component (*S-D*) was 28% of *S* on an average. Then, 18% of *S* was attributed to *M*. Thus, only 10% of *S* was assigned to $C_{TS} - C_{TS}^{ini}$. The population displayed large genotypic variations in the relative importance of these components. Indeed, in the population, the relative importance of the components (*S-D*), *M* and ($C_{TS} - C_{TS}^{ini}$) varied from 9 to 56%, 2 to 45% and 2 to 22% of *S*, respectively. Many genotypes in the population showed greater importance for ($C_{TS} - C_{TS}^{ini}$) than the parents (4, 7, and 5% of *S* for P1908, SG and ZE, respectively), which means that for a given amount of assimilate supplied during the monitored period, *S*, they accumulate higher sugar contents than the parents.

Discussion

The simple model we have developed allowed us to simulate the genotypic variation in C_{TS}within a breeding population. Through the input variables, the model takes into account the influence of assimilate supply and dilution on C_{TS} and through the parameter *k*, it considers the effect of metabolic activity. The rank correlation between observations for 2001 and model predictions reached 0.72. Without doubt, the model improved C_{TS} predictions in comparison with predictions based on extrapolation of C_{TS} observations from one year to another.

Such a model could be used to improve the efficiency of breeding strategies. Instead of considering C_{TS} values which result from numerous processes and which are highly affected by environmental factors, it appears worthwhile to describe the elaboration of C_{TS} via a model. This should allow the identication of key variables which might be considered as traits of interest in a

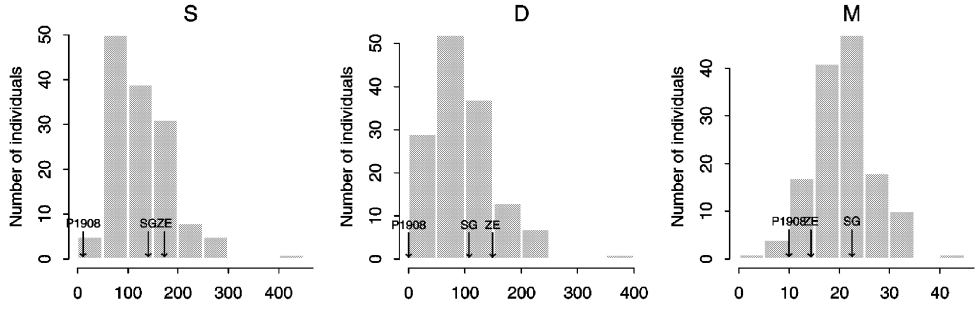

Fig. 6 Distribution of genotype means for cumulated assimilate supply (*S*), metabolic transformation of carbon into compounds other than sugars (*M*) and dilution attributable to the change in fruit volume (*D*). Simulated components were cumulated from the beginning of the monitored period to maturity and expressed in the same units as contents [$g(100\ g_{FM})^{-1}$]. Total sugar content at the end of the monitored period ($C_{TS} - C_{TS}^{ini}$) is equal to *S-D-M*. The values for the parents, P1908, Summergrand (*SG*) and Zephyr (*ZE*), are indicated by *arrows*

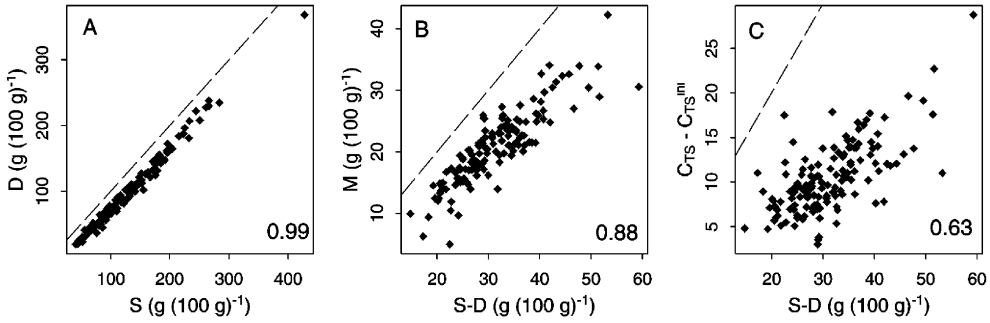

Fig. 7a–c Relationships between the different components of the equation $C_{TS} - C_{TS}^{ini} = S\text{-}D\text{-}M$, where $C_{TS} - C_{TS}^{ini}$ is the total sugar content accumulated during the monitored period, *S* the cumulated assimilate supply, *M* the cumulated metabolic transformation of carbon into compounds other than sugars and *D* the cumulated dilution attributable to the change in fruit volume. Spearman correlation coefficients between the plotted components are indicated

breeding strategy. Indeed, the rank correlation between the *k* values for 2001 and 2002 was higher than that for C_{TS} suggesting an expectation of genetic progress. Moreover, we can consider that the similarity between the two experimental years (very light-loading level treatment) maximized the correlation for C_{TS}, since fruit growth highly influences C_{TS}. Indeed, traits linked to total sugar content, as percent soluble solids, have often been noted to have lower heritability (Souza and Byrne 2000; Rodriguez-Burruezo et al. 2002). The population displayed large variations in *k* values and no unfavorable rank correlation was found between the *k* values and the flesh fresh mass at maturity. Simulations with low *k* values always resulted in enhanced C_{TS}, whatever the growth pattern. Accordingly, selection for low *k* values should be efficient to improve C_{TS} whatever the fruit growth pattern. However, efficiency of selection for low *k* values on C_{TS} increase will depend on the fruit growth pattern and duration, for two reasons. First, the distributions of the *k* values differed between the early- and late-maturing genotypes. For the early-maturing genotypes, *k* values, even though high, displayed enough variation to make genetic gain. For the late-maturing genotypes, the *k* values tended to be low and variations were slight, so that a more limited genetic advance is expected from selection of low *k* values. Second, simulations showed that although minimal *k* values always corresponded to the highest C_{TS} values, the actual level reached by C_{TS} depended on the growth pattern. For the early-maturing genotypes, higher C_{TS} values were observed for high growth, whereas for the late-maturing genotypes, higher C_{TS} values were observed for low growth. Further investigations will be necessary to analyze the effect of interactions between *k* value and the fruit growth pattern on C_{TS} variation. Since the model variables (dry flesh growth rate, fresh flesh mass at maturity, growth duration,$\cdots$) result from numerous processes and are significantly affected by environmental factors, we suggest that a similar modeling approach may be applied to each of these variables. Indeed, models of dry and fresh fruit growth could be used to describe the elaboration of the input variables and identify genotypic parameters that may be independent of the environment.

The model appeared useful to analyze C_{TS} genotypic variation. None of the three components, *S*, *D* or *M*, were highly correlated to C_{TS}. Indeed, differences in C_{TS} between the genotypes at maturity resulted in a balance between the three components. Low simultaneous variation of these components led to substantial variations in C_{TS}. From a practical point of view, this result can rapidly be applied to improve experiments. It underlines the necessity for evaluation of test plants under controlled conditions of fruit growth to make it possible to compare C_{TS} values between genotypes and years. Furthermore, it appears extremely important to consider the three components equally when studying variation in sugar contents. Most of the studies on sugars concern metabolic transformation and much is known about the metabolic pathways involved in sugar accumulation in plants (Ho 1996; Vizzotto et al. 1996; Grof and Campbell 2001). In contrast, few studies take into account assimilate supply and dilution. Metabolic control study (Kacser and Burns 1981) has been successfully used for theoretical genetic studies of metabolism (Bost et al. 1999). In the same way, detailed studies of the mechanisms involved in assimilate supply to fruit and in dilution are necessary to improve the understanding of the variation in sugar content in fruits. Such studies are required to dissect assimilate supply and dilution into components and parameters that might be under simpler genetic control.

Testing of the model on independent databases revealed the limitations of the model. Indeed, even though it allowed us to predict the rank of the genotypes from the BC201 database for C_{TS} at maturity, the predictions were underestimated. Moreover, tests of the observations during fruit growth (the kin01 and kin02 databases) were not accurate. Various reasons can be invoked to explain these limitations. First, the *k* estimations were not precise enough because of the lack of C_{TS} observations for each genotype, especially for observations during fruit growth. Considering the high within-tree variation in C_{TS} commonly observed, more observations, even if made only at maturity, would allow a more accurate estimation of the *k*

values. Second, the model is perhaps too simple, since it does not take into account the possible variation of k during fruit growth. Using the Suncrest cultivar Génard et al. (2003) showed that the relative rate of synthesis of compounds other than sugars increased with increasing relative mesocarp growth rate. This relationship arises from the synthesis of new cellular structures such as cell walls during periods of intense relative growth (Bouranis and Niavis 1992; Fishman et al. 1993). Lastly, some uncontrolled factors were probably involved in the variation of C_{TS}. Through the input variables, we took into account the main environmental factors that influence fruit growth. However, other factors, such as fruit microclimate may be involved in the variation of C_{TS}. Fruit micro-environment modifications are often studied through fruit bagging treatments that modify humidity, light intensity and ambient temperature. Such micro-environment modifications can change physiological process in fruits. For example, bagging often reduces the soluble fruit solids at harvest (Kikuchi et al. 1997). Li et al. (2001) reported changes in various fruit quality traits in response to bagging of peach fruits.

The work presented above is a simple illustration of the joint use of an ecophysiological model and of a breeding population displaying large variation for a trait. This approach is in keeping with the potential contribution of crop modeling, suggested by Hammer et al. (2002), to understand genetic regulation and to help crop improvement. This approach revealed an interesting outlook for both biological understanding and selection. It benefits from the modelling of decomposition of the C_{TS} elaboration, and from the enhanced C_{TS} variation in a breeding population. More knowledge must be integrated in the model to take into account the effect of environmental factors on the processes involved in the elaboration of a trait. Similar conclusions were drawn by Hunt et al. (2003), who failed to identify stable genotypic characteristics of wheat growth over successive seasons. They suggested the need to devote efforts to model improvement and interdisciplinary cooperation. However, such approaches may be extended towards the use of molecular tools and genetic analysis. The ecophysiological model makes it possible to identify the parameters explaining most of the variation of a trait of interest in the breeding population. These parameters could be considered as quantitative traits and used to characterize the genotypes. Thus, detection of quantitative trait loci (QTL) of these parameters may be performed. It is therefore expected that these QTL and their effects will not be environment-dependent. This interdisciplinary approach has already been tested with success to analyze the genetic variability of the responses of maize leaf growth to temperature and water (Reymond et al. 2003). It was also applied by Yin et al. (1999) in barley. Such a QTL analysis approach to ecophysiological variables that stem from models should make it possible to cope with the limitations of quality trait improvement.

Acknowledgements We gratefully acknowledge J. Hostalery for her assistance in the field experiments and T. Pascal for his advice on tree management. We thank E. Rubio and L. Gomez for the sugar analyses.

References

Abbott AG, Rajapakse S, Sosinski B, Lu ZX, Sossey-Alaoui K, Scorza R, Callahan A (1998) Construction of saturated linkage maps of peach crosses segregating for characters controlling fruit quality, tree architecture and pest disease. Acta Hortic 465:41–50

Bernacchi D, Beck-Bunn T, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD (1998) Advanced backcross QTL analysis in tomato I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. Theor Appl Genet 97:381–397

Black CC, Mustardy L, Sung SS, Kormanik PP, Xu DP, Pas N (1987) Regulation and roles for alternative pathways of hexose metabolism in plants. Physiol Plant 69:387–394

Boote KJ, Jones JW, Pickering NB (1996) Potential uses and limitations of crop models. Agron J 88:704–716

Boote KJ, Kropff MJ, Bindraban PS (2001) Physiology and modeling of traits in crop plants: implications for genetic improvement. Agric Syst 70:395–420

Bost B, Dillmann C, de Vienne D (1999) Fluxes and metabolic pools as model traits for quantitative genetics. I. The L-shaped distribution of gene effects. Genetics 153:2001–2012

Bouranis DL, Niavis CA (1992) Cell wall metabolism in growing and ripening stone fruit. Plant Cell Physiol 33:999–1008

Byrne DH (2002) Peach breeding trends: a world wide perspective. Acta Hortic 592:49–59

Causse M, Saliba-Colombani V, Lesschaeve I, Buret M (2001) Genetic analysis of organoleptic quality in fresh market tomato 2. Mapping QTLs for sensory attributes. Theor Appl Genet 102:273–283

Chambers JM, Hastie TJ (1992) Statistical models. Wadsworth & Brook/Cole Computer Science Series, Pacific Grove, p 608

Corelli-Grappadelli L, Coston DC (1991) Thinning pattern and light environment in peach tree canopies influence fruit quality. Hortic Sci 26:1464–1466

Crisosto CH, Johnson RS, Luza JG, Crisosto GM (1994) Irrigation regimes affect fruit soluble solids concentration and rate of water loss of 'O'Henry' peaches. Hortic Sci 29:1169–1171

Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, Plomion C, Monet R (1999) Mapping QTLs controlling fruit quality in peach [*Prunus persica* (L) Batsch]. Theor Appl Genet 98:18–31

Etienne C, Moing A, Dirlewanger E, Raymond P, Monet R, Rothan C (2002) Isolation and characterization of six peach cDNAs encoding key proteins in organic acid metabolism and solute accumulation:involvement in regulating peach fruit acidity. Physiol Plant 114:259–270

Fishman ML, Levaj B, Gillespie R, Scorza R (1993) Changes in the physicochemical properties of peach fruit pectin during on-tree ripening and storage. J Am Soc Hortic Sci 118:343–349

Fulton TM, Beck-Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD (1997) QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. Theor Appl Genet 95:881–894

Génard M (1992) Influence du nombre de feuilles et de la répartition des fruits sur la production et la qualité des pêches. Can J Plant Sci 72:517–525

Génard M, Bruchou C, Souty M (1991) Variabilité de la croissance et de la qualité chez la pêche (*Prunus persica* L Batsch) et liaison entre croissance et qualité. Agronomie 11:829–845

Génard M, Bruchou C (1992) Multivariate analysis of within-tree factors accounting for the variation of peach fruit quality. Scientia Hortic 52:37–51

Génard M, Souty M (1996) Modeling the peach sugar contents in relation to fruit growth. J Am Soc Hortic Sci 121:1122–1131
Génard M, Lescourret F, Gomez L, Habib R (2003) Changes in fruit sugar concentrations in response to assimilate supply, metabolism and dilution: a modeling approach applied to peach fruit (*Prunus persica*). Tree Physiol 23:373–385
Gomez L, Rubio E, Augé M (2002) A new procedure for extraction and measurement of soluble sugars in ligneous plants. J Sci Food Agric 82:360–369
Grof CPL, Campbell JA (2001) Sugarcane sucrose metabolism: scope for molecular manipulation. Aust J Plant Physiol 28:1–12
Gurrieri F, Audergon JM, Albagnac G, Reich M (2001) Soluble sugars and carboxylic acids in ripe apricot fruit as parameters for distinguishing different cultivars. Euphytica 117:183–189
Hammer GL, Butler DG, Muchow RC, Meinke H (1996) Integrating physiological understanding and plant breeding via crop modeling and optimization. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. CAB International, Wallingford, pp 419–441
Hammer GL, Kropff MJ, Sinclair TR, Porter GW (2002) Future contributions of crop modeling from heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement. Eur J Agron 18:15–31
Hashizume T, Shimamoto I, Hirai M (2003) Construction of a linkage map and QTL analysis of horticultural traits for watermelon [*Citrullus lanatus* (THUNB) MATSUM & NAKAI] using RAPD, RFLP and ISSR markers. Theor Appl Genet 16:779–785
Ho LC (1996) The mechanism of assimilate partitioning and carbohydrate compartmentation in fruit in relation to the quality and yield of tomato. J Exp Bot 47:1239–1243
Hubbard NL, Pharr DM, Huber SC (1991) Sucrose synthase and other sucrose metabolizing enzymes in fruits of various species. Physiol Plant 82:191–196
Huet S, Bouvier A, Gruet M-A, Jolivet E (1996) Statistical tools for nonlinear regression. Springer, Berlin Heidelberg New York, p 154
Hunt LA, Reynolds MP, Sayre KD, Rajaram S, White JW, Yan W (2003) Crop modeling and the identification of stable coefficients that may reflect significant groups of genes. Agron J 95:20–31
Kacser H, Burns JA (1981) The molecular basis of dominance. Genetics 97:639–666
Kikuchi T, Arakawa O, Norton RN (1997) Improving skin color of 'Fuji' apples in Japan. Fruit Var J 51:71–75
Kliewer WM, Weaver RJ (1971) Effect of crop level and leaf area on growth, composition and coloration of Tokay grapes. Am J Enol Viticult 22:172–177
Kobayashi K, Us Salam M (2000) Comparing simulated and measured values using mean squared deviation and its components. Agron J 92:345–352
Ladizinsky G (1985) Founder effect in crop-plant evolution. Econ Bot 39:191–199
Leonard S, Luh BS, Hinreiner E (1953) Flavor evaluation of canned cling peaches. Food Technol 480–485
Li SH, Huguet JG, Schoch PG, Orlando P (1989) Response of peach tree growth and cropping to soil water deficit at various phenological stages of fruit development. J Hortic Sci 64:541–552
Li SH, Génard M, Bussi C, Huguet JG, Habib R, Besset J (2001) Fruit quality and leaf photosynthesis in response to microenvironment modification around individual fruit by covering the fruit with plastic in nectarine and peach trees. J Hortic Sci Biotechnol 76:61–69
Marini RP, Sowers DL, Marini MC (1991) Peach fruit quality is affected by shade during final swell of fruit growth. J Am Soc Hortic Sci 116:383–389
Moriguchi T, Ishizawa Y, Sanada T (1990) Differences in sugar competition in *Prunus persica* fruit and the classification by the principal component analysis. J Jpn Soc Hortic Sci 59:307–312
Pascal T, Kervella J, Pfeiffer F, Sauge MH, Esmenjaud D (1998) Evaluation of the interspecific progeny *Prunus persica* cv Summergrand × *Prunus davidiana* for disease resistance and some agronomic features. Acta Hortic 465:185–192
Quarta R, Dettori MT, Verde I, Gentile A, Broda Z (1998) Genetic analysis of agronomic traits and genetic linkage mapping in a BC1 peach population using RFLPs and RAPDs. Acta Hortic 465:51–60
Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F (2003) Combining quantitative trait loci: analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. Plant Physiol 131:664–675
Reynders S, Monet R (1987) Evolution au cours du temps de la consanguinité des variétés de pêchier. Etude des distances génétiques entre quelques géniteurs. Fruit 42:529–535
Robertson JA, Meredith FI, Forbus WR, Lyon BG (1992) Relationship of quality characteristics of peaches (cv Loing) to maturity. J Food Sci 57:1401–1404
Rodriguez-Burruezo A, Prohens J, Nuez F (2002) Genetic analysis of quantitative traits in pepino (*Solanum muricatum*) in two growing seasons. J Am Soc Hortic Sci 127:271–278
Saliba-Colombani V, Causse M, Langlois D, Philouze J, Buret M (2001) Genetic analysis of organoleptic quality in fresh market tomato 1. Mapping QTLs for physical and chemical traits. Theor Appl Genet 102:259–272
Shorter R, Lawn RJ, Hammer GL (1991) Improving genotypic adaptation in crops: a role for breeders, physiologists and modelers. Exp Agric 2:155–175
Souza VA, Byrne DH (2000) Predicted breeding values for nine plant and fruit characteristics of 28 peach genotypes. J Am Soc Hortic Sci 125:460–465
Stepansky A, Kovalski I, Schaffer AA, Perl-Reves R (1999) Variation in sugar levels and invertase activity in mature fruit representing a broad spectrum of *Cucumis melo* genotypes. Genet Resour Crop Evol 46:53–62
Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor Appl Genet 92:213–224
Tardieu F (2003) Virtual plants: modeling as a tool for the genomics of tolerance to water deficit. Trends Plant Sci 8:9–14
Vicente MC de, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585–596
Vizzotto G, Pinton R, Varanin Z, Costa G (1996) Sucrose accumulation in developing peach fruit. Physiol Plant 96:225–230
Xiao J, Li J, Grandillo S, Ahn S, Yuan L, Tanksley SD, McCouch SR (1998) Identification of trait-improving QTL alleles from a wild rice relative, *Oryza rufipogor*. Genetics 150:899–909
Yin X, Kropff MJ, Stam P (1999) The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. Heredity 82:415–421