Journal of Experimental Botany, Vol. 64, No. 16, pp. 5099–5113, 2013
doi:10.1093/jxb/ert297 Advance Access publication 8 October, 2013

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

New insights for estimating the genetic value of segregating apple progenies for irregular bearing during the first years of tree production

Jean-Baptiste Durand[1,2,*], Baptiste Guitton[3], Jean Peyhardi[2,4], Yan Holtz[5], Yann Guédon[2], Catherine Trottier[4] and Evelyne Costes[5]

[1] Grenoble University, Laboratoire Jean Kuntzmann, BP53, F-38041 Grenoble Cedex 9, France
[2] CIRAD/Inria, Virtual Plants Team, UMR AGAP, F-34398 Montpellier, France
[3] CIRAD, UMR AGAP, F-34398 Montpellier, France
[4] Institut de Mathématiques et de Modélisation de Montpellier, Université Montpellier 2, F-34095 Montpellier Cedex 9, France
[5] INRA, UMR AGAP, Equipe Architecture et Fonctionnement des Espèces Fruitières, F-34398 Montpellier, France

* To whom correspondence should be addressed. E-mail: jean-baptiste.durand@imag.fr

Received 19 February 2013; Revised 26 July 2013; Accepted 12 August 2013

Abstract

Because irregular bearing generates major agronomic issues in fruit-tree species, particularly in apple, the selection of regular cultivars is desirable. Here, we aimed to define methods and descriptors allowing a diagnostic for bearing behaviour during the first years of tree maturity, when tree production is increasing. Flowering occurrences were collected at whole-tree and (annual) shoot scales on a segregating apple population. At both scales, the number of inflorescences over the years was modelled. Two descriptors were derived from model residuals: a new biennial bearing index, based on deviation around yield trend over years and an autoregressive coefficient, which represents dependency between consecutive yields. At the shoot scale, entropy was also considered to represent the within-tree flowering synchronicity. Clusters of genotypes with similar bearing behaviours were built. Both descriptors at the whole-tree and shoot scales were consistent for most genotypes and were used to discriminate regular from biennial and irregular genotypes. Quantitative trait loci were detected for the new biennial bearing index at both scales. Combining descriptors at a local scale with entropy showed that regular bearing at the tree scale may result from different strategies of synchronization in flowering at the local scale. The proposed methods and indices open an avenue to quantify bearing behaviour during the first years of tree maturity and to capture genetic variations. Their extension to other progenies and species, possible variants of descriptors, and their use in breeding programmes considering a limited number of years or fruit yields are discussed.

Key words: Alternation indices, biennial bearing, cultivar breeding, linear mixed models, *Malus*×*domestica*, multiscale model, QTL detection, synchronicity.

Introduction

Numerous economically important fruit-tree species, including apple, are prone to irregular bearing, i.e. irregular fruit load of a tree over consecutive years. Irregular bearing is characterized by numerous small-sized and low-quality fruits during years of high production ('ON' years) and by few oversized fruits during years of low yield ('OFF' years) (Monselise and Goldschmidt, 1982). Alternate bearing denotes a specific case of irregular bearing for which fruiting pattern is biennial, this case being predominant in apple, even though more complex patterns are possible. In addition to the fluctuation

Abbreviations: BBI, biennial bearing index; FDA, factorial discriminant analysis; GS, Granny Smith; LAS, long axillary shoot; LPAS, long proleptic axillary shoot; LSAS, long sylleptic axillary shoot; QTL, quantitative trait loci; SAS, short axillary shoots; SD, standard deviations; STK, Starkrimson.

of production, this alternation generates major agronomic issues such as an increased demand for labour and chemical products to regulate crop loads (Jonkers, 1979). However, because of their impact on the environment, fewer chemical solutions are available for growers today. This implies the need for innovating solutions to attenuate irregular bearing. The main hypothesis behind the biennial bearing phenomenon is that fruit yield in one year has a negative effect on floral formation for the subsequent year (Jonkers, 1979; Monselise and Goldschmidt, 1982; Neilsen and Dennis, 2000). A previous study showed that biennial bearing is inheritable and segregates in an apple progeny ('Starkrimson'×'Granny Smith' population) (Guitton *et al.*, 2012), suggesting that selecting new varieties with intrinsic regular bearing is a possible strategy. However, breeding programmes for fruit-tree species do not yet consider this trait, because it expresses exclusively during tree mature phase, i.e. after the tree started flowering and fruiting, and requires observations over several years before its value for a given genotype can be assessed. Development of methods and tools for a faster diagnostic of the bearing tendency of a genotype during its first years of production is thus highly desirable.

Several parameters have been proposed to quantify biennial bearing. These parameters are based on either: (i) indices quantifying the stability of yields through time, or (ii) repeated measure analysis to characterize the relationships between successive yields. The first approach, based on indices, has been developed to assess bienniality and intensity of alternation, as well as its synchronicity in different locations. Hoblyn *et al.* (1936) proposed an index to estimate the intensity of deviation in yield during successive years, which has been renamed by Wilcox (1944) as the biennial bearing index (BBI). BBI has become the accepted standard to describe biennial bearing and has been applied to yield (mass of fruit) at different scales—whole areas, individual trees, and branches—on apple and other fruit-tree species (Wilcox, 1944; Singh, 1948; Pearce and Dobersek-Urbane, 1967; Jonkers, 1979; Reddy *et al.*, 2003; Smith *et al.*, 2004; Rosenstock *et al.*, 2010). Huff (2001) highlighted that the distribution of BBI strongly depends on the mean and variance of yields, under the hypothesis that they are a random sample. Therefore, the accepted interpretation of BBI as a measure of the magnitude of irregular bearing is questionable. To address this issue, Huff (2001) proposed a significance test of BBI, but it has been used only once, on citrus (Smith *et al.*, 2004). Pearce and Dobersek-Urbane (1967) investigated empirical properties of indices for bienniality in practical situations, using data simulated with several models of patterns. In particular, they highlighted that simulated trended series without alternation had positive BBI values. As a result, using BBI on trended series may lead to confounding alternation and trend.

Indices for alternation have also been proposed by analysing successions of shoot types (including vegetative, flowering but not fruiting, and fruiting) along branches over consecutive years (Lauri *et al.*, 1995, 1997). Two indices have been proposed, a local index reflecting the ability to produce bourse over bourse, and a global index quantifying the alternation synchronicity within the tree. While characterizing genotypic variations in fruiting patterns in various apple cultivars, Lauri *et al.* (1995) noticed that alternation-to-fruit patterns can be completely hidden when alternating sequences are desynchronized. The index developed to assess this phenomenon (named *alternation synchronism*) quantifies the balance between the number of 2-year-long sequences with fruiting alternation, depending on whether fruiting occurs in even or odd years.

The second approach, referred to as *longitudinal data analysis*, relies on mixed models to analyse multi-year datasets (Cnaan *et al.*, 1997; Verbeke and Molenberghs, 2000). Their main strength is to provide a model for the serial correlation between successive observations for an individual (Diggle and Kenward, 1994; Hand and Crowder, 1996). Diggle and Kenward (1994) identified three sources of variation in longitudinal data: serial correlations, random effects, and measurement error. As a consequence, such models have been used to decompose the global variability among these sources.

Recently, an experiment studying yield stability over time on *Coffea canephora* considered both strategies, i.e. longitudinal data analysis and indices (Cilas *et al.*, 2011). Longitudinal data analysis was used to model the residual correlation structure among tree yields over time, so as to estimate genotypic values associated with yield as accurately as possible. Classical indices (BBI and the number of sign changes in the difference between consecutive yields) were computed on yield directly, with the prospect of deriving heritability values and correlations between traits. These indices and the longitudinal data analysis were considered independently.

Our claim is that the genetic variability of correlations between annual yields contains relevant information, in the context of selection. Thus, our analysis was based on an integrated model that combined both correlations and their genetic variability. Based on a previous study showing that the changes in harvested fruits were less impacted by the genotype than by the number of inflorescences per year (Guitton *et al.*, 2012), it focused on the estimation of genotype ability to flower regularly. Flowering occurrences were collected through several consecutive years during the beginning of the mature phase of an apple tree progeny, at two scales: whole-tree (referred to as *global*) and annual shoot (referred to as *local*) scales. The first objective of the present study was the early identification of irregular genotypes during the first years of tree production, when it is affected by a trend (i.e. a gradual increase in the mean level of yield over years). Thus, quantification of biennial bearing at global scale relied on mixed models to separate this trend from the alternating pattern related to biennial bearing. The degree of alternate bearing was quantified by a dependency model of the successive deviations of yields from the trend. The intensity of biennial bearing was assessed using either the model parameters or indices measuring the relative amplitudes of the deviations. Our second objective was to investigate the possibility of early identification of irregular genotypes using retrospective measurements performed on annual

shoots in individuals. The use of such samples, consisting of a limited number of sequences, would make phenotyping easier and faster than exhaustive measurements. Thus, we established and validated a methodology to predict bearing behaviour from annual shoot samples. Moreover, combining descriptors of alternation computed at the local scale and of synchronization between axes brought new (but still partial) insights into how regularity at the tree scale may result from different strategies of synchronization in flowering at the local scale.

Materials and methods

Plant material

A segregating population described by Guitton *et al.* (2012) was obtained from a cross between 'Starkrimson' and 'Granny Smith', and was used to study the bearing behaviour of genotypes. Two tree replicates (or just one in rare cases) were available for each genotype.

The Starkrimson × Granny Smith (STK×GS) progeny was the first generation after crossing (F1) and was composed of 123 genotypes. Because both parents were highly heterozygous and polymorphic, the F1 population showed segregation of both tree architectures during the strictly vegetative phase (Segura *et al.*, 2006, 2007, 2009) and biennial bearing during the beginning of the mature phase (Guitton *et al.*, 2012). Most of the trees of the population (87.6%) were able to set flower during their third year after grafting (2006). Contrasted bearing behaviours were observed within the population, ranging from regular to strongly biennial bearing, with some genotypes exhibiting irregular patterns (Guitton *et al.*, 2012).

Phenotyping

Flowering recurrence was measured at global and local scales. At the global scale, the total number of inflorescences was observed during 6 consecutive years, from their second to their seventh year after grafting. Trees for which less than 2 years of substantial yield were available were discarded. In some cases, flowering did not occur in the first years for any replication: such years were also discarded.

At the local scale, successive vegetative versus floral annual shoots (defined as the portion of axes developed during the same year) were observed over the same consecutive years along different types of axes: trunk, and long and short axillary shoots (LAS and SAS, respectively) (Supplementary Fig. S1 at *JXB* online). Axillary shoots were classified depending on their annual shoot lengths: with all annual shoot lengths <5 cm in SAS, with annual shoot length between 5 and 20 cm in brindles, and with at least one annual shoot with length ≥20 cm in LAS (see Segura *et al.*, 2006). Within LAS, proleptic and sylleptic branches were distinguished according to whether the meristem outgrowth occurred after a dormant period or not [long proleptic axillary shoot (LPAS) and long sylleptic axillary shoot (LSAS), respectively]. Vegetative and floral annual shoots were distinguished according to the presence/absence of an inflorescence. Floral annual shoots were composed of a 'bourse' (leafy basal part ending with an inflorescence) and a 'bourse shoot' (vegetative growth units originating from a lateral meristem of the bourse).

Flowering occurrence was observed along the trunk of each tree, as well as along two LAS sampled along the first trunk's annual shoots (grown in 2004): one LSAS and one LPAS (Supplementary Fig. S1). For each LAS, two SAS per annual axillary shoot were also sampled and phenotyped the same way. As a consequence, ten SAS of 5 to 1 years were recorded on LSAS, and eight SAS of 4 to 1 years were recorded on LPAS. The flowering pattern was described by recording the presence/absence of flowering events on annual shoots (six possible flowering occurrences on the trunk and LSAS, and five on LPAS). The data thus consisted of vegetative versus floral annual shoots in 6- to 1-year sequences.

Statistical modelling

Modelling biennial bearing at the global scale

Trend models. To dissociate the increase in the number of inflorescences per tree from biennial bearing, a trend model based on a linear mixed model was first applied. Two linear mixed models (referred to as A and B) were considered. Their slope and intercept can be decomposed into three terms: mean, genotype, and tree replication effects. Both models can be generically written as:

$$Y_{g,r,t} = \underbrace{\beta + \beta_g + \zeta_{g,r}}_{\text{Intercept}} + \underbrace{(\alpha + \alpha_g + \xi_{g,r})}_{\text{Slope}} t + \varepsilon_{g,r,t}$$

where $Y_{g,r,t}$ is the production of tree replicate r of genotype g at year t, $\beta + \alpha t$ represents the mean trend (treated as fixed effect), $\beta_g + \alpha_g t$ the trend deviation for genotype g, $\zeta_{g,r} + \xi_{g,r} t$ the trend deviation for replicate r of genotype g, and where $\varepsilon_{g,r,t}$ is the Gaussian residual of tree replicate r of genotype g at time t. Genotype effects β_g and α_g were considered as fixed effects in model A and as Gaussian random effects in model B, with variances τ^2_β and τ^2_α, respectively. Replication-specific parameters ζ_{gr} and ξ_{gr} were not directly of interest since we aimed at quantifying irregular bearing at the genotype scale. Therefore, they were treated as random effects, with respective variances τ^2_ζ and τ^2_ξ. Our interest being to characterize each genotype for its bearing behaviour, modelling genotype effect as fixed (model A) rather than as random effect (model B) was preferred *a priori*. Selection of the most suitable model among A and B was performed using minimization of AIC and BIC criteria (Verbeke and Molenberghs, 2000).

Quantifying deviation around trend

Construction of indices inspired by BBI. Following Huff's suggestion (2001) that yields should be detrended in BBI calculations, we proposed to characterize irregular bearing in the context of trended yield by incorporating the empirical residuals $\hat{\varepsilon}_{g,r,t}$ of trend models A or B into BBI-related indices.

Let us recall that BBI is defined for a sample by:

$$\mathrm{BBI} = \frac{2}{\sum_r (T_{g,r} - 1)} \sum_r \sum_{t=2}^{T_{g,r}} \frac{\left| Y_{g,r,t} - Y_{g,r,t-1} \right|}{Y_{g,r,t-1} + Y_{g,r,t}}$$

where $T_{g,r}$ denotes the number of measurements for replication r of genotype g. Compared with the usual presentation of BBI, a multiplying factor of value 2 was introduced to make BBI comparable in scale to the indices introduced below. The justification is that, in this way, the elementary terms that are averaged take the form of a ratio between an absolute difference $\left| Y_{g,r,t} - Y_{g,r,t-1} \right|$ and a mean $\left(Y_{g,r,t-1} + Y_{g,r,t} \right)/2$.

The BBI suffers from the shortcomings mentioned in the Introduction. A detailed mathematical analysis of these shortcomings is given in Supplementary material A at *JXB* online. This analysis led us to propose the following variants of BBI, called BBI_norm and BBI_res_norm, which consisted of a normalization of the residual fluctuations by total yield:

$$\mathrm{BBI\_norm} = \frac{\sum_r \sum_{t=2}^{T_{g,r}} \left| Y_{g,r,t} - Y_{g,r,t-1} \right| / \sum_r (T_{g,r} - 1)}{\sum_r \sum_{t=1}^{T_{g,r}} Y_{g,r,t} / \sum_r T_{g,r}};$$

$$\mathrm{BBI\_res\_norm} = \frac{\sum_r \sum_{t=2}^{T_{g,r}} \left| \hat{\varepsilon}_{g,r,t} - \hat{\varepsilon}_{g,r,t-1} \right| / \sum_r (T_{g,r} - 1)}{\sum_r \sum_{t=1}^{T_{g,r}} Y_{g,r,t} / \sum_r T_{g,r}}.$$

BBI_norm is the ratio between the mean absolute difference between successive yields and the average yield. In BBI, each absolute difference is given a different weight; in contrast, the differences are simply summed and normalized in BBI_norm. This normalization scheme was expected to render BBI less sensitive to deviation from the assumption of alternating residuals with amplitudes roughly proportional to the corresponding trend level (see proposition P2 in Supplementary material at *JXB* online). If the trend variation between $t-1$ and t is small with respect to $\hat{\varepsilon}_{g,r,t-1}$ and $\hat{\varepsilon}_{g,r,t}$, BBI_res_norm is expected to be a close approximation of BBI_norm (see proposition P3 in Supplementary material). The role of BBI_res_norm is to measure the amplitude of successive variations of yield around the trend, relative to total yield. Although this is a measure of yield irregularity, biennial or irregular series with same BBI_res_norm could be built, showing that BBI_res_norm is not sufficient to discriminate between those patterns. Distinguishing between such patterns requires modelling the dependencies between successive residuals.

Modelling residual structure. The hypothesis of regular bearing for a given genotype g is associated with the assumption of independent residuals $(\varepsilon_{g,r,t})_{0\le t\le T}$ and small BBI_res_norm. In contrast, the hypothesis of biennial bearing is associated with dependent residuals, negative correlation between successive residuals (since they are expected to have opposite signs), and high BBI_res_norm. The individuals combining independent residuals and high BBI_res_norm were expected to be associated with irregular bearing.

Dependency between successive residuals can be characterized using first-order autoregressive (AR1) models. Two models (referred to as I and II) are considered; they can be generically written as:

$$\varepsilon_{g,r,t} = (\gamma + \gamma_g + \gamma_{g,r})\varepsilon_{g,r,t-1} + u_{g,r,t}$$

where $\varepsilon_{g,r,t}$ is the same residual as in trend model A with fixed effects, γ is the mean AR1 coefficient (common to every genotype and replicate), γ_g the fixed deviation from the mean AR1 coefficient for genotype g, $\gamma_{g,r}$ the random deviation from the AR1 coefficient $\gamma+\gamma_g$ for tree replicate r of genotype g, and $u_{g,r,t}$ the residual of the AR1 model for tree replicate r of genotype g at time t. Let τ_1^2 and ρ^2 denote the variance of $\gamma_{g,r}$ and $u_{g,r,t}$, respectively. The models do not have intercepts, since by hypothesis the $\varepsilon_{g,r,t}$ have zero mean.

Regular genotypes were expected to have values of $\gamma+\gamma_g$ close to zero, whereas biennial bearing genotypes would have values close to –1. Model II is the model defined by the above equation, whereas model I is obtained by letting the variance τ_1^2 of $\gamma_{g,r}$ vanish (thus removing these effects from the model). The variances of $\varepsilon_{g,r,t}$ and $Y_{g,r,t}$ induced by combining model A with model I are given by:

$$\mathrm{var}(\varepsilon_{g,r,t}) = \rho^2\frac{\left(1-(\gamma_g)^{2t}\right)}{1-\gamma_g^2};\mathrm{var}(Y_{g,r,t}) = \tau_\xi^2 t^2 + \rho^2\frac{\left(1-(\gamma_g)^{2t}\right)}{1-\gamma_g^2}. \quad (*)$$

Thus, the residual variance tends increasingly to $\rho^2/1-\gamma_g^2$. In the analysis, for similar reasons as those detailed above for trend models, model I was preferred *a priori* to model II.

In practice, to estimate these models, the empirical residuals $\hat{\varepsilon}_{g,r,t}$ were used, rather than the actual (unknown) residuals $\varepsilon_{g,r,t}$ of trend model A with fixed effects. The model parameters were estimated by maximum likelihood and the best linear unbiased predictors for the random effects were computed using the R software (R Development Core Team, 2008) and specifically the function *lmer* of package *lme4* (Bates *et al.*, 2011).

Prediction of biennial bearing using subsamples of annual shoot sequences

Biennial bearing at the local scale was characterized using 1- to 6-year sequences of vegetative versus flowering shoots $(F_{g,r,t,\ell})_{\ell\ge 0}$, ($F_{g,r,t,\ell}$=0) denoting the absence and ($F_{g,r,t,\ell}$=1) the presence of flower for replication r of genotype g at year t, at location (or shoot) ℓ in the tree. Information contained in this subsample of sequences was used to predict both indices developed in 'Modelling biennial bearing at the global scale' above.

First, for each replication r of each genotype g, the total number of inflorescences

$$Y'_{g,r,t} = \sum_l F_{g,r,t,l}$$

contained in the subsample of sequences was computed. These quantities were used as in 'Modelling biennial bearing at the global scale' to compute both indices BBI_res_norm and the genotype AR coefficient γ_g, replacing $Y_{g,r,t}$ by $Y'_{g,r,t}$ in the above formulas. It was hypothesized that the indices obtained using the subsample of sequences would be sufficiently good approximations of the same indices computed from the data at global scale.

Secondly, we have developed a third index dedicated to the quantification of synchronicity in flowering among all shoots of a same year. This index is based on the entropy of a random variable, which is a measure of uncertainty concerning its values. In the case of a binary variable F with Bernoulli distribution with parameter p, the entropy is $-p\log p-(1-p)\log(1-p)$. Its minimum 0 is reached when p=0 or p=1 (no uncertainty on F and perfect synchronicity in flowering), and its maximum log 2 is reached when p=0.5 (maximal uncertainty on F and perfectly erratic flowering). The entropy of $F_{g,r,t,m,\ell}$ averaged over years t is given by:

$$Ent_{g,r} = -\frac{1}{\sum_t n_{g,r,t}}\sum_t\left(n_{g,r,t,0}\log\frac{n_{g,r,t,0}}{n_{g,r,t}} + n_{g,r,t,1}\log\frac{n_{g,r,t,1}}{n_{g,r,t}}\right)$$

where $n_{g,r,t} = n_{g,r,t,0} + n_{g,r,t,1}$ is the number of shoots for replication r of genotype g at year t, $n_{g,r,t,0}/n_{g,r,t}$ the estimated probability of vegetative shoot, and $n_{g,r,t,1}/n_{g,r,t}$ the estimated probability of flowering shoot. The actual interpretation and behaviour of *Entg,r* are illustrated in Supplementary Table S1 at *JXB* online on replications of regular, irregular, and biennial bearing genotypes. This index can be directly extended to characterize synchronicity at the genotype scale, replacing $n_{g,r,t,}$ by the number $n_{g,t}$ of shoots for genotype g at year t, and $n_{g,r,t,v}$ by $n_{g,t,v}$ accordingly. In practice, we chose to characterize synchronicity at the genotype scale by computing the mean value $\overline{Ent}_g$ of value $Ent_{g,r}$ averaged over r, to favour synchronicity within replications rather than between replications.

Alternation in yield for each genotype was characterized using two indices computed at the global scale. They were derived from trend model A with fixed effects and residual AR1 model I with fixed effects: BBI using the estimated residuals with normalization by average yields (BBI_res_norm) and the genotype-related component in the AR1 model of the residual structure (γ_g, referred to in brief as *genotype AR coefficient* in the sequel). These two indices will be referred to as *global indices*. The same indices computed from the subsamples of sequences (denoted respectively by B^{loc} and γ^{loc}), plus the entropy index, will be referred to as *local indices*.

Clustering of genotypes from global indices and ranking

In order to identify groups of genotypes with similar bearing behaviours, clustering was performed using the vector of global indices for each genotype. The model used for clustering (into K clusters) was a Gaussian mixture model with K components (McLachlan and Peel, 2000) assuming that within each cluster, BBI_res_norm and γ_g are independent Gaussian random variables whose means and variances depend on the cluster. The detailed variant of the model is described formally in Supplementary material M1 at *JXB* online.

A ranking of the genotypes was obtained using a factorial discriminant analysis (FDA; see Tabachnick and Fidell, 2007). FDA provides a dimensionality reduction described in Supplementary material M3

at *JXB* online, such that the classes are as separated as possible after projection of the genotypes (characterized by BBI_res_norm and γ_g) into linear subspaces. If the classes are well separated on the first axis, the genotypes can be scored and ranked using their coordinates on this axis (the first ranks correspond to the most regular genotypes and the last ranks to the alternate bearing genotypes).

Prediction of classes and global indices from local indices

In order to characterize the relationship between alternation in flowering at both local and global scales, two kinds of statistical methods were combined. The first consisted of predicting the global indices of alternation I_g from the local indices B^{loc}, γ^{loc} and $\overline{Ent}_g$. This was performed using Gaussian linear models, written generically as:

$$I_g = a_0 + a_1 B_g^{loc} + a_2 \gamma_g^{loc} + a_3 \overline{Ent}_g + \eta_g$$

where I_g denotes either BBI_res_norm or γ_g and where η_g is a Gaussian residual. *P* values of regressions and (multiple) correlation coefficients were computed, with confidence intervals.

The second relied on the prediction of the classes yielded by Gaussian mixture clustering. The aim of this approach was to predict (i.e. to retrieve) the class of each genotype, using information contained in the subsamples of sequences only. This issue corresponds to the statistical framework of supervised classification. The three local indices were used (simultaneously) as predictors. Thus, in the case of *K* classes, a classification provides a partition of $\mathbf{R}^3$ where each of the *K* components of the partition is associated with one given class. Supervised classification was performed using feed-forward neural networks, which essentially are non-linear multivariate regression models (see Chapter 5 of Bishop, 2006). These models output probabilities for each genotype to belong to each possible class. The R-based implementation *nnet* was used (Venables and Ripley, 2002).

Support vector machines were also considered (see Chapter 7 of Bishop, 2006). Support vector machines provide partitions determined by optimizing geometric criteria, so as to maximize the distance between the boundaries and the genotypes closest to these boundaries. The R-based implementation *kernlab* was used (Karatzoglou *et al.*, 2004). Further precisions on parameter estimation or model selection and validation are given in Supplementary material M2, at *JXB* online.

Model validation: prediction of future yields at tree scale from past yields

The statistical trended AR model A-I can be used to predict future yields at tree scale. The predictions were validated, using separate learning and test samples. Six years of production were available, but model A-I was re-estimated using the first 5 years of production only (i.e. production between 2005 and 2009 as the learning sample), isolating the 2010 data as a test sample. The conditional distribution of $Y_{g,r,t}$ at year 2010 (corresponding to t=5) given the data up to 2009 was computed. This conditional distribution is Gaussian with mean $\beta + \beta_g + [\alpha + \alpha_g + E(\xi_{g,r}|\boldsymbol{Y})]t + \gamma_g \varepsilon_{g,r,t-1}$ and variance $t^2 \,\mathrm{var}(\xi_{g,r}|\boldsymbol{Y}) + \rho^2$, where the best linear unbiased predictor $E(\xi_{g,r}|\boldsymbol{Y})$ and the conditional variance $\mathrm{var}(\xi_{g,r}|\boldsymbol{Y})$ can be computed directly, using the R package *lme4*. Then a prediction interval that contains $Y_{g,r,t}$ with probability 1 - *p* (*p*=0.05) was deduced, and the actual frequency of $Y_{g,r,t}$ in year 2010 that actually belonged to this interval of prediction was computed.

The clusters based on the new estimates of BBI_res_norm and γ_g (using the first 5 years of production) were updated for each genotype. The comparison between these clusters and those using 6 years of production (see 'Clustering of genotypes from global indices and ranking' above) was performed through a contingency table. Taking the latter clusters as a reference, an error rate was computed. The genotype ranks were updated and, taking the ranks obtained using 6 years of production as a reference, the rankings were compared using Kendall's τ coefficient (see Supplementary material E at *JXB* online).

To investigate the robustness of our approach, the stability of the classes and rankings was also assessed by using 3 years of yields (years 2006–2008 or 2008–2010) only, and using either the number of inflorescences or fruit mass.

Quantitative trait loci (QTL) mapping

At the term of model definition and index construction, six descriptors were used for QTL detection. Two groups of descriptors were distinguished: (i) those linked to biennial bearing at the global scale (global descriptors), which were the BBI using residuals of trend model (BBI_norm) with normalization by total yields (BBI_res_norm) and the genotype AR coefficient (γ_g); and (ii) those linked to the local scale (local descriptors), which were B^{loc}, γ^{loc}, and entropy.

QTL analyses were performed using the STK×GS consensus genetic map (Guitton *et al.*, 2012). QTL mapping was carried out using MapQTL® 5.0. (Van Ooijen, 2004) through a classical analysis including a permutation test to determine the LOD threshold (genome-wide error rate of 0.05), an interval mapping analysis to detect potential genomic regions associated to the trait, a cofactor selection, and a multiple QTL mapping (MQM) analysis.

Results

Quantification of alternation at the global scale

First, we present the models fitted using the global scale dataset, including year 2010. The random effect $\zeta_{g,r}$ (replication-specific random intercept) was assessed as non-significant in trend models A and B (AIC=16 902; BIC=18 208 in model A) and new models without $\zeta_{g,r}$ were finally estimated (AIC=16 845; BIC=18 138 in model A). Selection of a trend model among fixed-effects model A and random-effect model B was strongly in favour of model A rather than model B (AIC=18 972; BIC=18 998). The random effect $\gamma_{g,r}$ (replication-specific random deviation from the AR coefficient) was assessed as non-significant in residual AR1 model II (BIC=16 058), compared with residual AR1 model I without random effect $\gamma_{g,r}$ (BIC=16 036). These results appeared as an *a posteriori* confirmation of our preference of modelling genotypes as fixed effects, as in model A. The constraint of Gaussian slopes in model B treating genotypes as random effects led some estimated slopes to be too far from the measurements. This produced incoherent slope estimates for some tree replicates, and thus incoherent indices for the corresponding genotypes. The estimates of fixed effects and variances are given in Table 1. Confidence intervals are also given at the 95% level, computed using a bootstrap approach (Baayen *et al.*, 2008).

Values fitted for trended AR model (A-I) are given by $\hat{Y}_{g,r,t} = \beta + \beta_g + (\alpha + \alpha_g + E[\xi_{g,r}|\boldsymbol{Y} = \boldsymbol{y}])t$, where $\boldsymbol{Y}$ denotes the full vector of measurements $(Y_{g,r,t})_{g,r,t}$. The estimated model is illustrated in Figs 1–3 for genotypes with contrasted bearing habits. Figure 1 represents the fitted values in trend model A (Fig. 1a) and fitted residuals in residual AR1 model I (Fig. 1b) for genotype 85 representative of regular bearing. The residuals as a function of year are depicted in Supplementary Fig. S3 at *JXB* online. The indices and coefficients were BBI=1.14, BBI_res_norm=0.45, and γ_g=–0.07, this value indicating that the residuals $\varepsilon_{g,r,t}$ for genotype 85 were independent.

Table 1. *Estimates of the parameters of the trend and AR models using years 2005–2010*

The parameters are given for the genotype-independent fixed effects and variances (with 95% confidence intervals). The means and SD are given for the genotype-dependent fixed effects α_g of the trend model and γ_g of the AR1 model. Intercept β, residual SD α of the residuals ε and ρ of the AR model are expressed as number of flowers. Slope α and random effect standard deviation τ_ξ are given as number of flowers per year.

Parameter	Trend model (A)					AR Model (I)	
	β	α	τ_ξ	σ	α_g	ρ	γ_g
Estimate (genotype-independent parameters)	43.1 (–108, 194)	90.7 (36, 145)	20.9 (13.7, 24.0)	178 (172, 185)		150 (134,172)	
Mean/SD (genotype-dependent parameters)					–9.23/48.16		–0.40/0.39

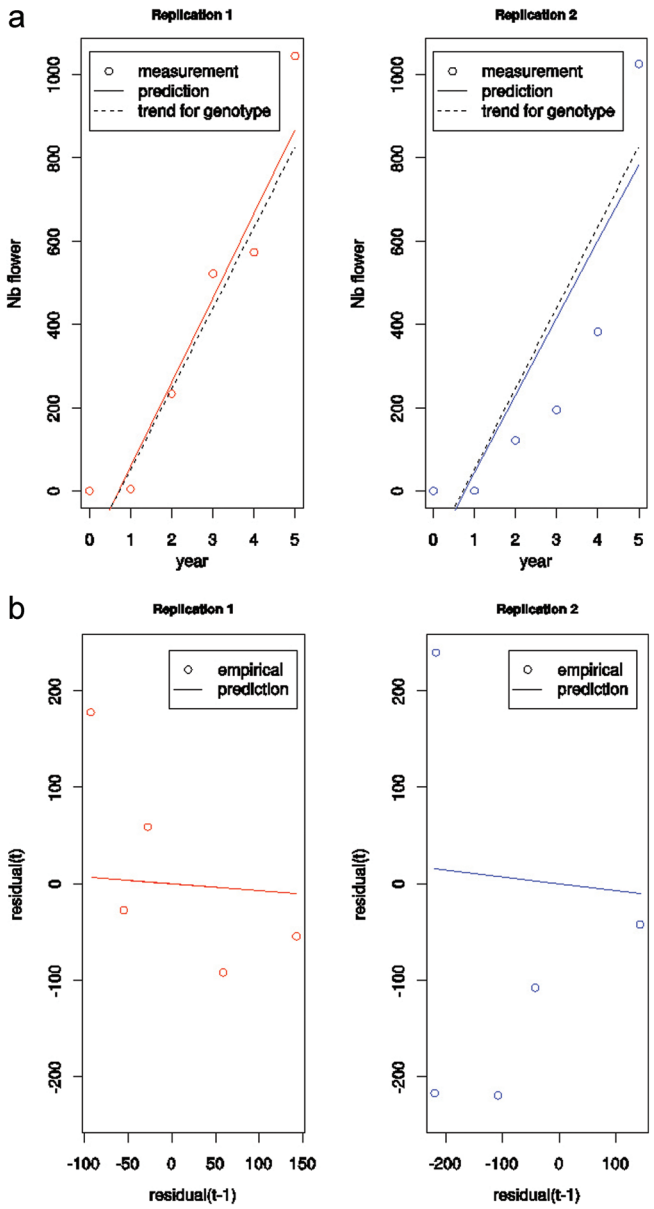

Fig. 1. Measurements and fitted values of yield $Y_{g,r,t}$ of replication *r* of genotype *g* at year *t* (a), empirical and predicted residuals as a function of previous residual (b) for regular bearing genotype *g*=85. (This figure is available in colour at *JXB* online.)

Figure 2 represents the fitted values and residuals for genotype 107 representative of a biennial bearing. The residuals as a function of year are depicted in Supplementary Fig. S4 at *JXB* online. The indices and coefficients were BBI=1.43, BBI_res_norm=1.21, and γ_g=–0.88, this indicating that the residuals alternate over time.

Figure 3 represents the fitted values and residuals for genotype 108 which is representative of irregular bearing. The residuals as a function of year are depicted in Supplementary Fig. S5 at *JXB* online. The indices and coefficients were BBI=1.10, BBI_res_norm =1.17, and γ_g=–0.28. These values of BBI_res_norm and γ_g are characteristic of unstructured residuals, the variability of which is high with regard to production. Genotypes 85 and 108 were quite similar from the point of view of BBI, however.

Clusters of genotypes

Using BIC as a criterion to select the number of clusters led us to consider models with two to four clusters, which fitted the data equally well. However, the three-cluster model was far more easily interpretable than the two- and four-cluster models, and had a slightly lower BIC value. The three-cluster model was thus selected (Fig. 4) and the clusters could be interpreted as follows:

- Cluster 1 contained 36 regular bearing genotypes. These genotypes mostly had values of genotype AR coefficient γ_g either positive or close to 0 (0.08 on average) and small values of BBI_res_norm (0.56 on average). Genotype 85 belonged to this cluster.
- Cluster 2 contained 31 biennial bearing genotypes. These genotypes mostly had negative values of γ_g (–0.82 on average) and high values of BBI_res_norm (1.47 on average). Genotype 107 belonged to this cluster. Note that, in absolute terms, 1.47 should be considered as a rather high value of BBI_res_norm. It is shown in proposition P4 in Supplementary material that if yield is linearly growing with biennial alternation, BBI_res_norm tends to 2 as the length of the time series increases.
- Cluster 3 contained the other 55 irregular bearing genotypes. These genotypes mostly had intermediate values of

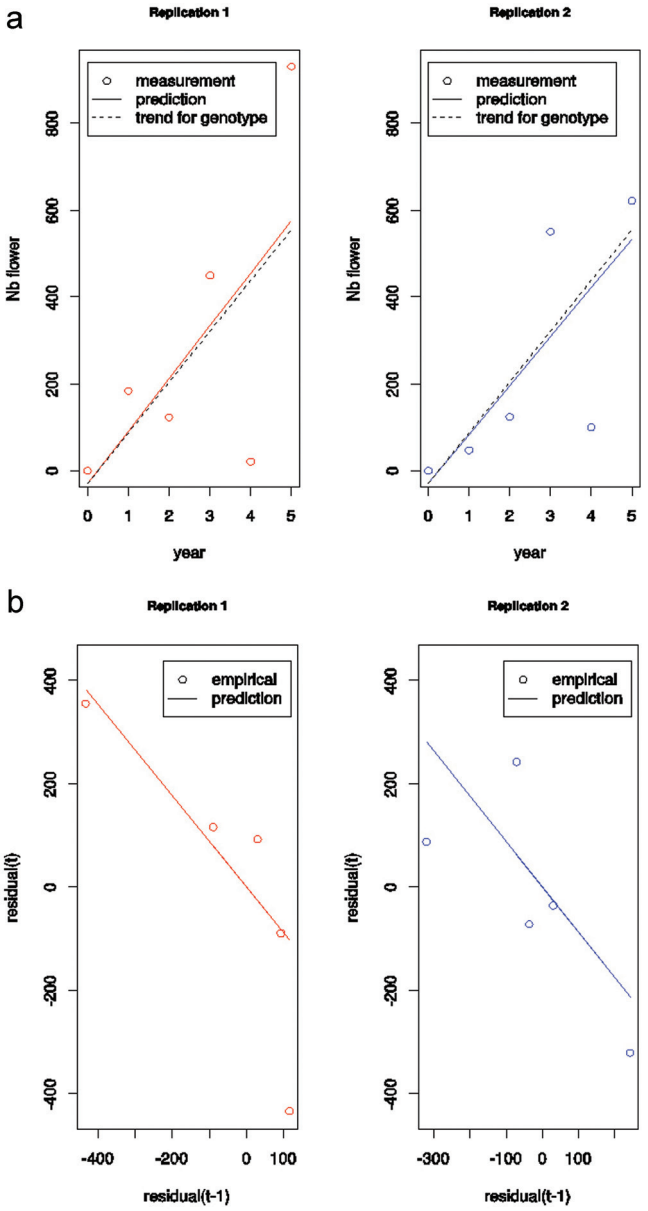

Fig. 2. Measurements and fitted values of $Y_{g,r,t}$ of replication *r* of genotype *g* at year *t* (a), empirical and predicted residuals as a function of previous residual (b) for biennial bearing genotype *g*=107. (This figure is available in colour at *JXB* online.)

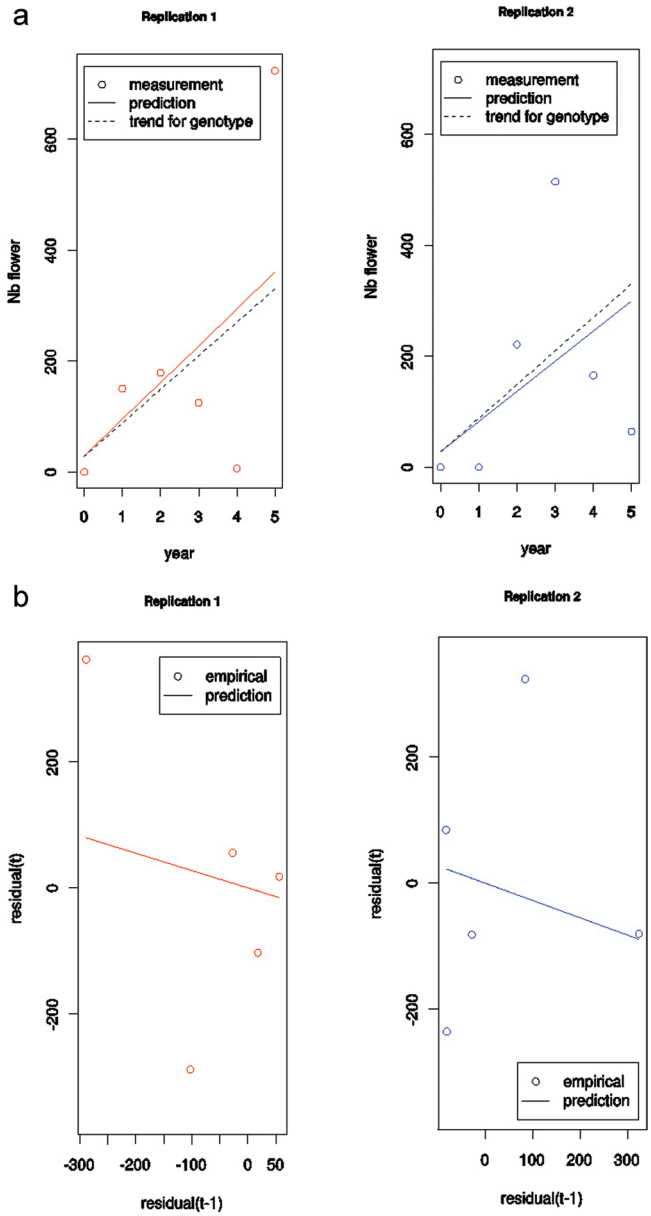

Fig. 3. Measurements and fitted values of $Y_{g,r,t}$ of replication *r* of genotype *g* at year *t* (a), empirical and predicted residuals as a function of previous residual (b) for (non-biennial) irregular bearing genotype *g*=108. (This figure is available in colour at *JXB* online.)

γ_g (–0.48 on average) and intermediate values of BBI_res_norm (0.95 on average). Genotype 108 belonged to this cluster.

Consequently, the three-cluster model allowed regular, irregular, and biennial bearing genotypes to be clearly discriminated, whereas in the two-cluster model, the irregular genotypes would be distributed between the irregular and biennial bearing genotypes. The four-cluster model would comprise an additional cluster containing the two genotypes with highest values of γ_g, which would not substantially improve the interpretation of the data. Within the cluster of regular (respectively, alternate bearing) genotypes, the correlations between fruit mass and number of inflorescences at successive years (given in Supplementary material C at *JXB* online) were shown to be significantly positive (respectively, negative) at the 5% level.

Model validation

The estimates for the fixed effects and variances of the model estimated on the first 5 years of production only and thus excluding year 2010 (referred to as the prediction model) are given in Table 2. It can be noted that the estimates were within

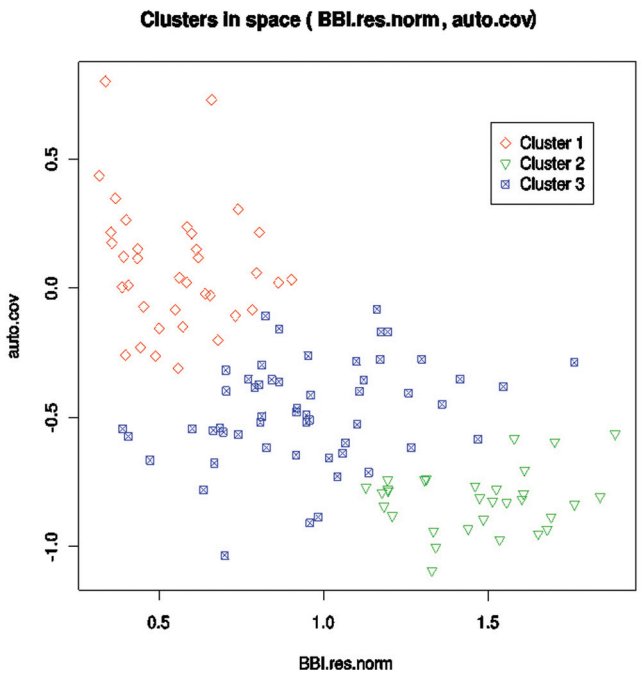

Fig. 4. Clustering obtained using a three-component mixture of Gaussian distributions of BBI_res_norm (*x*-axis) and genotype AR coefficient γ_g (*y*-axis named *auto.cov*). Cluster 1 can be interpreted as regular bearing genotypes, cluster 2 as biennial bearing genotypes, and cluster 3 as irregular bearing genotypes. (This figure is available in colour at *JXB* online.)

the confidence intervals given in Table 1 (estimates using the whole dataset).

Prediction intervals for the sixth year yield (2010) were computed for each replication of each genotype, at level 0.95. The frequency of actual yields in 2010 that were within the prediction interval was 0.74. This shows that the variance of the number of flowers in 2010 tends to be underestimated by the models. This is consistent with the comparison between Tables 1 and 2, which shows that the variance parameters σ^2 and ρ^2 increased if year 2010 is considered in the analysis. As expected, the percentage of observations within the prediction interval was higher in regular or biennial bearing genotypes than in irregular genotypes. This results from the trended AR model (A-I) being essentially dedicated to modelling regular and biennial flowering. In contrast, irregular bearing can be seen as poor adequacy to this model (i.e. low autocorrelation but high residuals). An illustration of the predicted number of flowers in 2010 for genotype 107 (biennial bearing) is provided in Fig. 5. The predictions associated with genotypes 85 (regular) and 108 (irregular) are provided in Supplementary Fig. S6 at *JXB* online.

Changes in the estimates in the model parameters and random effects caused switches in the clusters of several genotypes. Taking clusters obtained using the 6 years of production as a reference (Fig. 4), five genotypes switched from irregular to regular bearing (among which two were not significant since they lay at the boundary of both clusters), five switched from irregular to biennial bearing, and six switched from biennial bearing to irregular. Only one regular genotype switched to biennial bearing, and seven switched to irregular (among which two were not significant). No biennial bearing genotype was assessed as regular. The class switches can be summed up by a contingency table (Supplementary Table S2 at *JXB* online). In particular, the clusters of the three genotypes illustrated in Figs 1–3 did not change. Kendall's τ coefficient was 0.70.

The results based on 3 years of yields (years 2006–2008 or 2008–2010), and using either the number of inflorescences or fruit mass are given in Supplementary material E at *JXB* online (see also Supplementary Tables S3–S8). These analyses showed that using fruit mass instead of the number of inflorescences, or using 3 or 5 years of production instead of 6 years, generated the same results, qualitatively. The confusion between clusters of genotypes as well as the number of rank inversions tended to increase when using the last 3 years of yields, to increase further when using the first 3 years of yields, and to increase still further when using fruit mass. The confusion involves majority switches from or to the cluster of irregular genotypes. When using fruit mass, however, switches between regular and alternate bearing genotypes tend to occur more often.

Quantification of alternation using local indices

Gaussian linear models with predictors B^{loc}, γ^{loc}, and $\overline{Ent}_g$ were estimated to assess their correlations with BBI_res_norm

Table 2. *Estimates of the parameters of the trend and AR models using years 2005–2009*

The parameters are given for the genotype-independent fixed effects and variances (with 95% confidence intervals). The means and SD are given for the genotype-dependent fixed effects α_g of the trend model and γ_g of the AR1 model. Intercept β, residual standard deviations σ of the residuals ε and ρ of the AR model are expressed as number of flowers. Slope α and random effect standard deviation τ_ξ are given as number of flowers per year.

Parameter	Trend model (A)					AR Model (I)	
	β	α	τ_ξ	σ	α_g	ρ	γ_g
Estimate (genotype-independent parameters)	44.2 (–101, 188)	90.0 (29, 154)	23.0 (10.9, 25.8)	160.4 (155, 168)		138.7 (123, 159)	
Mean/SD (genotype-dependent parameters)					–9.73/62.84		0.19/0.44

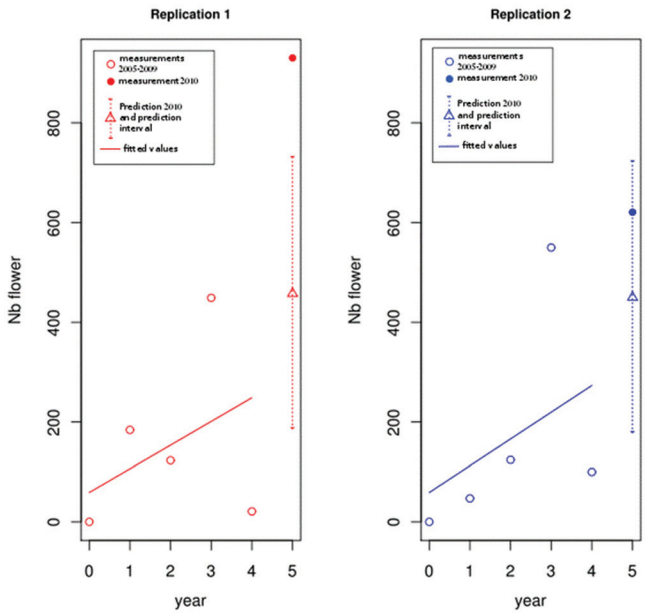

Fig. 5. Measurements and predicted yield values in 2010 (year number 5) for irregular genotype *g*=107. Circles are the measured values, and triangles are the predicted values, located in the middle of prediction intervals (dotted segments). (This figure is available in colour at *JXB* online.)

and γ_g. The predictor $\overline{Ent}_g$ was not assessed as significant to predict either BBI_res_norm (*P* value of the *t*-test on nullity of coefficient: 0.98) or γ_g (*P* value of the *t*-test: 0.37). Excluding $\overline{Ent}_g$ from the set of predictors, the coefficient of multiple correlation was 0.74 in predicting BBI_res_norm and –0.65 in predicting γ_g (with *P* values of the overall F-test for regression smaller than 1e–13 in both cases). The correlations of each predictor considered separately with BBI_res_norm and γ_g are given in Table 3. Every correlation was significantly positive at $P<0.05$.

The correlations of mean entropy $\overline{Ent}_g$ with the quantities characterizing regularity of flowering at the whole-tree scale (BBI_res_norm and γ_g) suggested that the genotypes with highest synchronicity (lowest value of mean entropy) tend to be of biennial bearing type (high values of BBI_res_norm and low values of γ_g). These genotypes also tended to show alternation at the local scale (B^{loc} and γ^{loc}). An analysis of variance performed on $\overline{Ent}_g$ to assess its separation on the three clusters provided a *P* value of 1e–6, which proves a significant separation of $\overline{Ent}_g$ with respect to clusters. Considering the means, the biennial bearing genotypes (mean entropy: 0.23) could be discriminated from the irregular (mean entropy: 0.38) and the regular bearing genotypes (mean entropy: 0.38). However, the latter two categories could not be separated by their mean entropies.

During the global model validation step (see 'Model validation' above), the correlation matrix between the indices at both scales (Table 3) was updated. The updated matrix is given in Supplementary Table S9 at *JXB* online. The correlations were within the confidence intervals given in Table 3, which shows that the above conclusions do not qualitatively change if year 2010 is included or excluded from the analysis.

A Gaussian mixture clustering was used to identify clusters from the local indices, as in 'Model validation' above. This approach did not benefit from the knowledge of the clusters obtained using the global indices, nor from the relations between the global indices and these clusters. Hence, its performance should be equivalent to clustering new progenies, from different parents from 'Starkrimson' and 'Granny Smith'. The error rate was 0.4, which must be compared with that of the worst possible classifier (best random classifier independent of the three indices), which is necessarily larger (0.56 here). The confusion between clusters is illustrated by the contingency Table 4.

If a new cross between the same 'Starkrimson' and 'Granny Smith' parents was considered, in contrast, the statistical distributions of the indices and clusters of these genotypes would be comparable to those presented here. The ability of the three local indices to accurately predict the bearing behaviour of future genotypes in such a context could be assessed by the cross-validated error rate of supervised classification. The cross-validated error rate obtained using support vector machines was 0.38, which is only slightly better than unsupervised classification (i.e. Gaussian mixture clustering). This should be considered as a quantification of the accuracy of the best possible predictor of the bearing behaviour from local indices. The cross-validated error rate provided by neural networks was 0.41. The errors systematically involved confusion between irregular bearing genotypes on the one hand, and either regular or biennial bearing genotypes on the other hand, as illustrated by the contingency Supplementary Table S10 at *JXB* online. A visual representation the overlap between the classes of genotypes, as characterized by the three local indices, is provided in Supplementary Fig. S7 at *JXB* online by the FDA. In such representation where the classes

Table 3. *Correlation coefficients between indices at global and local scales, with 95% confidence intervals*

	Genotype AR coefficient γ_g	Local BBI_res_norm B_g^{loc}	Local genotype AR coefficient γ_g^{loc}	Mean entropy $\overline{Ent}_g$
BBI_res_norm	–0.66 (–0.75, –0.55)	0.72 (0.61, 0.80)	–0.58 (–0.69, –0.45)	–0.49 (–0.62, –0.34)
γ_g	1	–0.61 (–0.71, –0.49)	0.55 (0.41, 0.67)	0.33 (0.16, 0.48)
B_g^{loc}		1	–0.63 (–0.73, –0.51)	–0.66 (–0.75, –0.54)
γ_g^{loc}			1	0.33 (0.15, 0.48)

Table 4. *Contingency table for the clusters of each genotype. The clusters are R(egular), A(lternate) and I(rregular) bearing*

		Cluster based on local indices		
		R	A	I
Cluster based on global indices	R	33	0	2
	A	2	18	9
	I	29	4	18

Clusters in lines correspond to those obtained using the global indices (reference clusters to compute an error rate). Clusters in columns correspond to those obtained using the local indices.

are optimally separated, the irregular genotypes seemed to be uniformly distributed between both other classes.

QTL mapping

Concerning the descriptors at the whole-tree scale, altogether four QTLs were identified, for BBI_norm and BBI_res_norm, in two separated genomic regions (Fig. 6), on LG1 and LG8 respectively. No significant QTL was detected for genotype AR coefficient γ_g, by contrast. QTLs for BBI_norm and BBI_res_norm were characterized by high LOD scores, from 5.27 to 6.19 (Table 5). The same cofactors were used for both variables, MdGA20ox1a_S and MdEFL3a_G, respectively, for LG1 and LG8 (Table 5). The trait variability explained by these QTLs ranged from 17.7 to 20%.

Concerning the descriptors at the local scale, one QTL was mapped for B^{loc} on LG8, with an LOD score of 4.32 (Table 5) with the cofactor CH01c06_G. This QTL is located in the same region as the QTL mapped for BBIs at the global scale, since the cofactor CH01c06_G is located less than 1 cM away from the marker MdEFL3a_G. Two QTLs were also detected for genotype AR coefficient at the local scale γ^{loc}. These QTL, located on LG11 and LG14, had high LOD scores (7.24 and 4.55, respectively), and explained 22.5 and 13.5% of the total variability, respectively. By contrast, no QTL was detected for entropy.

Discussion

Phenotyping method, genetic determination, and practical applications for breeding

In the present study, we investigated biennial bearing using the number of inflorescences, even though previous studies have been usually focused on the number or the mass of harvested fruits (Cilas *et al.*, 2011). This choice resulted from the fact that the number of inflorescences is less subject to environmental variation than variables related to fruits, e.g. fruit set, and fruit drop caused by climatic conditions or pests and diseases. Consistently, when the same models, as for the number of inflorescences, were applied to harvested fruit variables, the tree replicate and the residual variability was higher than the genetic variability (data not shown). Similarly, the confusion between clusters and the degradation in the rankings increased when using fruit mass instead of the number of inflorescences (see Supplementary material E at *JXB* online). We could thus conclude that phenotyping the number of inflorescences per tree is more efficient for quantifying the genetic parameters of biennial bearing in apple.

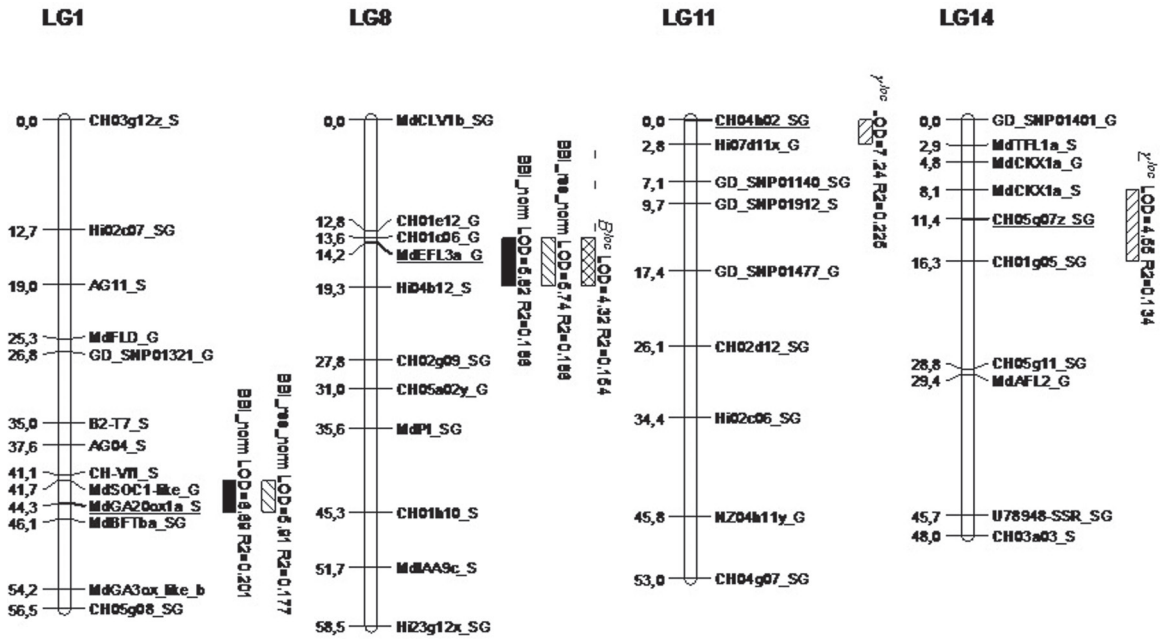

Fig. 6. Genomic positions of the QTLs detected on the consensus 'Starkrimson'×'Granny Smith' (STK×GS) genetic map. QTLs are represented by boxes, whose length represents the LOD-1 confidence interval. Parameters at the global scale are BBI_norm (BBI normalized by total yield) and BBI_res_norm (using residuals of trend model). Parameters at the local scale are B^{loc} and γ^{loc} (respective BBI_res_norm and genotype AR coefficient computed from data at local scale). Underlined genetic markers correspond to cofactors used for QTL mapping.

Table 5. *QTLs detected on the consensus STK×GS genetic map by MQM mapping for variables related to biennial bearing at global (BBI_norm and BBI_res_norm) and local scales (B^{loc} and γ^{loc})*

Variable	LG	LOD	R^2	Cofactor
BBI_norm	1	6.69	0.201	MdGA20ox1a_S
	8	5.82	0.186	MdEFL3a_G
BBI_res_norm	1	5.91	0.177	MdGA20ox1a_S
	8	5.74	0.186	MdEFL3a_G
B^{loc}	8	4.32	0.154	CH01c06_G
γ^{loc}	11	7.24	0.225	CH04h02_SG
	14	4.55	0.134	CH05g07z_SG

The table displays the linkage group (LG) the QTL has been detected on, the LOD score (LOD), the proportion of genetic variability (R^2) explained by the QTL and the cofactor used for the QTL mapping.

This suggests that, in apple trees, biennial behaviour is more closely linked to floral induction than to pollination, fruit set, or fruit drop. This is consistent with the assumption that gibberellins synthesized by the developing fruits and/or competition for carbohydrates between fruits and shoots could inhibit flower induction (Monselise and Goldschmidt, 1982; Bangerth, 2009). This finding was re-enforced by co-localizations found between QTL for annual yield with positions of hormone-related genes such as GA20-ox and GA3-ox-like (Guitton *et al.*, 2012). However, the physiological causes of biennial bearing pattern are likely to differ among fruit-tree species, since yield correlates more to fruit set than to flowering occurrence in other crops, for instance in the olive tree (Ben Sadok *et al.*, 2013).

Phenotyping at the local scale constitutes a faster strategy and is made easier by the retrospective observation of flowering occurrences in the apple, due to the terminal position of flowering and the possibility of distinguishing bourse from vegetative growth units. Similar phenotyping could be applied to other fruit-tree species with terminal flowering and an easy observation of flowering sites, such as pear (Huet, 1972) or walnut (Sabatier *et al.*, 1998). However, it may not be applicable to species with lateral flowering such as *Prunus* species (Fournier, 1994; Costes *et al.*, 2006), due to the lack of morphological markers for a retrospective observation of flowering occurrences.

QTL detection was undertaken, using these descriptors as quantitative traits. This led us to identify genomic regions involved in biennial bearing. Indeed, QTL were detected in two genomic regions for BBI_norm and BBI_res_norm, which corroborate zones previously identified (Guitton *et al.*, 2012). The QTL cluster on LG1 for BBI_norm (global scale), and BBI_res_norm (global and local scale) also co-located with QTLs for inflorescence yield of a given year and fruit yield QTLs for the year before. This LG1 QTL seems to be linked to the antagonism between fruit production of the current year and inflorescence development for the year after, as reported by Bangerth (2009).

The two QTL revealed on LG11 and LG14 for the autocorrelation index at the local scale were located on zones that were not previously associated with flowering or bearing traits in this progeny, even though QTL were detected in the same regions for branching and internode length, respectively (Segura *et al.*, 2009). The recent discovery of interactions between florigen and genes inhibiting lateral shoot growth (Niwa *et al.*, 2013) suggest a possible common mechanism underlying these associations. The absence of QTL detection for the same descriptor, at the tree scale and for the entropy may be due to the population size. Indeed, 120 individuals can be limiting to detect QTL with small effects (Bernardo, 2004), and only major QTL were confirmed in the present study. However, an LOD peak (3.04) was observed for the autocorrelation index on the global scale on LG8 at the same position as for BBI_res and BBI_res_norm, but was lower than the threshold at 95% (3.7). Similarly, a QTL for BBI_res_norm at local scale was found on LG1 but was below the threshold. Both QTL were not considered in our results.

The practical application of the proposed method in breeding programmes is still challenging because quantifying irregular bearing implies the collection of data over consecutive years. Phenotyping based on BBI_res_norm and the genotype AR coefficient calculated on the number of inflorescences per tree will certainly not be performed by breeders. Phenotyping at the local scale is more likely to be feasible, because these retrospective measures would be less time-consuming than weighing the production tree by tree over several years. However, to make it more acceptable, an estimation of fruiting behaviour of genotypes could be performed in the last steps of the selection scheme only, on pre-selected genotypes (i.e. genotypes that have overcome the first individual selections) (Laurens *et al.*, 2000).

From our results, it is suspected that, after a regular increase in yield, some genotypes would start exhibiting irregular or biennial bearing at some age (unknown in advance). This hypothesis is supported by Smith *et al.* (2004), who noticed while studying biennial bearing on citrus that there was a clear increase in the intensity of the biennial bearing pattern as the trees aged. This is also supported by the increase in the number of changes in clusters and the degradation in rankings caused by using the first 3 years, the last 3 years, or the last 5 years of yields, respectively, to compute the indices at the global scale. This shows that the last 3 years are of great importance in the model and that the first 3 years, associated with low yield for every genotype, had no significant discriminating power. Moreover, prediction of yields in 2010 (see 'Model validation' above) showed that, although an increase of yield variability was anticipated by the model, this increase was underestimated. This suggests that models considering more complex correlation structures could be further tested.

Moreover, a good discrimination between regular and biennial bearing genotypes was obtained by unsupervised classification based on local descriptors, i.e. with no *a priori* information related to the clusters from global descriptors. This suggests that retrospective measurements on annual shoots could be performed on new families issuing from different parents than 'Starkrimson' and 'Granny Smith'.

As a consequence, we propose that breeding for regular genotypes could be performed in two steps: (i) during the beginning of the mature phase when tree production is increasing, progressively suppressing biennial or irregular genotypes after the first observation of a large decrease in flowering; and (ii) to confirm the regular fruiting behaviour of the pre-selected genotypes after the production and the intensity of biennial pattern have stabilized, through phenotyping at the local scale. It must be noticed that such measures at the local scale would allow practically every biennial bearing genotype to be discarded. Moreover, cluster validation using 3–5 years with local or global descriptors and FDA have shown a strong overlap between the cluster of irregular genotypes and both other clusters. Thus, part of the irregular genotypes would also be discarded by this approach. Eventually, the methods proposed here could make the estimation process more robust, even though the robustness and reliability of model predictions will depend on the number of years considered.

Further research could investigate the possibility of including both steps in a single modelling approach. Presently, the trend model is dedicated to a period beginning with the mature phase and ending just before the average production is stabilized. Thus, the data cannot include this stable period to avoid any bias in the model parameter and in the quantification of biennial bearing. To address such issues, a possibility would be to resort to a regime switching model, including two stationary regimes: one regime with trend (at the beginning of mature phase) and one regime without trend (after the production is stabilized on average). Each regime would be associated with the state of a Markovian model, which would model the times of switches between regimes, as proposed previously by Chaubert-Pereira *et al.* (2009).

New descriptors: limits and possible extensions or variants

We have proposed three descriptors having a biological meaning to quantify biennial bearing that are defined at either the global or the local scale, and are deduced from models, either directly (parameter γ_g of a linear mixed model) or indirectly (using estimated residuals in the case of BBI_res_norm). In the same way, the positive intercept on the x-axis of the trend model could be used to predict precocity.

At the global scale, Cilas *et al.* (2011) estimated the genotypic values associated with yield by linear mixed models with residual correlations having a genotype-invariant structure. A compound symmetry model (i.e. with constant correlation between years) with heterogeneous variances was assessed to achieve the best fit to the data. To assess regularity of flowering in the prospect of deriving heritability values and correlations between traits, classical indices (BBI and the number of sign changes in the difference between consecutive yields) were computed on yield directly, independently from the linear mixed model. Their approach was justified by the particular context of their study, where yield was not affected by a trend. Our approach, in contrast, aims at estimating the bearing behaviour of new genotypes during their first years of fruiting while yield increases, and each genotype has a specific correlation structure between consecutive years. This led us to estimate residuals from linear mixed models and incorporate them into a modified version of the BBI (BBI_res_norm). Our results (see 'Quantification of alternation at global scale' above and Supplementary material R at *JXB* online) highlighted that BBI is not able to discriminate irregular from alternate bearing genotypes, and confirmed that this index should not be used to quantify irregularity of production on trended data. However, it must be noticed that, like BBI, BBI_res_norm is sensitive to the residual variance under the hypothesis of random fluctuations. Thus, generally, its value for a given individual cannot be interpreted directly, and it must be either compared with the values for other individuals of similar nature (as has been the case in our work), or compared to its expected value under some given hypothesis, for example using a resampling approach, as in Huff (2001). Moreover, because genotypes with different behaviours may have quite similar values of BBI_res_norm (see 'Quantification of alternation at global scale' above), its association with the AR coefficient is necessary to discriminate biennial bearing from irregular genotypes.

At the local scale, an entropy index has been proposed to characterize synchronicity in flowering. This proposal was initially inspired by Lauri *et al.* (1995), who summarized the information of 5–6 years sequences by an alternation synchronicity index that characterized synchronous patterns of alternation over two successive years. We instead propose to study full-length sequences and to use repetition of sequences within the tree in order to estimate genetic parameters related to alternation. As a complement, the new entropy index aims at characterizing how frequently the shoots are either simultaneously flowering the same year or stay vegetative the same year. Entropy indices could be computed at the tree scale, but this would require measurement of the annual total number of vegetative shoots per tree.

It can be noted that the subsamples of sequences contain structural information that is not fully exploited by the local indices. Successive shoots along sequences are subject to dependencies (which depend *a priori* on the bearing behaviour) that could be taken as a benefit. Markov models could achieve this goal, as illustrated by Costes and Guédon (2012), still in the case of apple trees, but for sequences of growth units. Their approach may be adapted to let the parameters of this statistical model depend on the genotypes, and to predict classes of bearing behaviour using a different Markov model for each class.

The analysis of correlations between global and local descriptors of alternation and synchronicity leads us to hypothesize that biennial bearing at the tree scale results from the conjunction of two phenomena: synchronicity in flowering at the global scale between shoots in a given year, and biennial alternation at the local scale between successive years. Regular bearing genotypes at the global scale, in contrast, can be obtained from either asynchronous alternating flowering or from regular flowering at the local scale. Irregular

genotypes exhibit intermediate values for each of the five descriptors, related to bienniality and synchronicity. Thus, it can be hypothesized that these genotypes are characterized by partial biennial alternation at the local scale and partial synchronicity. However, more complex within-tree organization of synchronicity could exist, especially at the branch scale, as proposed previously for a range of fruit species (Couranjou, 1978; Monselise and Goldschmidt, 1982). Such complex patterns were not considered in the present study.

Conclusion

In the present study, we revisited the indices and descriptors available for estimating the bearing behaviour of fruit trees in a breeding context, i.e. at the genotype scale. Using an innovative modelling approach, we have proposed new descriptors at both whole-tree and annual shoot scales. This approach opens up new possibilities to accelerate evaluation processes and allow a first diagnostic during the first years of tree production when yield is still increasing. Further investigations could contribute to testing multivariate classifying methods and to extending the estimation at the genotype scale to the entire mature phase of trees.

Supplementary data

Supplementary data are available at *JXB* online.

Supplementary Fig. S1. Schematic representation of observations performed on 6-years-old trees on the trunk, long sylleptic axillary shoot (LSAS), long proleptic axillary shoot (LSAS) and short axillary shoot (SAS).

Supplementary Fig. S2. Absolute difference between consecutive yields $|Y_t - Y_{t-1}|$ as a function of Y_t.

Supplementary Fig. S3. Empirical and predicted residuals of yields as a function of time for regular bearing genotype g=85.

Supplementary Fig. S4. Empirical and predicted residuals of yields as a function of time for biennial bearing genotype g=107.

Supplementary Fig. S5. Empirical and predicted residuals of yields as a function of time for irregular bearing genotype g=108.

Supplementary Fig. S6. Measurements and predicted yield values in 2010 (year number 5) for regular bearing genotype g=85 (a) and irregular bearing genotype g=108 (b).

Supplementary Fig. S7. Plot of genotypes in the first FDA plane, based on mean entropy and local indices B^{loc} and local genotype AR coefficient γ^{loc}.

Supplementary Fig. S8. Measurements and fitted values of number of inflorescences for regular bearing genotype g=123.

Supplementary Fig. S9. Measurements and fitted values of number of inflorescences for irregular bearing genotype g=5.

Supplementary Fig. S10. Measurements and fitted values of fruit mass for regular genotype g=85.

Supplementary Fig. S11. Measurements and fitted values of number of inflorescences (a) and fruit mass (b) for regular genotype g=62.

Supplementary Fig. S12. Measurements and fitted values of number of inflorescences (a) or fruit mass (b) for regular genotype g=64.

Supplementary Table S1. Computation of entropies to quantify synchronism in flowering for three genotypes g: regular bearing (g=85), biennial bearing (g=107) and irregular bearing (g=108).

Supplementary Table S2. Contingency table for the clusters of each genotype obtained using either years 2005–2010 or years 2005–2009.

Supplementary Table S3. Contingency table for the clusters of each genotype obtained using either years 2005–2010 or years 2006–2008.

Supplementary Table S4. Contingency table for the clusters of each genotype obtained using either years 2005–2010 or years 2008–2010.

Supplementary Table S5. Contingency table for the clusters of each genotype obtained using either the number of inflorescences or fruit mass.

Supplementary Table S6. Contingency table for the clusters of each genotype obtained using either the number of inflorescences and years 2005–2010 or fruit mass and years 2006–2008.

Supplementary Table S7. Contingency table for the clusters of each genotype obtained using either the number of inflorescences and years 2005–2010 or fruit mass and years 2008–2010.

Supplementary Table S8. Contingency table for the clusters of each genotype obtained using either BBI_res_norm and the genotype AR1 coefficient, or BBI only.

Supplementary Table S9. Correlation coefficient between indices at whole tree and AS scales, with 95% confidence intervals.

Supplementary Table S10. Contingency table for the clusters of each genotype obtained using either the global or the local indices.

Supplementary Table S11. Precision $|BBI(t) - \alpha(t)|/BBI(t)$ of approximation of the BBI by $\alpha(t)=\log t/(t-1)$ in the case of affine growth of Y_t, as a function of the slope a and the length t of the time series.

Supplementary Table S12. Precision $|BBI\_res\_norm(t) - 2|/BBI\_res\_norm(t)$ of the approximation of BBI_res_norm by its limit 2 in the case of linear growth of alternate yield Y_t, as a function of the slope a and the length t of the time series.

Acknowledgements

The authors thank Christian Lavergne, Christian Cilas, and Pierre-Éric Lauri for fruitful discussions, Bernard Ycart and David Chagné for their critical reading of the manuscript and help with the English. The PhD scholarship of BG was supported by the Plant Breeding Department of the National Institute of Agronomic Research of France (INRA), by Plant & Food Research (Pipfruit Internal Investment Project), and by the New Zealand Ministry of Science and Innovation [Horticultural Genomics programme (CO6X0810)]. The PhD of JP was founded by the University of Montpellier 2 (UM2).

References

Baayen RH, Davidson DJ, Bates DM. 2008. Mixed-effect modelling with crossed random effects for subjects and items. *Journal of Memory and Language* **59,** 390–412.

Bangerth F. 2009. Floral induction in mature, perennial angiosperm fruit trees: similarities and discrepancies with annual/biennial plants and the involvement of plant hormones. *Scientia Horticulturae* **122,** 153–163.

Bates D, Maechler M, Dai B. 2011. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375–28: http://lme4.r-forge.r-project.org [accessed 5 October 2011].

Ben Sadok I, Celton J-M, Essalouh L, El Aabidine AZ, Garcia G, Martinez S, Grati-Kamoun N, Rebai A, Costes E, Khadari B. 2013. QTL mapping of flowering and fruiting traits in olive. *PLoS One* **8,** e62831.

Bernardo R. 2004. What proportion of declared QTL in plants are false? *Theoretical and Applied Genetics* **109,** 419–424.

Bishop CM. 2006. *Pattern recognition and machine learning* . New York: Springer Verlag.

Chaubert-Pereira F, Caraglio Y, Lavergne C, Guédon Y. 2009. Identifying ontogenetic, environmental and individual components of forest tree growth. *Annals of Botany* **104,** 883–896.

Cilas C, Montagnon C, Bar-Hen A. 2011. Yield stability in clones of *Coffea canephora* in the short and medium term, longitudinal data analyses and measures of stability over time. *Tree Genetics and Genomes* **7,** 421–429.

Cnaan A, Laird NM, Slasor P. 1997. Tutorial in biostatistics, using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in Medicine* **16,** 2349–2380.

Costes E, Guédon Y. 2012. Deciphering the ontogeny of a sympodial tree. *Trees—Structure and Function* **26,** 865–879.

Costes E, Lauri PE, Régnard, JL. 2006. Tree architecture and production. *Horticultural Reviews* **32,** 1–60.

Couranjou J. 1978. Studies on the genetic causes and mechanisms of alternate bearing in plum, *Prunus domestica* L. II. Effect of fruit load in one part of the tree on floral induction in the rest of the tree, from which the fruits have been removed; level of automony between the two parts according to cultivar. *Physiologie Végétale* **16,** 505–520.

Diggle P, Kenward MG. 1994. Informative drop-out in longitudinal data analysis. *Journal of the Royal Statistical Society C (Applied Statistics)* **43,** 49–93.

Fournier D. 1994. Analyse et modélisation des processus de croissance et développement qui contribuent aux performances agronomiques du pêcher *Prunus persica* (L.) Batsch. PhD thesis, ENSA Montpellier, France.

Guitton B, Kelner J-J, Velasco R, Gardiner SE, Chagné D, Costes E. 2012. Genetic control of biennial bearing in apple. *Journal of Experimental Botany* **63,** 131–149.

Hand DJ, Crowder MJ. 1996. *Practical longitudinal data analysis* . London: Chapman & Hall.

Hoblyn TN, Grubb NH, Painter AC, Wates BL. 1936. Studies in biennial bearing. *International Journal of Pomology and Horticultural Science* **14,** 39–76.

Huet J. 1972. Etude des effets des feuilles et des fruits sur l'induction florale des brachyblastes du Poirier. *Physiologie Végétale* **10,** 529–545.

Huff A. 2001. A significance test for biennial bearing using data resampling. *Journal of Horticultural Science and Biotechnology* **76,** 534–535.

Jonkers H. 1979. Biennial bearing in apple and pear: a literature survey. *Scientia Horticulturae* **11,** 303–307.

Karatzoglou A, Smola A, Hornik K, Zeileis A. 2004. kernlab—an S4 package for Kernel Methods in R. *Journal of Statistical Software* **11**(9), 1–20.

Laurens F., Audergon J-M, Claverie J, Duval H, Germain E, Kervella J, Lelezec M, Lauri P-É, Lespinasse J-M. 2000. Integration of architectural types in French programmes of ligneous fruit species genetic improvement. *Fruits* **55,** 141–152.

Lauri P-É, Térouanne E, Lespinasse J-M, Regnard J-L, Kelner J-J. 1995. Genotypic differences in the axillary bud growth and fruiting pattern of apple fruiting branches over several years—an approach to regulation of fruit bearing. *Scientia Horticulturae* **64,** 265–281.

Lauri P-É, Térouanne E, Lespinasse J-M. 1997. Relationship between the early development of apple fruiting branches and the regularity of bearing—an approach to the strategies of various cultivars. *Journal of Horticultural Science* **72,** 519–530.

McLachlan GJ, Peel D. 2000. *Finite mixture models* . New York: Wiley.

Monselise SP, Goldschmidt EE. 1982. Alternate bearing in fruit trees. *Horticultural Reviews* **4,** 128–173.

Neilsen JC, Dennis FG. 2000. Effects of seed number, fruit removal, bourse shoot length and crop density on flowering in 'Spencer Seedless' apple. *Acta Horticulturae* **527,** 137–146.

Niwa M, Daimon Y, Kurotani K, *et al.* 2013. BRANCHED1 interacts with FLOWERING LOCUS T to repress the floral transition of the axillary meristems in *Arabidopsis*. *Plant Cell* **25,** 1228–1242.

Pearce SC, Dobersek-Urbane S. 1967. The measurement of irregularity in growth and cropping. *Journal of Horticultural Science* **42,** 295–305.

R Development Core Team. 2008. *R, A language and environment for statistical computing* . Vienna: R Foundation for Statistical Computing.

Reddy YTN, Kurian RM, Ramachander PR, Singh G, Kohli RR. 2003. Long-term effects of rootstocks on growth and fruit yielding patterns of 'Alphonso' mango (*Mangifera indica L.*). *Scientia Horticulturae* **97,** 95–108.

Rosenstock TS, Rosa UA, Plant RE, Brown PH. 2010. A reevaluation of alternate bearing in pistachio. *Scientia Horticulturae* **124,** 149–152.

Sabatier S, Barthélémy D, Ducousso I, Germain É. 1998. Modalités d'allongement et morphologie des pousses annuelles chez le noyer commun, *Juglans regia L.* 'Lara' (Juglandaceae). *Canadian Journal of Botany* **76,** 1253–1264.

Segura V, Cilas C, Laurens F, Costes E. 2006. Phenotyping progenies for complex architectural traits, a strategy for 1-year-old apple trees (*Malus×domestica* Borkh.). *Tree Genetics and Genomes* **2,** 140–151.

Segura V, Denancé C, Durel C-E, Costes E. 2007. Wide range QTL analysis for complex architectural traits in a 1-year-old apple progeny. *Genome* **50,** 159–171.

Segura V, Durel C-E, Costes E. 2009. Dissecting apple tree architecture into genetic, ontogenetic and environmental effects, QTL mapping. *Tree Genetics and Genomes* **5,** 165–179.

Singh L. 1948. Studies on biennial bearing. III. Growth studies on 'on' and 'off' year trees. *Journal of Horticultural Science* **24,** 123–148.

Smith MW, Shaw RG, Chapman JC, Owen-Turner J, Slade Lee L, Bruce McRae K, Jorgensen KR, Mungomery WV. 2004. Long-term performance of 'Ellendale' mandarin on seven commercial rootstocks in sub-tropical Australia. *Scientia Horticulturae* **102,** 75–89.

Tabachnick B, Fidell L. 2007. *Using multivariate statistics* . New York: Harper and Row Publishers.

Van Ooijen JW. 2004. MapQTL® 5, software for the mapping of quantitative trait loci in experimental populations. Wageningen, The Netherlands: Kyazma B.V.

Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S* , 4th edn. New York: Springer.

Verbeke G, Molenberghs G. 2000. *Linear mixed models for longitudinal data* . New York: Springer.

Wilcox J. 1944. Some factors affecting apple yields in the Okanagan Valley, tree size, tree vigor, biennial bearing, and distance of planting. *Scientific Agriculture* **25**, 189.