

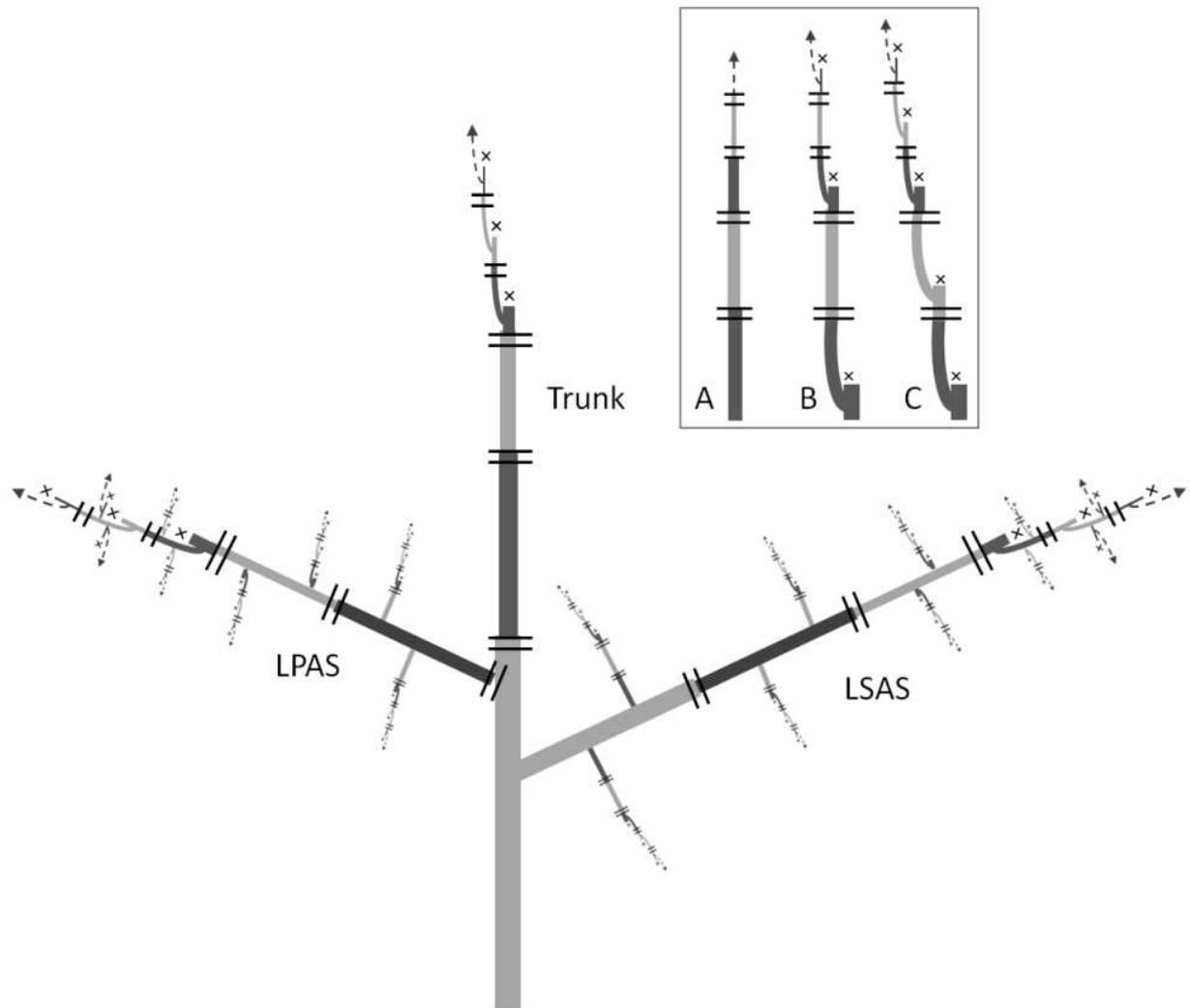
**New insights for estimating the genetic value of segregating apple progenies  
for irregular bearing during first years of tree production**

**Jean-Baptiste Durand, Baptiste Guitton, Jean Peyhardi, Yan Holtz, Yann Guédon,  
Catherine Trottier, Evelyne Costes**

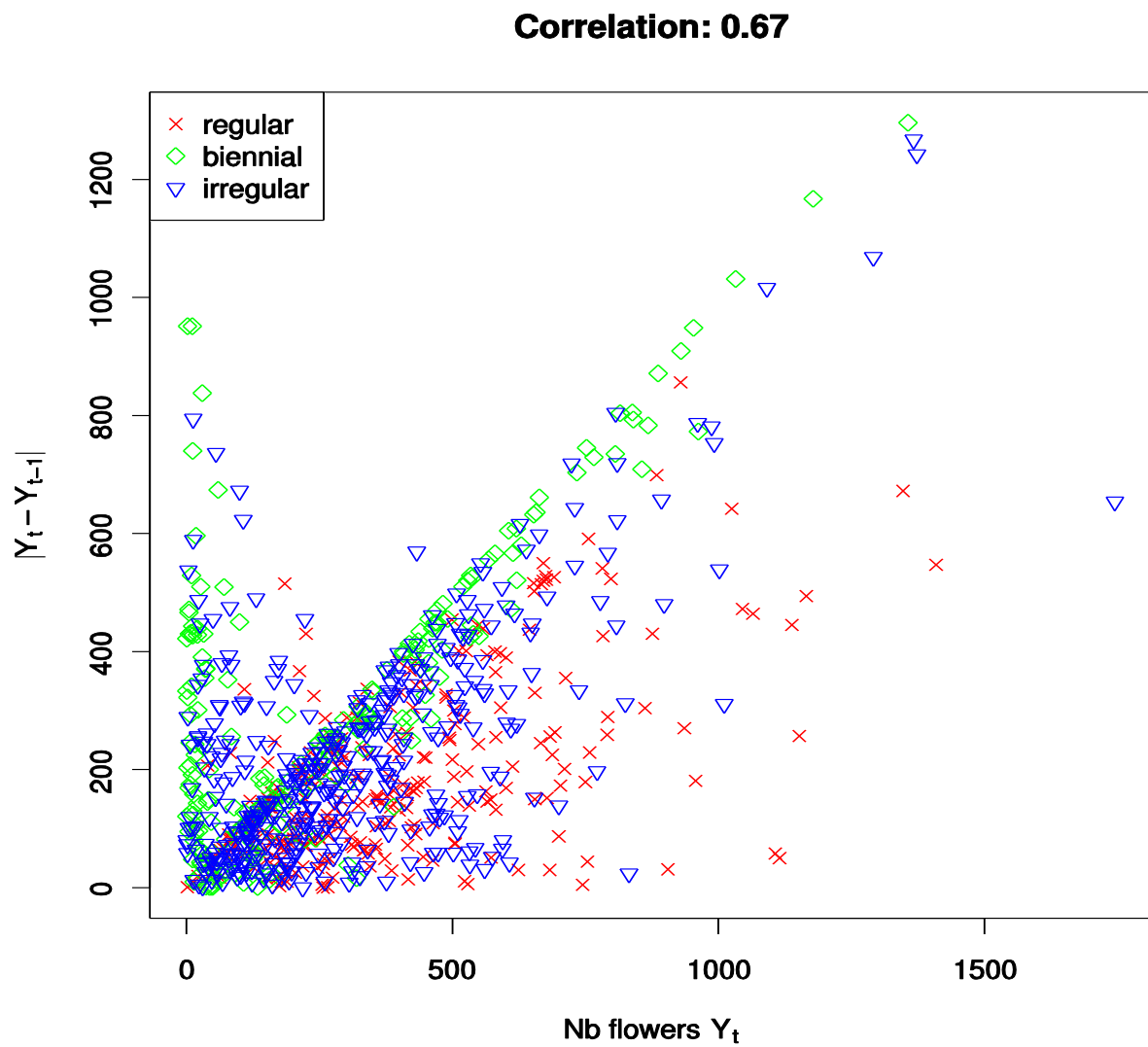
Corresponding author: Jean-Baptiste Durand (+33 4 76 63 57 09), [jean-baptiste.durand@imag.fr](mailto:jean-baptiste.durand@imag.fr)

**Supplementary Information**

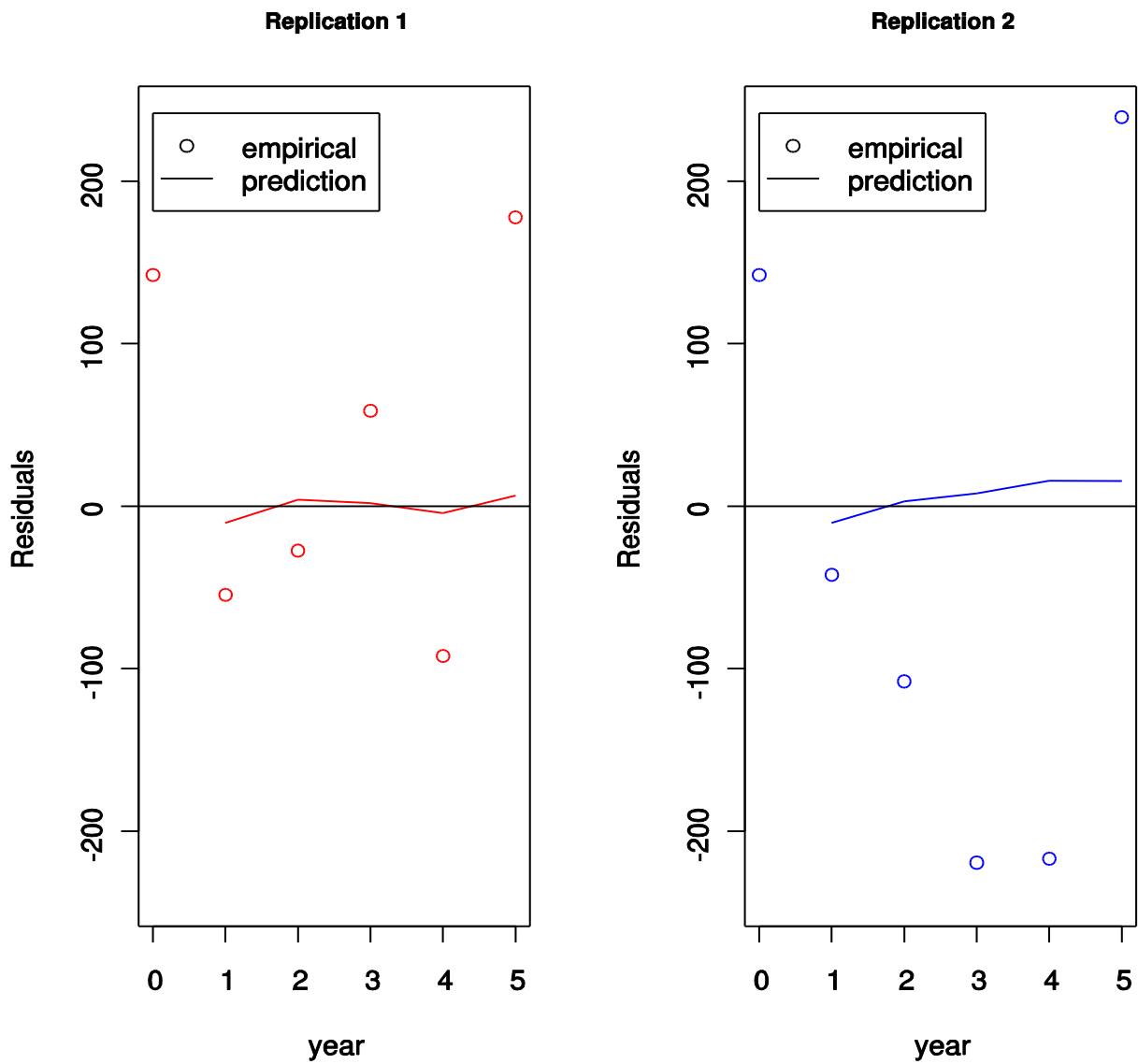
**Fig. S1.** Schematic representation of observations performed on 6 years old trees on the trunk, Long Sylleptic Axillary Shoot (LSAS), Long Proleptic Axillary Shoot (LSAS) and Sort Axillary Shoot (SAS). Annual shoots are delimited by “=” and death of the shoot apical meristem is represented by “x”. (A) Example of a 5 years long sequence bearing no flower, (B) biennial sequence and (C) regular sequence bearing only flowers.



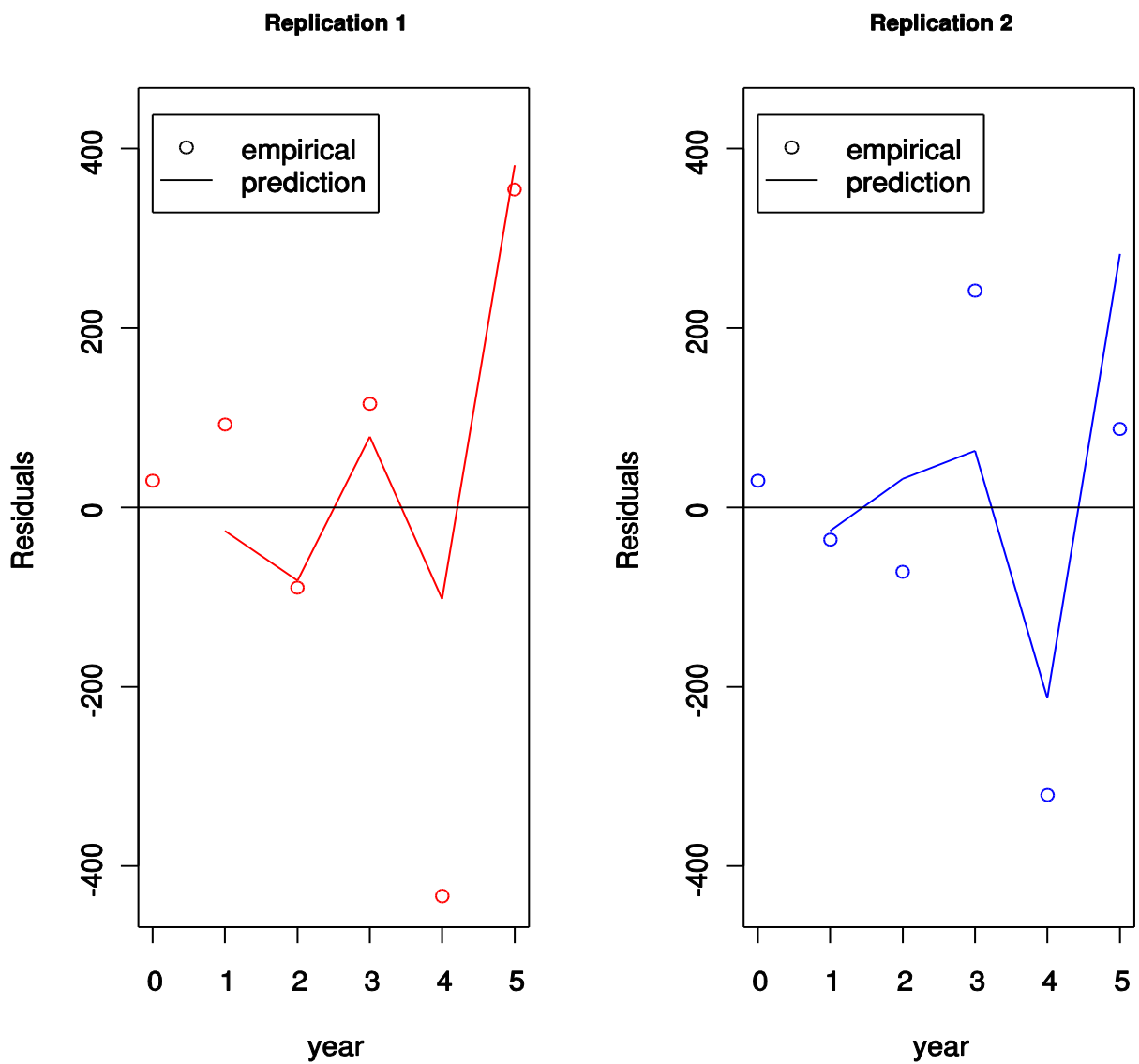
**Fig. S2.** Absolute difference between consecutive yields  $|Y_t - Y_{t-1}|$  as a function of  $Y_t$ . Points with  $Y_t = 0$  or  $Y_{t-1} = 0$  have been removed. The correlation is 0.67, with 95% confidence interval (0.63;0.70), which shows that the implicit hypothesis underlying BBI, i.e. the alternation amplitudes given by the residuals are roughly proportional to the corresponding trend level, is roughly satisfied. Observations associated with regular genotypes are in red, biennial genotypes in green and irregular genotypes in blue. Two main directions are of particular significance:  $Y_t \approx 0$  ( $|Y_t - Y_{t-1}| \approx Y_{t-1}$  in this case) and  $Y_{t-1} \approx 0$  ( $|Y_t - Y_{t-1}| \approx Y_t$  in this case), which both are typical cases of alternation (most points aligned on these directions are from biennial or irregular genotypes).



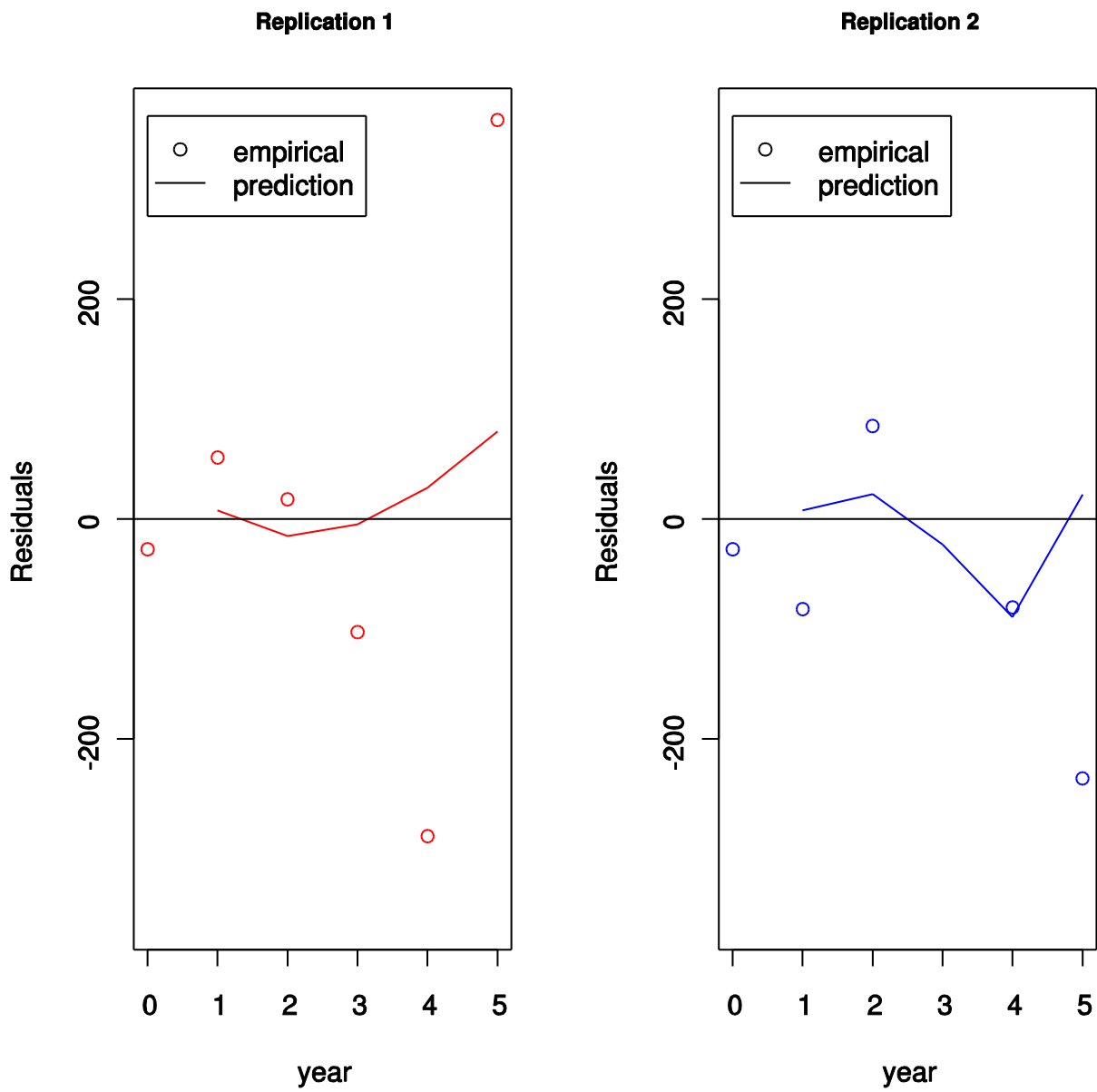
**Fig. S3.** Empirical and predicted residuals of yields as a function of time for regular bearing genotype g=85.



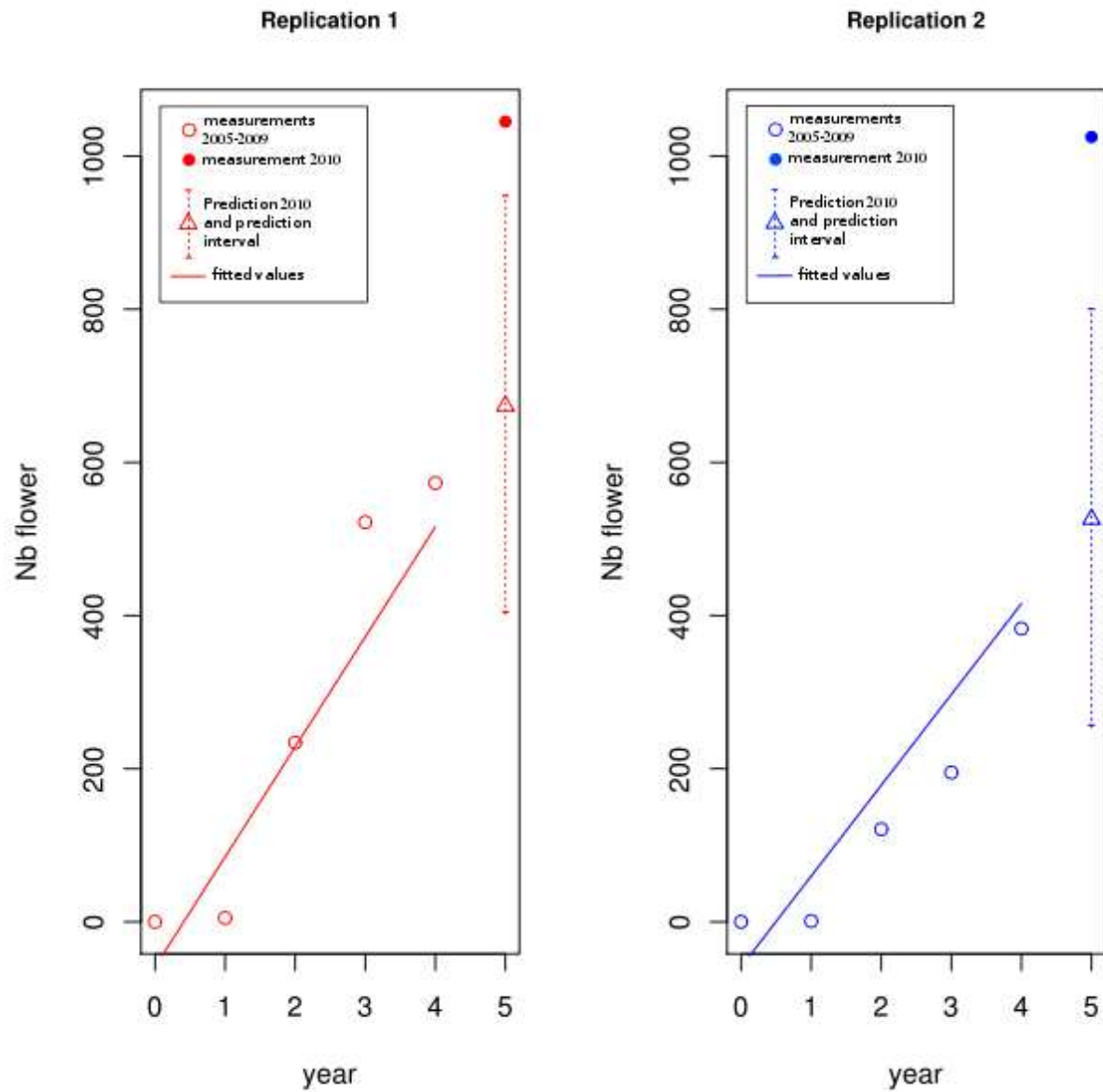
**Fig. S4.** Empirical and predicted residuals of yields as a function of time for biennial bearing genotype g=107.



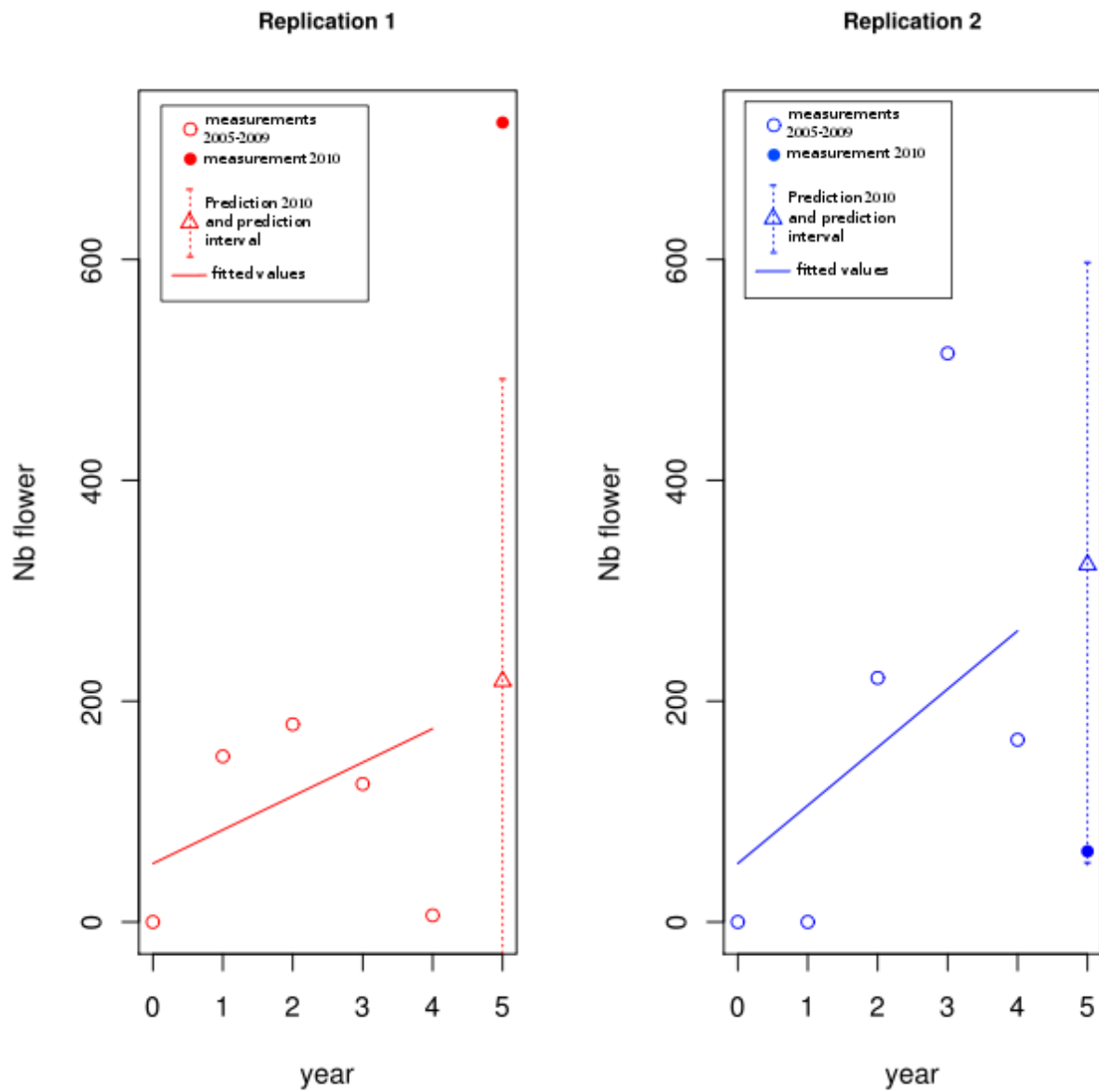
**Fig. S5.** Empirical and predicted residuals of yields as a function of time for irregular bearing genotype g=108.



**Fig. S6a.** Measurements and predicted yield values in 2010 (year number 5) for regular bearing genotype  $g=85$  (a). Circles are the measured values; triangles are the predicted values, located in the middle of prediction intervals (dotted segments).

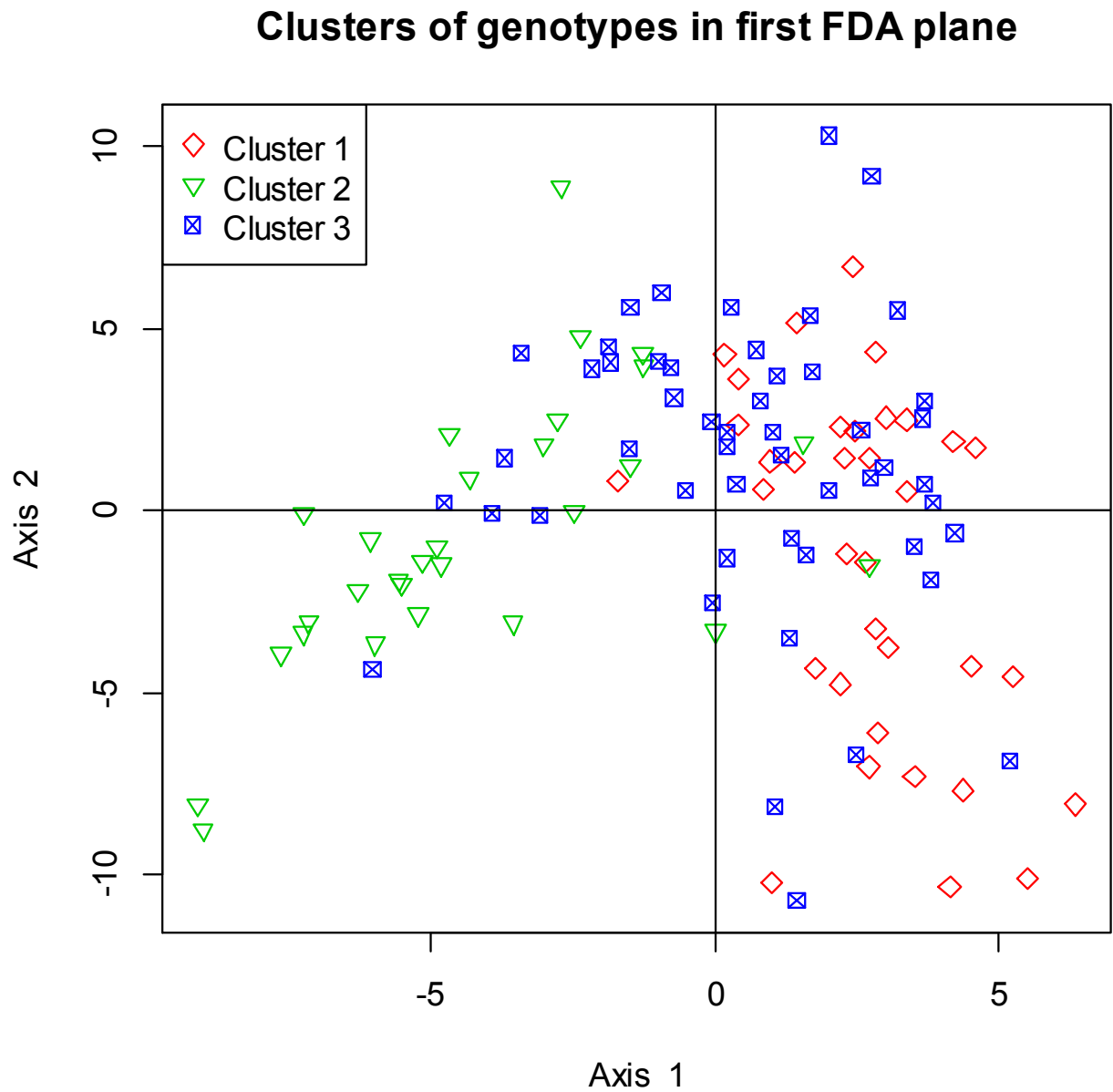


**Fig. S6b.** Measurements and predicted yield values in 2010 (year number 5) for irregular bearing genotype g=108 (b). Circles are the measured values; triangles are the predicted values, located in the middle of prediction intervals (dotted segments).

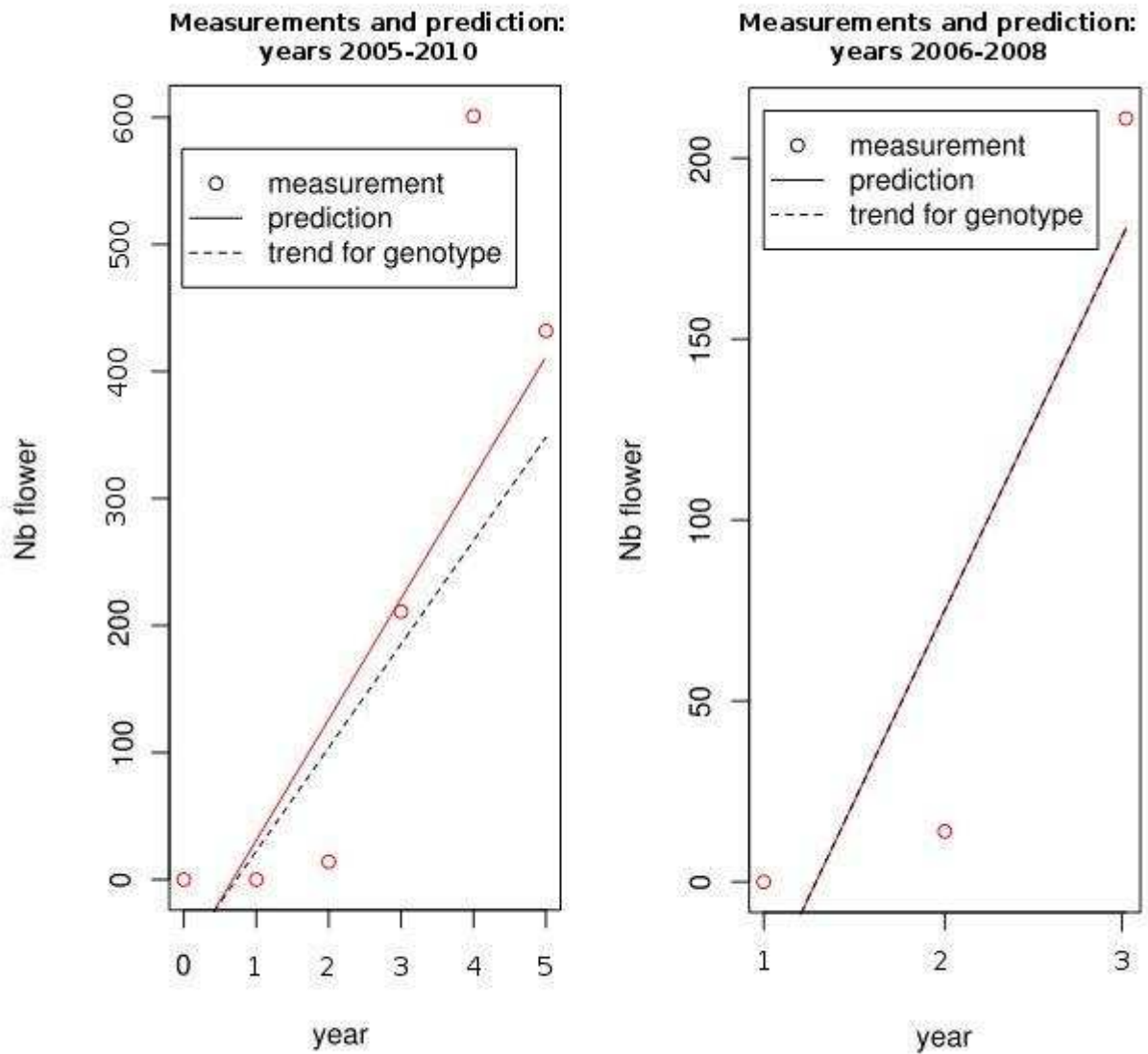




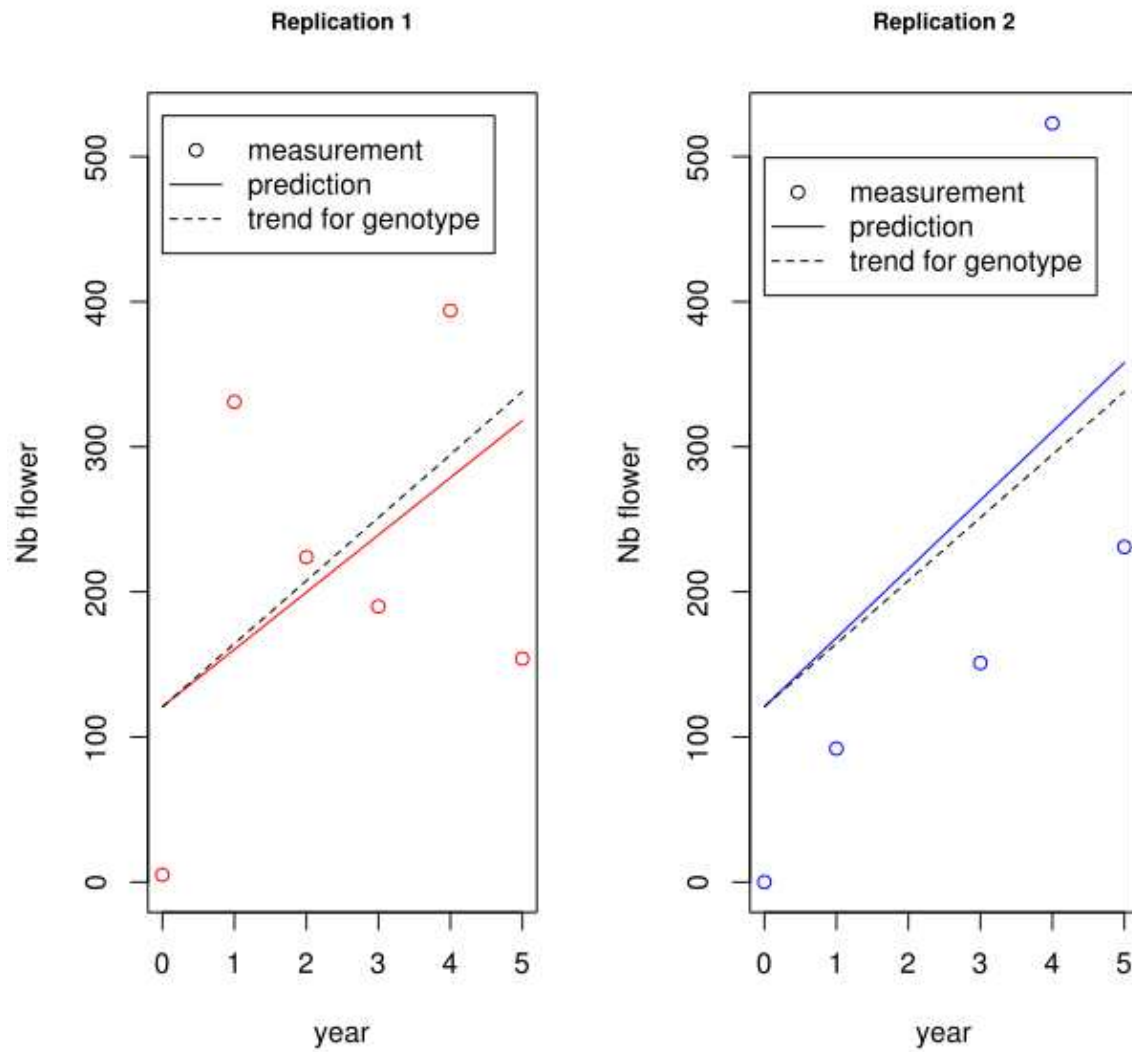
**Fig. S7.** Plot of genotypes in the first FDA plane, based on mean entropy and local indices  $B^{\text{loc}}$  and local genotype AR coefficient  $\gamma^{\text{loc}}$ . The three colours indicate to which cluster each genotype belongs, according to the previous clustering performed with genotype AR coefficient and BBI\_res\_norm. The first FDA axis explains 99% of the variation of intra-cluster inertia, and the second FDA axis 1%.



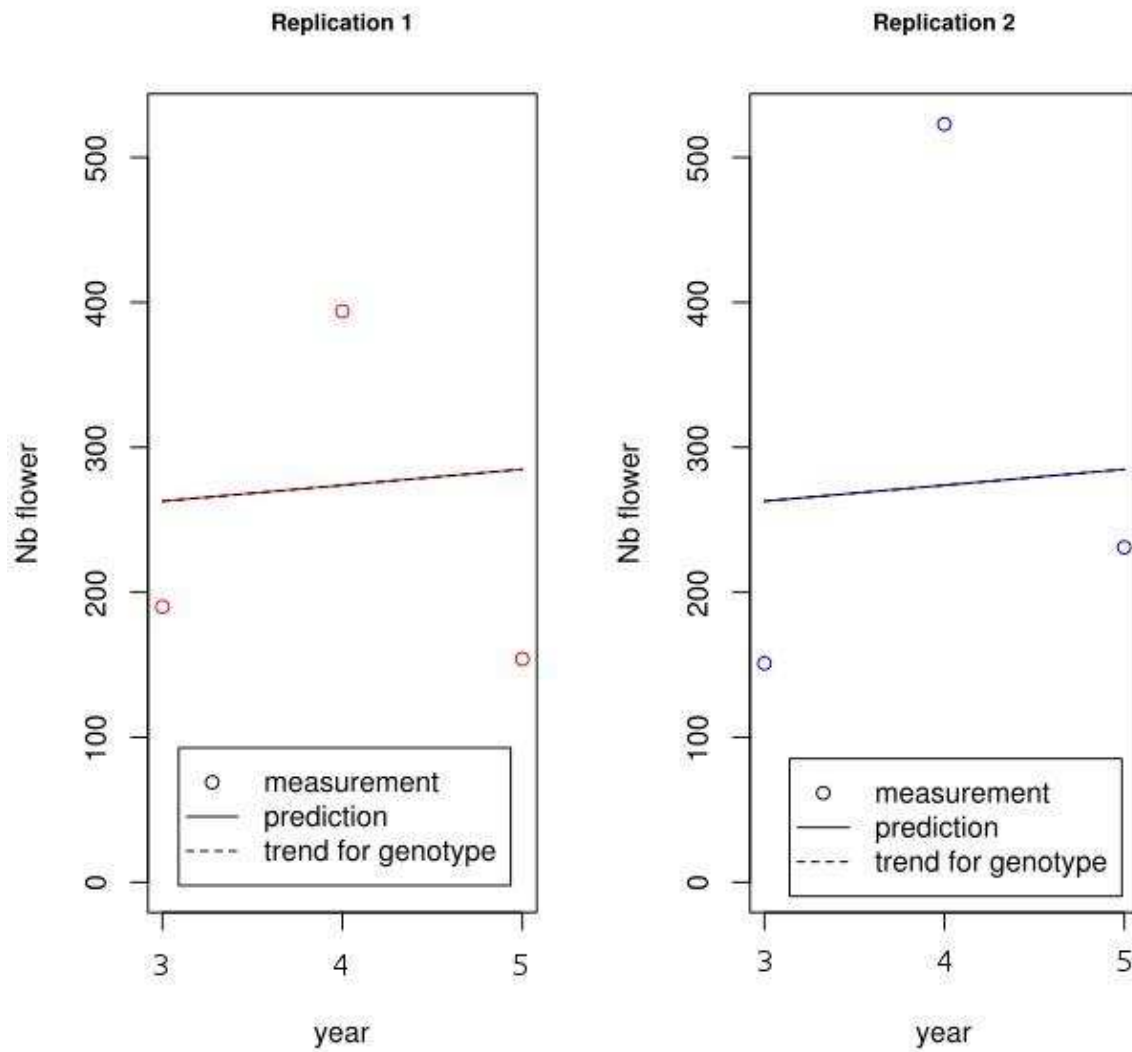
**Fig. S8.** Measurements and fitted values of number of inflorescences for regular bearing genotype  $g=123$ . The left-hand part of the figure corresponds to a model fitted using years 2005 to 2010. The right-hand part of the figure corresponds to a model fitted using years 2006 to 2008.



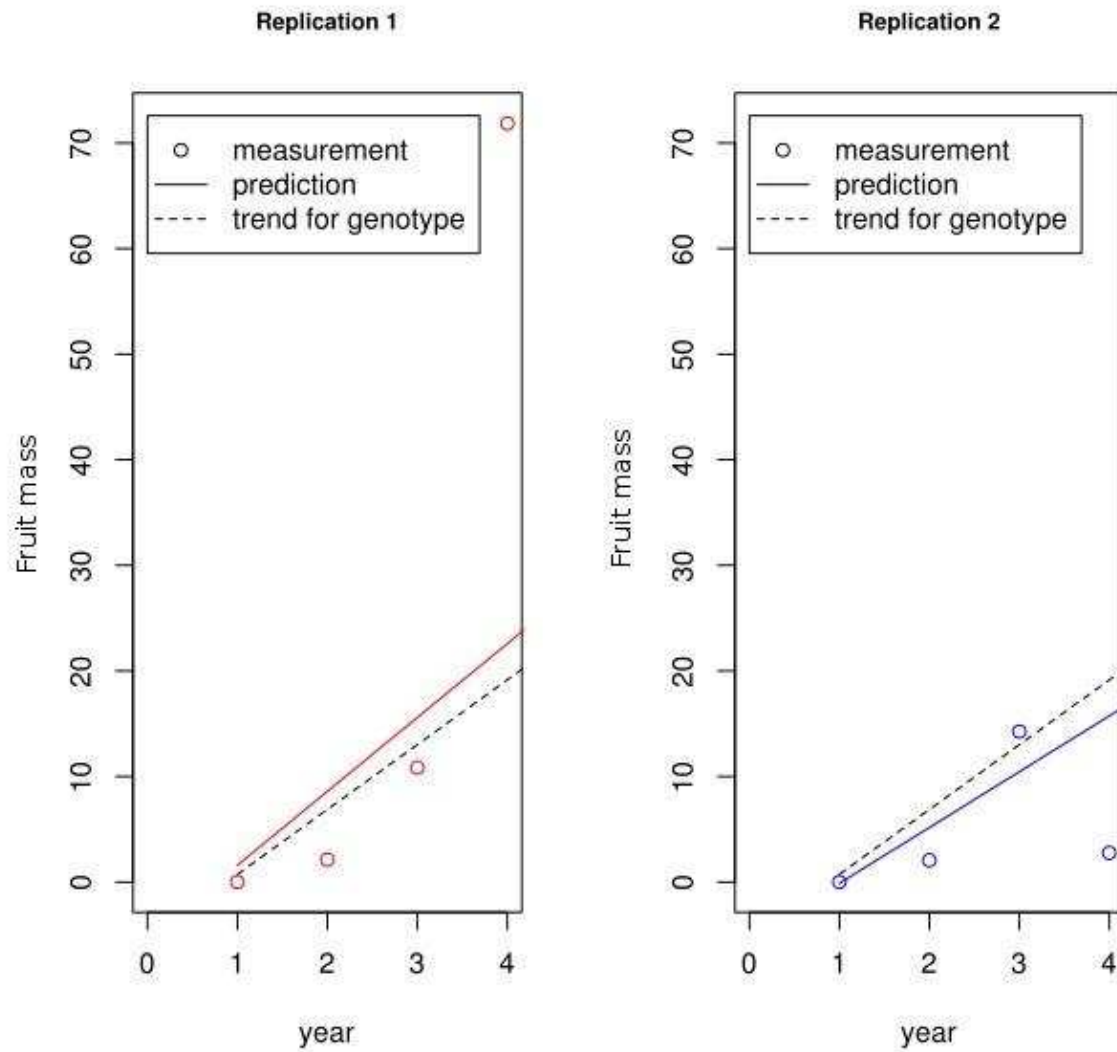
**Fig. S9a.** Measurements and fitted values of number of inflorescences for irregular bearing genotype  $g=5$ . The model was fitted using years 2005 to 2010.



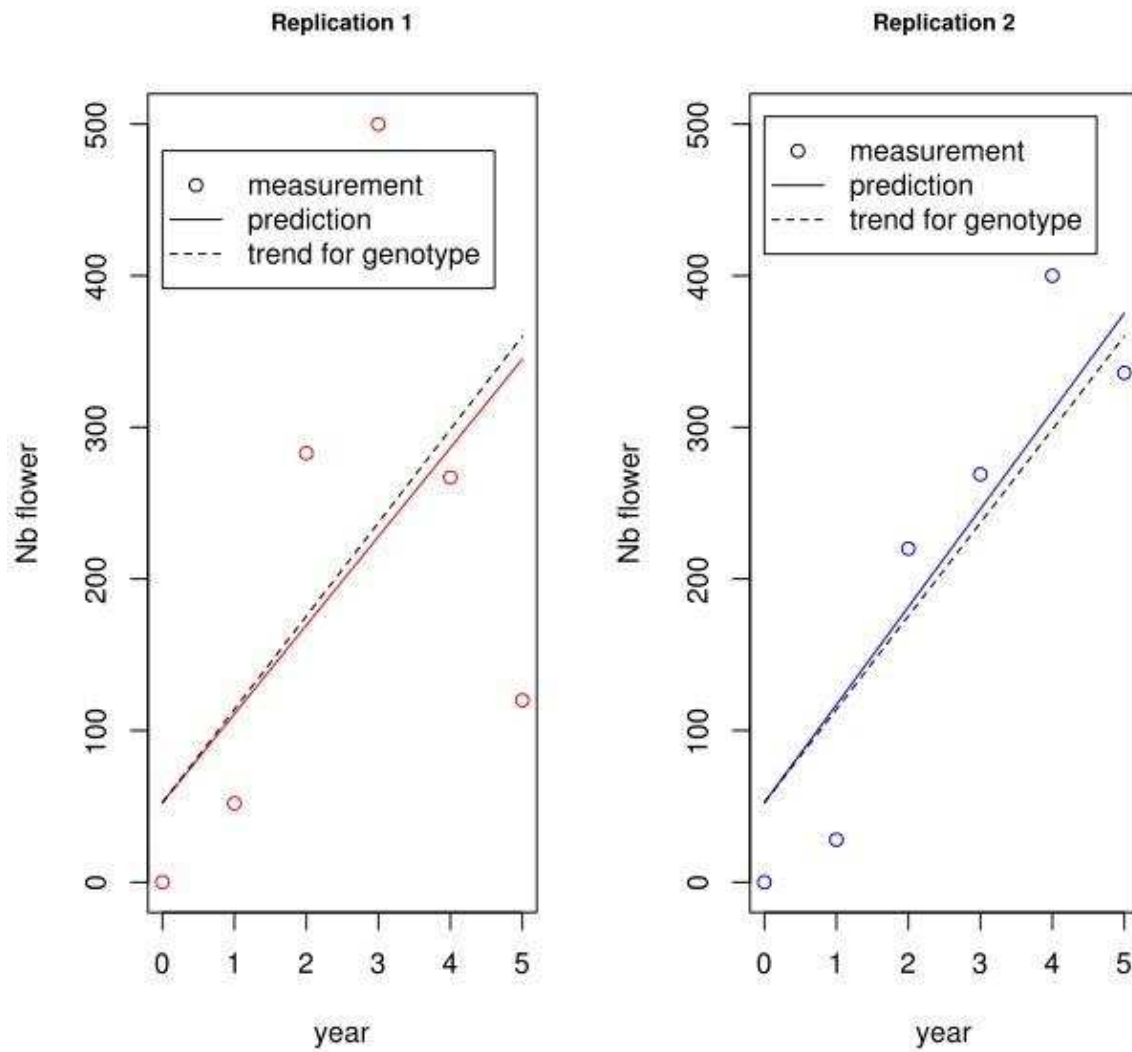
**Fig. S9b.** Measurements and fitted values of number of inflorescences for irregular bearing genotype  $g=5$ . The model was fitted using years 2008 to 2010, which led this genotype to be assessed as alternate bearing.



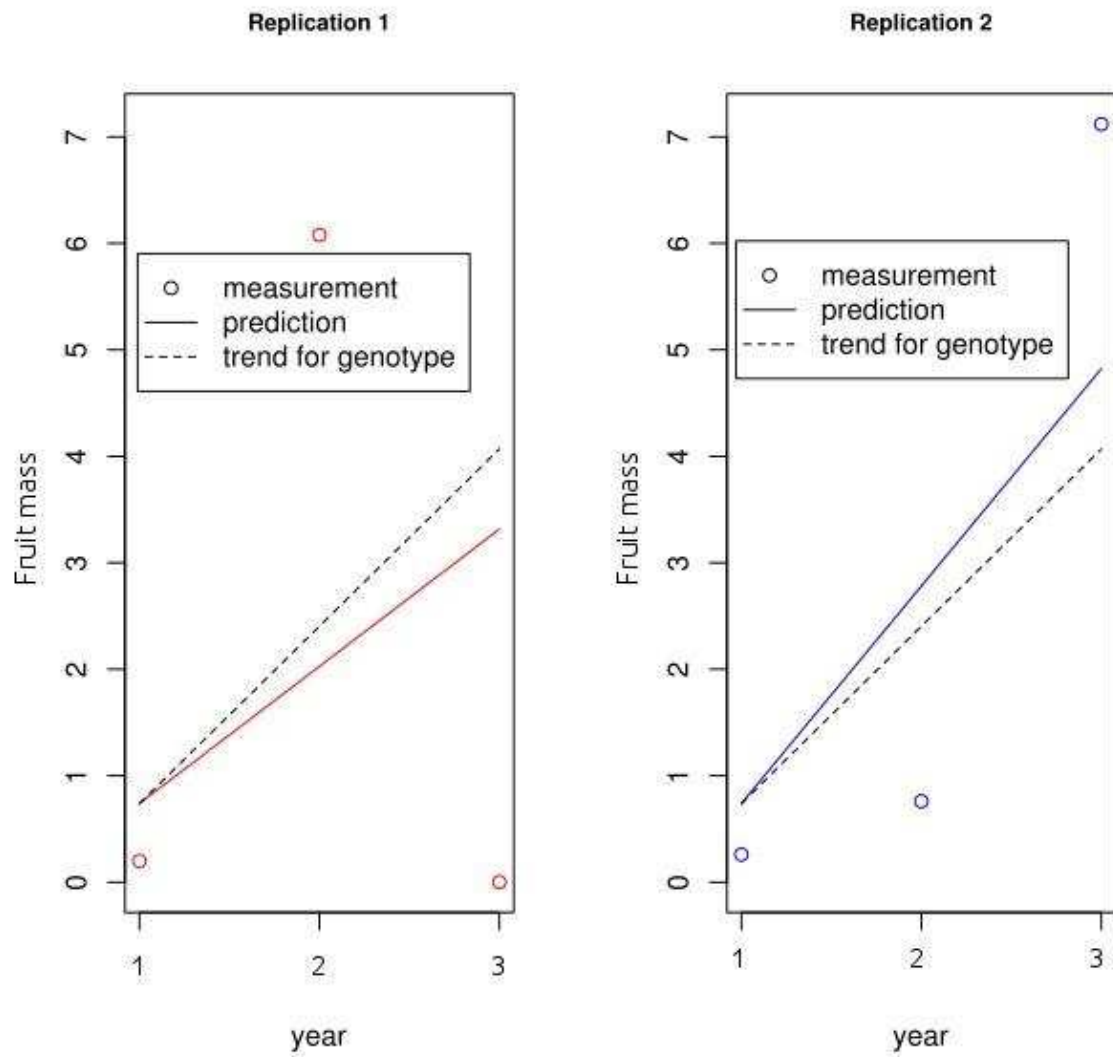
**Fig. S10.** Measurements and fitted values of fruit mass for regular genotype g=85, which led this genotype to be assessed as alternate bearing. The model was fitted using years 2005 to 2010.



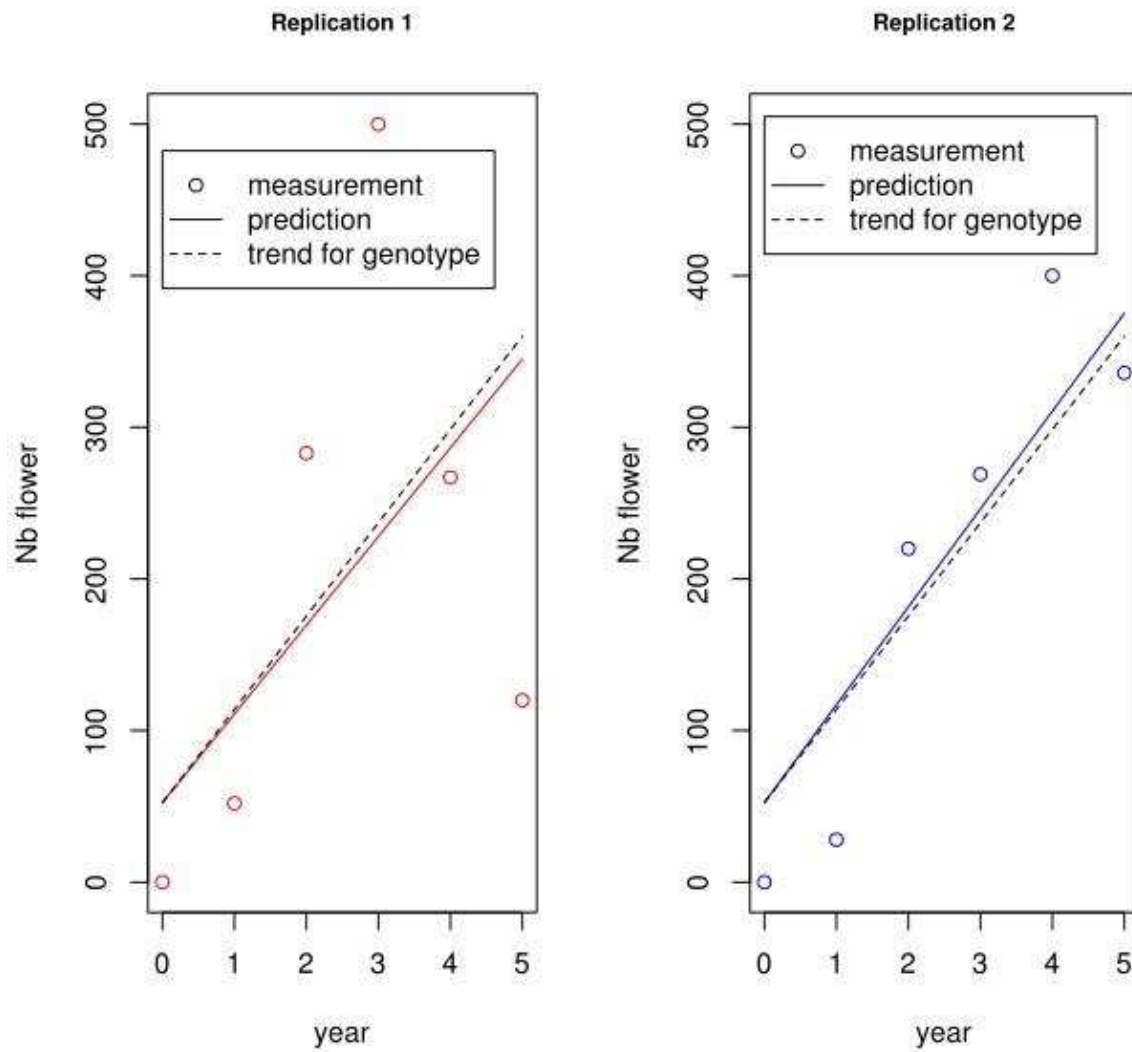
**Fig. S11a.** Measurements and fitted values of number of inflorescences for regular genotype g=62. The model was fitted using years 2005 to 2010.



**Fig. S11b.** Measurements and fitted values of fruit mass for regular genotype g=62, which led this genotype to be assessed as alternate bearing. The model was fitted using years 2006 to 2008.

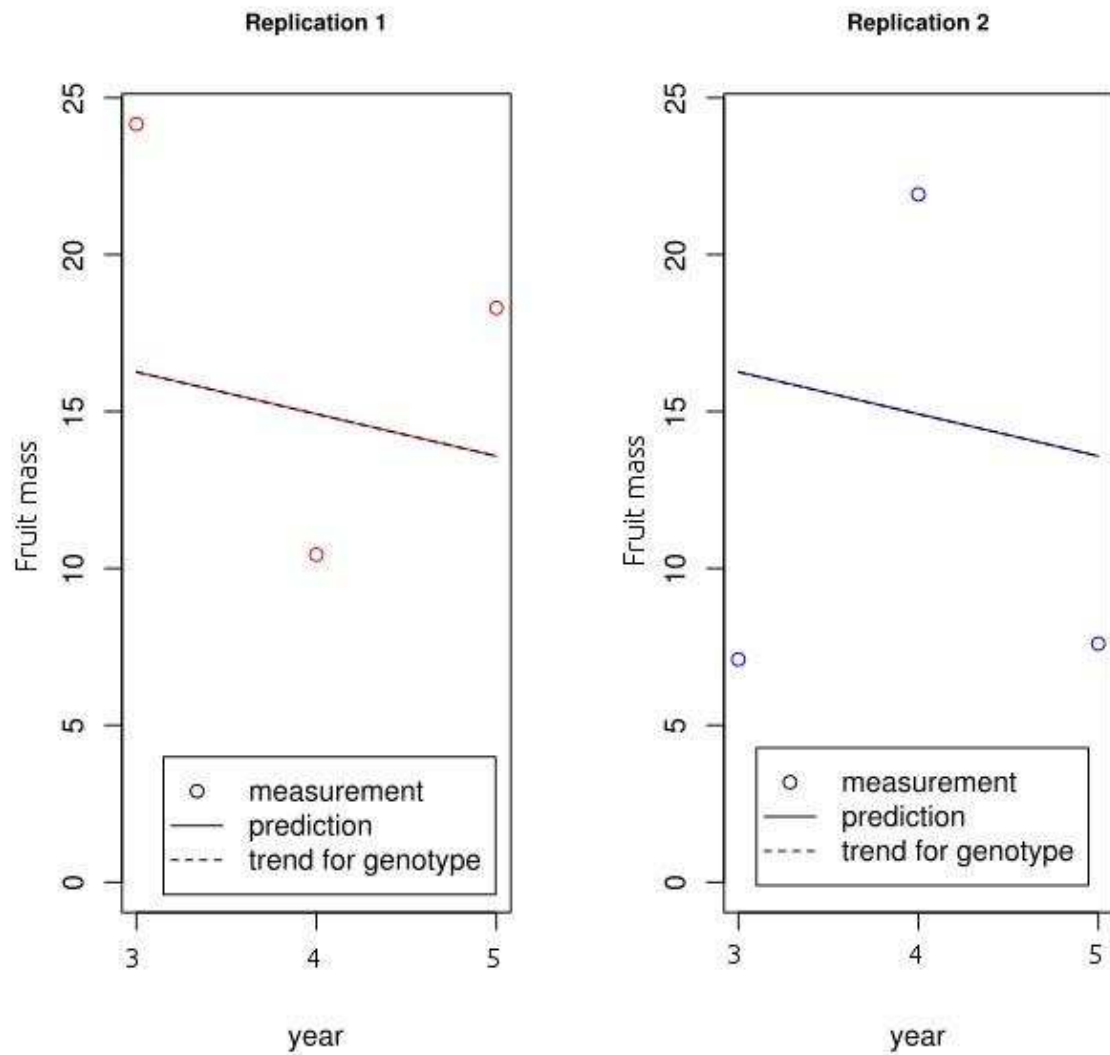


**Fig. S12a.** Measurements and fitted values of number of inflorescences for regular genotype g=64. The model was fitted using years 2005 to 2010.





**Fig. S12b.** Measurements and fitted values of fruit mass for regular genotype g=64, which led this genotype to be assessed as alternate bearing. The model was fitted using years 2008 to 2010.



## Tables

**Table T1.** 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). For each year, the frequency of flowering  $F_{g,r,t}$  and the contribution  $Ent_{g,r}$  to the average entropy are given.

Genotypes		Year					Entropy
		2005	2006	2007	2008	2009	
g=85	Number of GUs	2	6	10	14	18	Total 50
	$F_{g,r,t}$	0.00	0.00	0.20	0.64	0.72	
	$Ent_{g,r}$	0.00	0.00	0.50	0.65	0.59	0.50
g=107	Number of GUs	2	5	9	11	11	Total 38
	$F_{g,r,t}$	0.00	0.60	0.0	1.0	0.0	
	$Ent_{g,r}$	0.00	0.67	0.00	0.00	0.00	0.09
g=108	Number of GUs	5	9	13	15	12	Total 54
	$F_{g,r,t}$	0.00	0.22	0.15	0.53	0.00	
	$Ent_{g,r}$	0.00	0.53	0.43	0.69	0.00	0.38

**Table T2.** Contingency table for the clusters of each genotype. Clusters in lines correspond to those obtained using a Gaussian mixture model, based on years 2005-2010 (reference clusters to compute an error rate). Clusters in columns correspond to those obtained using years 2005-2009. The clusters are R(egular), A(lternate) and I(rregular) bearing. The numbers in parentheses indicate non-significant switches for genotypes at the boundary between clusters.

		Cluster based on years 2005-2009		
		R	A	I
Cluster based on years 2005-2010	R	28	1	5(2)
	A	0	25	6
	I	3 (2)	5	44

**Table T3.** Contingency table for the clusters of each genotype. Clusters in lines correspond to those obtained using the number of inflorescences for years 2005-2010 (reference clusters to compute an error rate). Clusters in columns correspond to those obtained using the number of inflorescences for years 2006-2008. The clusters are R(egular), A(lternate) and I(rregular) bearing.

		Cluster based on years 2006-2008 and the number of inflorescences		
		R	A	I
Cluster based on years 2005-2010 and the number of inflorescences	R	12	5	19
	A	0	21	9
	I	3	13	37

**Table T4.** Contingency table for the clusters of each genotype. Clusters in lines correspond to those obtained using the number of inflorescences for years 2005-2010 (reference clusters to compute an error rate). Clusters in columns correspond to those obtained using the number of inflorescences for years 2008-2010. The clusters are R(egular), A(lternate) and I(rregular) bearing.

		Cluster based on years 2008-2010 and the number of inflorescences		
		R	A	I
Cluster based on years 2005-2010 and the number of inflorescences	R	32	4	0
	A	0	31	0
	I	11	22	22

**Table T5.** Contingency table for the clusters of each genotype. Clusters in lines correspond to those obtained using the number of inflorescences for years 2005-2010 (reference clusters to compute an error rate). Clusters in columns correspond to those obtained using fruit mass for years 2005-2010. The clusters are R(egular), A(lternate) and I(rregular) bearing.

		Cluster based on years 2005-2010 and fruit mass		
		R	A	I
Cluster based on years 2005-2010 and the number of inflorescences	R	21	6	9
	A	3	21	7
	I	25	10	20

**Table T6.** Contingency table for the clusters of each genotype. Clusters in lines correspond to those obtained using the number of inflorescences for years 2005-2010 (reference clusters to compute an error rate). Clusters in columns correspond to those obtained using fruit mass for years 2006-2008. The clusters are R(egular), A(lternate) and I(rregular) bearing.

		Cluster based on years 2006-2008 and fruit mass		
		R	A	I
Cluster based on years 2005-2010 and the number of inflorescences	R	20	5	10
	A	5	19	6
	I	22	12	18

**Table T7.** Contingency table for the clusters of each genotype. Clusters in lines correspond to those obtained using the number of inflorescences for years 2005-2010 (reference clusters to compute an error rate). Clusters in columns correspond to those obtained using fruit mass for years 2008-2010. The clusters are R(egular), A(lternate) and I(rregular) bearing.

		Cluster based on years 2008-2010 and fruit mass		
		R	A	I
Cluster based on years 2005-2010 and the number of inflorescences	R	2	20	14
	A	0	27	4
	I	6	38	11



**Table T8.** Contingency table for the clusters of each genotype. Clusters in lines correspond to those obtained using the number of inflorescences for years 2005-2010 (reference clusters to compute an error rate) and both BBI\_res\_norm and the genotype AR coefficient. Clusters in columns correspond to those obtained using BBI. The clusters are R(egular), A(lternate) and I(rregular) bearing.

		Cluster based on years 2005-2010, the number of inflorescences and BBI		
		R	A	I
Cluster based on years 2005-2010, the number of inflorescences, BBI_res_norm and $\gamma_g$	R	2	0	34
	A	0	27	4
	I	1	9	45

**Table T9.** Correlation coefficient between indices at whole tree and AS scales, with 95% confidence intervals. Indices at whole tree scale are computed on the validation set (first 5 years of yield).

	<b>Genotype AR coefficient <math>g_g</math></b>	<b>Local BBI_res_norm</b>	<b>Local genotype AR coefficient</b>	<b>Mean entropy</b>
<b>BBI_res_norm</b>	-0.62 (-0.72;-0.50)	0.65 (0.54;0.75)	-0.47 (-0.60;-0.31)	-0.52 (-0.65;-0.38)
<b><math>g_g</math></b>	1	-0.52 (-0.65, -0.38)	0.55 (0.41;0.67)	0.34 (0.17;0.49)

**Table T10.** Contingency table for the clusters of each genotype. 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. The clusters are R(egular), A(lternate) and I(rregular) bearing.

		Cluster based on the local indices		
		R	A	I
Cluster based on the global indices	R	17	0	18
	A	1	20	8
	I	9	5	37

**Table T11.** 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.

t	a			
	0.1	1	10	100
5	4.19	0.48	0.07	0.03
25	1.78	0.26	0.04	0.01
400	0.66	0.13	0.02	0.008

**Table T12.** 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.

t	a			
	0.1	1	10	100
5	0.2	0.2	0.2	0.2
25	0.04	0.04	0.04	0.04
400	0.002	0.002	0.002	0.002

## A. Analytic properties of BBI

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

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

where  $T_{g,r}$  denotes the number of measurements for replication  $r$  of genotype  $g$ , and with the convention that

$$\frac{|Y_{g,r,t} - Y_{g,r,t-1}|}{Y_{g,r,t-1} + Y_{g,r,t}} = 0$$

if  $Y_{g,r,t-1} = Y_{g,r,t} = 0$ . Compared to the usual presentation of BBI, a multiplying factor of value 2 is introduced to make BBI comparable in scale to the indices introduced below. The justification is that, in this way, the elementary terms which are averaged take the form of a ratio between an absolute difference  $|Y_{g,r,t} - Y_{g,r,t-1}|$  and a mean  $(Y_{g,r,t-1} + Y_{g,r,t})/2$ .

Using BBI in trended series generates confusion between alternation and trend, as developed in proposition P1. This proposition shows that the BBI of a series on length  $T$  with affine growth has order of magnitude  $\log T / (T - 1)$ . Moreover, interpreting BBI in the framework of linear filtering (Diggle, 1990; Chatfield, 2003) highlights a restrictive assumption underlying this index, related to the interpretation of BBI as the sum of the absolute values of the residuals obtained by first-order differencing normalized by the sum of the two successive values involved in the differencing (that can be interpreted as a very local trend). The underlying implicit hypothesis is that the alternation amplitudes given by the residuals are roughly proportional to the corresponding trend level (the relevance of this hypothesis for our dataset is highlighted in Fig. S2). In the case where the residuals are independent from the trend level, BBI scales the residuals as a function of the trend level. Each absolute difference will be weighted differently, and BBI will be irrelevant. This should be considered its main shortcoming, as index for alternation.

## P. Proofs of propositions

### P1/ BBI of a time series with affine growth

If  $Y_t = at + b$  then the BBI is asymptotically equivalent to  $\log t / (t-1)$ . We assume that  $a > 0$  and  $b > 0$  to ensure that  $Y_t > 0$ , but the proof can be easily adapted to the other cases, replacing  $Y_t + Y_{t-1}$  by  $|Y_t| + |Y_{t-1}|$  in the definition of BBI.

Proof

For any  $t$ ,  $|Y_t| + |Y_{t-1}| \frac{|Y_t - Y_{t-1}|}{Y_t + Y_{t-1}} = \frac{|a|}{(at+b) + (at-a+b)}$ . We have

$$\begin{aligned} \sum_{n=1}^t \frac{|Y_n - Y_{n-1}|}{Y_n + Y_{n-1}} &= \sum_{n=1}^t \frac{a}{2an - a + 2b} \\ &= \sum_{n=1}^t \frac{1}{2n(1 - \frac{1}{2n} + \frac{b}{an})} = \frac{1}{2} \sum_{n=1}^t \frac{1}{n} (1 + \frac{1}{2n} - \frac{b}{an} + o(\frac{1}{n})), \end{aligned}$$

where  $\sum_{n=1}^t \frac{1}{n}$  is asymptotically equivalent to  $\log t$  and where  $\sum_{n=1}^t \frac{1}{n^2} = \frac{\pi^2}{6}$ . Hence, the BBI is asymptotically equivalent to  $\alpha(t) = \log t / (t-1)$ . The difference between the BBI and this equivalent increases with ratio  $a/b$ . The precision  $|BBI(t) - \alpha(t)| / BBI(t)$  of approximation  $\alpha(t)$  is given in Table T11 for different values of  $a$  and  $t$ , in the case where  $b=1$ . This table shows that for large values of  $a$ , the precision of approximation  $\alpha(t)$  is quite good even for small values of  $t$ .

## P2/ BBI and BBI\_norm for stationary time series with constant amplitude alternation

Recall that BBI and BBI\_norm are defined as

$$\text{BBI} = \frac{2}{T-1} \sum_{t=2}^T \frac{|Y_t - Y_{t-1}|}{Y_{t-1} + Y_t},$$
$$\text{BBI\_norm} = \frac{\sum_{t=2}^T |Y_t - Y_{t-1}| / (T-1)}{\sum_{t=1}^T Y_t / T},$$

where it is assumed that these indices apply to series of non-negative values.

For a linear trend with residuals corresponding to an alternation with amplitudes proportional to the trend level ( $0 < c \leq a$ )

$$Y_{2t} = (a + c)2t,$$
$$Y_{2t+1} = (a - c)(2t + 1),$$

we have

$$\text{BBI} = \frac{2}{T-1} \sum_{t=2}^T \frac{a + c(2t-1)}{a(2t-1) \pm c},$$
$$\text{BBI\_norm} = \frac{\sum_{t=2}^T \{a + c(2t-1)\} / (T-1)}{a(T+1)/2}$$
$$= \frac{a + c(T+1)}{a(T+1)/2}.$$

$\text{BBI} \rightarrow 2c/a$  and  $\text{BBI\_norm} \rightarrow 2c/a$  when  $t \rightarrow +\infty$ . While the ranges of possible values of the two indices are similar, we expected BBI\_norm to be more robust to outliers. This illustrates the fact that these indices are only relevant when the alternation amplitudes are roughly proportional to the corresponding trend level, a particular case being a stationary series with constant alternation amplitudes.

For a linear trend with residuals corresponding to an alternation with constant amplitude  $0 < c \leq a$

$$Y_{2t} = a2t + c,$$
$$Y_{2t+1} = a(2t+1) - c,$$



we have

$$\begin{aligned} \text{BBI} &= \frac{2}{T-1} \sum_{t=2}^T \frac{a+2c}{a(2t-1)} \\ &\approx \frac{(a+2c)\log 2T}{a(T-1)}, \\ \text{BBI}_{\text{norm}} &= \frac{2(a+2c)}{a(T+1)}, \\ \text{BBI}_{\text{res\_norm}} &= \frac{4c}{a(T+1)}. \end{aligned}$$

BBI takes the form of a sub-series of the harmonic series.  $\text{BBI} \rightarrow 0$ ,  $\text{BBI}_{\text{norm}} \rightarrow 0$  and  $\text{BBI}_{\text{res\_norm}} \rightarrow 0$  when  $t \rightarrow +\infty$ .

### P3/ Indices for stationary time series with constant amplitude alternation

For a stationary series with average  $a$  and residuals corresponding to an alternation with constant amplitude  $0 < c \leq a$

$$Y_{2t} = a + c,$$

$$Y_{2t+1} = a - c,$$

we have

$$\begin{aligned} \text{BBI} &= \frac{2}{T-1} \sum_{t=2}^T \frac{2c}{2a} = \frac{2c}{a}, \\ \text{BBI}_{\text{norm}} &= \frac{2c}{a}, \\ \text{BBI}_{\text{res\_norm}} &= \frac{2c}{a}. \end{aligned}$$

P4/ BBI\_res\_norm of an alternate time series with linear growth

If  $Y_{2t+1} = a(2t+1)$  and  $Y_{2t} = 0$  (with  $a > 0$ ) then the  $BBI\_res\_norm$  is asymptotically independent from  $a$  and tends towards 2.

Proof

From classical results in linear regression, the least square line has slope

$$\begin{aligned} \frac{\sum_{n=0, n \text{ odd}}^{2t+1} n Y_n}{\sum_{n=0}^{2t+1} n^2} &= \frac{\sum_{n'=0}^t a(2n'+1)^2}{\sum_{n=0}^{2t+1} n^2} = a \frac{4 \sum_{n'=0}^t n'^2 + 4 \sum_{n'=0}^t n' + t + 1}{\sum_{n=0}^{2t+1} n^2} \\ &= a \frac{\frac{2}{3} t(t+1)(2t+1) + 2t(t+1) + t + 1}{\frac{(2t+1)(2t+3)(4t+3)}{6}} = \frac{a}{2} + o(1) \end{aligned}$$

Thus, the predictor is  $\hat{Y}_t = \frac{a}{2}t + o(t)$  and the empirical residual is  $\hat{\varepsilon}_t = \frac{a}{2}t + o(t)$  if  $t$  is odd

and  $\hat{\varepsilon}_t = -\frac{a}{2}t + o(t)$  if  $t$  is even. Consequently,  $|\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}| = |a|t + o(t)$  and

$$\sum_{n=1}^t |\hat{\varepsilon}_n - \hat{\varepsilon}_{n-1}| = |a| \frac{t(t+1)}{2} + o(t^2) = a \frac{t^2}{2} + o(t^2).$$

Since (for example if  $t$  is odd and  $t = 2k+1$ )

$$\sum_{n=1}^t Y_n = a \sum_{n'=0}^k (2n'+1) = 2a \frac{k(k+1)}{2} + o(t^2) = a \frac{t^2}{4} + o(t^2),$$

we have

$$BBI\_res\_norm(t) = \frac{a \frac{t^2}{2} + o(t^2)}{a \frac{t^2}{4} + o(t^2)} \times \frac{t}{t-1} = 2 + o(1).$$

The precision  $|BBI\_res\_norm(t) - 2| / BBI\_res\_norm(t)$  of the approximation of  $BBI\_res\_norm(t)$  by its limit 2 is given in Table T12 for different values of  $a$  and  $t$ . This table shows that for any value  $a$ , the precision of the approximation is good for  $t > 25$ .

## M. Supplementary description of statistical models and methods

### M1/ Clustering using Gaussian independent mixture models

If  $z = z_g$  refers to the two-dimensional vector of indices, the Gaussian mixture model is defined by its probability density function (pdf)

$$f(z) = \sum_{k=1}^K \pi_k f_k(z; \mu_k, \Sigma_k),$$

where  $f_k(z; \mu_k, \Sigma_k)$  denotes the pdf of the bivariate Gaussian distribution with mean  $\mu_k$  and (diagonal) covariance matrix  $\Sigma_k$ . This pdf corresponds to the assumption that genotypes within cluster  $k$  follow the distribution  $f_k(z; \mu_k, \Sigma_k)$ . The clustering is obtained by estimating the model parameters  $\pi_k$ ,  $\mu_k$  and  $\Sigma_k$ , and by associating each genotype  $z_g$  with the most likely cluster. The number of clusters  $K$  was selected using BIC. We used our own implementation of mixture models, developed with the R software.

### M2/ Model estimation, selection and validation in neural networks and SVMs

Neural networks and SVMs depend on two kinds of parameters:

- Parameters that can be estimated automatically from the dataset by optimizing a criterion (the likelihood function in the case of NNs, or a geometric criterion in the case of SVMs). Estimation relies on a set of genotypes, referred to as “learning sample”, which classes are considered as known. In practice, the classes yielded by Gaussian mixture clustering were considered.
- A so-called regularisation parameter, denoted by  $\nu$ , which controls the ability of the model to predict correctly either the classes of the learning set, or those of future genotypes not in the learning set (and even if possible, classes of both types of genotypes). The regularisation parameter  $\nu$  has to be specified by the modeller.

In the case of classes comprising reasonably comparable numbers of genotypes, the performance of supervised classification methods can be assessed with the classification error

rate (or error rate, in short), which is the frequency of genotypes which class is not correctly predicted. The set of genotypes used to compute the error rate is referred to as “test sample”. Usually, a low classification error rate (perfect classification) can be achieved through a particular choice of  $\nu$ , if the whole dataset is used simultaneously as learning and test sample. However, this is an optimistic prediction of the actual error rate on future genotypes, since the same dataset is used both to estimate the parameters and to compute the error rate. A more reliable way to assess the possibility of classifying future genotypes accurately is the cross-validated error rate (Bishop, 2006, Chapter 1). One half of genotypes, chosen randomly, are used as learning sample and the other half as test sample. Then the roles of both sets are permuted, and this procedure is repeated several times (5 times in our case) to reduce the variability in estimating the average error rate. This variability is related to the random choice of both sets. This algorithm is applied to several values of the regularisation parameter  $\nu$ , so as to minimise the predicted error rate with respect to  $\nu$ .

### M3/ Factorial Discriminant Analysis (FDA)

FDA is a variant of principal component analysis that aims at providing linear subspaces (e.g., a plane) in which the classes are optimally separated (Tabachnick & Fidell, 2007). This subspace is obtained by maximising the separation between the centres of the classes, in regard to the dispersion of the data of each class around their means. The plane obtained by applying the FDA to each genotype characterised by the three local indices is depicted in Fig. 6, which shows a correct discrimination between the regular genotypes (represented with red diamonds) and the biennial alternating genotypes (green triangles). Note that axis 1 seems sufficient, essentially, to separate both classes. This justifies its use as a scoring and ranking method. The most regular genotypes have maximal coordinates along x-axis, and most alternate genotypes have minimal coordinates along this axis. The irregular genotypes (blue

squares) seem to be uniformly distributed on the plane. Since this is the plane where the classes are optimally separated, genotypes with irregular yields at tree scale cannot be discriminated using the local indices.

### **E. Effect of using 3 years of yields at tree scale (number of inflorescence or fruit mass) on the indices, clusters and ranking**

The clusters and ranks obtained using the annual numbers of inflorescences at tree scale were used as a reference. Thus, the clusters and rankings using fewer years of measurement or fruit mass were compared with this reference, using contingency tables and error rates for clusters, and Kendall's  $\tau$  for rankings. Kendall's  $\tau$  coefficient is in  $[-1,1]$ ; this is a measure of similarity between rankings, so that the maximal value represents perfect agreement between rankings. We give the rankings for the 3 reference genotypes: 85 (rank: 20/122), 107 (rank: 101/122), 108 (rank: 64/122).

#### **Effect of using the number of inflorescences for years 1-3**

The confusion matrix is given in Table T3. The associated error rate was 0.41. Kendall's  $\tau$  coefficient was 0.39. Genotype 85 was assessed regular bearing with rank 1, genotype 107 irregular (instead of alternate) bearing with rank 65 and genotype 108 irregular bearing with rank 24. Among the regular bearing genotypes assessed as alternate bearing was genotype 123, which ranking increased from 19 to 101. The yields of regular genotype 123 are given in Figure S8 (model fitted using every year and model fitted using the first three years of yields). Replicate 2 of this genotype was not used to compute the indices based on three years on yields because this replicate did not bear flowers for the first three years. The cluster

switch for this genotype can be attributed to a high variability of yields over years and also between replicates.

#### **Effect of using the number of inflorescences for years 3-5**

The confusion matrix is given in Table T4. The associated error rate was 0.30. Kendall's  $\tau$  coefficient was 0.71. Genotype 85 was assessed regular with rank 21, genotype 107 alternate bearing with rank 95 and genotype 108 alternate (instead of irregular) bearing with rank 83. Among the irregular bearing genotypes assessed as alternate bearing was also genotype 5, which ranking increased from 70 to 80. The yields of regular genotype 5 are given in Figures S9 a) (model fitted using every year) and b) (model fitted using the last three years of yields). The cluster switch for this genotype can be attributed to a high variability of yields over years and also between replicates.

#### **Effect of using fruit mass for years 0-5**

The confusion matrix is given in Table T5. The associated error rate was 0.49. Kendall's  $\tau$  coefficient was 0.42. Genotype 85 was assessed alternate bearing with rank 56, genotype 107 alternate bearing with rank 89 and genotype 108 alternate bearing with rank 38. The yields of regular genotype 85 are given in Figure S10 (to be compared with Figure 1a). When using fruit mass to compute the indices, this genotype is placed at the boundary between alternate bearing and irregular genotypes. Its BBI\_res\_norm value is close to 1.5, which led the clustering to place it with alternate bearing genotypes.

#### **Effect of using fruit mass for years 1-3**

The confusion matrix is given in Table T6. The associated error rate was 0.54. Kendall's  $\tau$  coefficient was 0.26. Genotype 85 was assessed irregular (instead of regular) bearing with

rank 82, genotype 107 alternate bearing with rank 81 and genotype 108 regular (instead of irregular) bearing with rank 39. Among the regular bearing genotypes assessed as alternate bearing was genotype 62, which ranking increased from 10 to 117. The yields of regular genotype 62 are given in Figures S11 a) (model fitted using the number of inflorescences) and b) (model fitted using fruit mass). The cluster switch for this genotype can be attributed to the fact that for replicate 1, the number of inflorescences was particularly high in year 3 while fruit mass was particularly low.

### **Effect of using fruit mass for years 3-5**

The confusion matrix is given in Table T7. The associated error rate was 0.67. Kendall's  $\tau$  coefficient was 0.26. Genotype 85 was assessed irregular (instead of regular) bearing with rank 66, genotype 107 alternate bearing with rank 68 and genotype 108 irregular bearing with rank 40. Among the regular bearing genotypes assessed as alternate bearing was genotype 64, which ranking increased from 24 to 47. The yields of regular genotype 62 are given in Figures S12 a) (model fitted using the number of inflorescences) and b) (model fitted using fruit mass). The cluster switch for this genotype can be attributed to the fact that for replicate 1, the number of inflorescences is rather irregular (while it is regular for replicate 2). Fruit mass becomes alternate for both replicates in years 3-5, which can be attributed to environmental variations (not characterised by measurements in this study).

### **Further remarks of the effect of using 3 years of yields at tree scale or using or fruit mass**

The conclusions of using fruit mass instead of the number of inflorescences, or of using three years of production instead of 6 years, are qualitatively the same as using the number of inflorescences for first five years only (see Section 2.3.5). The confusion between clusters of

genotypes as well as the number of rank inversions tend to increase when using the last three years of yields, to increase further when using the first three years of yields, and to increase still further when using fruit mass. The confusion involves in 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. In particular, this is the case for switches from alternate bearing to regular genotypes, which is the error with most prejudicial consequences in our context.

### **C. Correlations between fruit mass and number of inflorescences at successive years**

The correlation between the number of inflorescences  $Y_{g,r,t}$  at replicate scale at year  $t$  and fruit mass  $M_{g,r,t+1}$  the year after  $t+1$  is 0.03. This is not significant at level 0.05, since the confidence interval is  $[-0.03, 0.09]$ . This is due to  $Y_{g,r,t}$  and  $Y_{g,r,t+1}$  being positively correlated for regular genotypes (due to the trend) and negatively correlated for alternate bearing genotypes. This is highlighted by the within-cluster correlations between  $Y_{g,r,t}$  and  $M_{g,r,t+1}$ , which is 0.33 for regular genotypes (95% confidence interval  $[0.23, 0.42]$ ), -0.26 for alternate bearing genotypes (95% confidence interval  $[-0.36, -0.15]$ ) and 0.10 for irregular bearing genotypes (95% confidence interval  $[0.02, 0.18]$ ).

Similarly, the correlation between  $M_{g,r,t}$  and  $Y_{g,r,t+1}$  is -0.05 (95% confidence interval  $[-0.12, 0.01]$ ). This correlation is 0.26 for regular genotypes (95% confidence interval  $[0.15, 0.37]$ ), -0.28 for alternate bearing genotypes (95% confidence interval  $[-0.39, -0.16]$ ) and -0.02 for irregular bearing genotypes (95% confidence interval  $[-0.12, 0.07]$ ).



## **R. Clustering and ranking based on BBI and the number of inflorescences**

A Gaussian mixture model was estimated using BBI to estimate clusters of genotypes. Three clusters were identified. The mean BBI within each cluster was close to the mean of the associated cluster using BBI\_res\_norm. However, the cluster sizes were different: the cluster of regular genotypes contained only 3 genotypes (instead of 36), the cluster of alternate bearing genotypes contained 36 genotypes (instead of 31) and the cluster of irregular genotypes contained 83 genotypes (instead of 55). The confusion matrix is given in Table T8. The associated error rate was 0.39. Kendall's  $\tau$  coefficient was 0.63. Genotype 85 was assessed irregular (instead of regular) bearing with rank 65, genotype 107 alternate bearing with rank 96 and genotype 108 irregular bearing with rank 60 (see associated yields and BBI values in Section 3.1).

The cluster switches suggest that BBI does not allow discriminating between trended yields (e.g. for regular genotypes) and irregular yields. This confirms that this index should not be used to quantify irregularity of yields in the setting of trended series.

## **Literature Cited in Supplementary Information**

- Chatfield C.** 2003. The Analysis of Time Series: An Introduction, 6<sup>th</sup> edition. Chapman & Hall/CRC Press, Boca Raton.
- Diggle PJ.** 1990. Time Series: A Biostatistical Introduction. Oxford University Press, Oxford.