Journal of Experimental Botany Advance Access published June 14, 2016

Journal of Experimental Botany
doi:10.1093/jxb/erw203

RESEARCH PAPER

Integrating mixed-effect models into an architectural plant model to simulate inter- and intra-progeny variability: a case study on oil palm (*Elaeis guineensis* Jacq.)

Raphaël P.A. Perez[1], Benoît Pallas[2], Gilles Le Moguédec[3], Hervé Rey[1], Sébastien Griffon[1], Jean-Pierre Caliman[4], Evelyne Costes[2] and Jean Dauzat[1]

[1] CIRAD, UMR AMAP, Montpellier, F-34000 France
[2] INRA, UMR 1334 AGAP, 34398 Montpellier Cedex 5, France
[3] INRA, UMR AMAP, Montpellier, F-34000 France
[4] SMART Research Institute, Pekanbaru 28112, Indonesia

Correspondence: jean.dauzat@cirad.fr

Received 28 January 2016; Accepted 3 May 2016

Editor: Christine Raines, University of Essex

Abstract

Three-dimensional (3D) reconstruction of plants is time-consuming and involves considerable levels of data acquisition. This is possibly one reason why the integration of genetic variability into 3D architectural models has so far been largely overlooked. In this study, an allometry-based approach was developed to account for architectural variability in 3D architectural models of oil palm (*Elaeis guineensis* Jacq.) as a case study. Allometric relationships were used to model architectural traits from individual leaflets to the entire crown while accounting for ontogenetic and morphogenetic gradients. Inter- and intra-progeny variabilities were evaluated for each trait and mixed-effect models were used to estimate the mean and variance parameters required for complete 3D virtual plants. Significant differences in leaf geometry (petiole length, density of leaflets, and rachis curvature) and leaflet morphology (gradients of leaflet length and width) were detected between and within progenies and were modelled in order to generate populations of plants that were consistent with the observed populations. The application of mixed-effect models on allometric relationships highlighted an interesting trade-off between model accuracy and ease of defining parameters for the 3D reconstruction of plants while at the same time integrating their observed variability. Future research will be dedicated to sensitivity analyses coupling the structural model presented here with a radiative balance model in order to identify the key architectural traits involved in light interception efficiency.

Key words: Allometric relationship, *Elaeis guineensis*, genetic variability, mixed-model, plant architecture, three-dimensional reconstruction.

Introduction

Understanding how plants intercept and use solar radiation is a necessary step for enhancing their performance. Plant architecture, defined as the combination of plant topology and organ geometry (Godin *et al.*, 1999), plays a key role in collecting light. Many aerial architectural traits have been shown to influence light interception, such as internode and petiole

length (Takenaka 1994; Sarlikioti *et al.*, 2011), and leaf area density and spatial distribution of leaves (Falster and Westoby 2003; Willaume *et al.*, 2004; Parveau *et al.*, 2008). Plant architecture also affects microclimatic conditions (organ temperature, hygrometry, and light environment), which are known to influence biological and physiological processes such as photosynthesis and leaf transpiration (Niinemets 2007; Vos *et al.*, 2010). Moreover, since plant architecture changes over time, the relevant developmental stages along with temporally variable aspects of morphology and topology must be taken into account when describing plant architecture (Barthélémy and Caraglio, 2007).

Biophysical models (e.g. light interception models, energy balance models) can be applied to three-dimensional (3D) plant representations to evaluate the influence of architectural traits on plant performance. These models can be built from explicit descriptions of plant topology and organ geometry (Vos *et al.*, 2010). One strategy is to record 3D points of interests using digitizing methods (Sinoquet *et al*, 1997; Godin *et al.*, 1999; Sonohat *et al.*, 2006; Louarn *et al.*, 2008); however, digitizing whole-plant architecture is time-consuming and is not adapted to fully describe large plants (Parveau *et al.*, 2008) or many individuals. Alternatively, allometric relationships combined with sampling strategies can be used to reconstruct plant architecture from the scale of the single organ to the entire plant stand (Casella and Sinoquet, 2003; Rey *et al.*, 2008). Such allometric relationships reflect the morphological relationships between plant components at different scales of organization. Recent methods based on image processing or 3D LiDAR scanning are likely to improve data collection efficiency in the future (Phattaralerphong and Sinoquet, 2005; Côté *et al.*, 2009; Hackenberg *et al.*, 2014).

Reducing the time needed for data acquisition is crucial for quantitative genetic studies or plant breeding programmes aiming to study architectural traits (Sakamoto and Matsuoka, 2004; Segura *et al.*, 2006). Studies on different species have demonstrated large genotypic variability in architectural traits and revealed genetic polymorphism associated with this variability (Bradshaw and Stettler, 1995; Plomion *et al.*, 1996; Wu and Stettler, 1998; Wang and Li, 2005; Segura *et al.*, 2008b; Ben Sadok *et al.*, 2013; Li *et al.*, 2015). Inter- and intra-genotypic variability can be estimated by quantitative genetic models. These models allow the estimation of (co)variance components, that is, partitioning of the total observed variance into its causal components, in particular variance due to genetic and environmental effects (Gallais, 1990; Smith *et al.*, 2005). These analyses are mainly based on mixed-effect models and allow the estimation of genotypic values, trait heritability, and genetic correlations between variables (Segura *et al.*, 2008a). Currently, several crop models integrate genotype-dependent parameters related to plant phenology, light interception, light conversion efficiency, or responses to abiotic conditions (Hammer *et al.*, 2010; Casadebaig *et al.*, 2011). In such approaches, genotypes are represented by a set of parameters estimated directly through dedicated experiments and, for the most part, independently of each other (Tardieu, 2003; Lecœur *et al.*, 2011). Others studies also include genetic parameters, combining allelic effects from quantitative trait loci (QTL) with model parameters (Chenu *et al.*, 2009; Letort *et al.*, 2008). Pioneering studies were dedicated to simple plant functions, such as leaf expansion rate (Reymond *et al.*, 2003) or specific leaf area (Yin *et al.*, 1999). More recently, marker-based crop models, estimating values of ecophysiological parameters from genetic markers, were used to explore potential yield improvement and support breeding strategies (Gu *et al.*, 2014). Regarding 3D representations of plants, few models have been calibrated for different genotypes (Casella and Sinoquet, 2003; Rey *et al.*, 2008); so far, none of them have dealt with the genetic control of architectural variability.

Usually, models integrate genotypic differences by quantifying genetic parameters via phenotypic mean values, thus neglecting inter-individual variability (Louarn *et al.*, 2008). Such an approach can be applied when plants are genetically fixed, as in the case of many annual crops or some tree clones (e.g. rubber trees, *Eucalyptus*), but this might lead to oversimplification when progenies have been subjected to large genetic segregation and grown directly in field conditions [e.g. oil palm (*Elaeis guineensis*), maize (*Zea mays*) or *Coffea*)]. In such case, the use of mixed-effect models is particularly interesting because they take account of both inter- and intra-progeny variability.

The principal goal of this study was to account for the architectural variability among individuals and among progenies in a 3D modelling approach. Oil palm (*Elaeis guineensis* Jacq.) is a convenient model for such a study because it exhibits a simple architectural topology following the Corner model, characterized by a mono-axial shoot producing phytomers in regular succession (Halle and Oldman, 1970). An adult oil palm bears 30–50 opened leaves disposed in a radial symmetry (Rees, 1964). Its structural complexity results from its leaf geometry: each leaf is pinnate, being divided into a petiole and a rachis bearing leaflets. The junction of the petiole and rachis (called point C) is recognizable from the presence of small leaflets with vestigial laminae. The rachis cross section is wide and asymmetrical at point C (with a flat adaxial side and a convex abaxial side) and becomes gradually circular from so-called point B (mid-rachis) and point A (rachis extremity). In optimal growing conditions, the number of leaves produced per year varies from 30 to 40 in plants of 2–4 years of age and then declines to 20–25 leaves per year from 8 years old onwards (Corley and Tinker, 2003). Leaf size increases up to the adult stage (8 years) with the result that, for a given individual, leaf size is observed to increase distally along the stem.

The long duration between consecutive generations of oil palm together with the difficulty of producing clonal plants prevents the generation of fixed lines and thus obliges breeders to adopt complex breeding schemes based on biparental crosses between heterozygous parents. Hence, most oil palms cultivated in the world are dura × pisifera crosses, displaying large intra-genotypic variability. Genetic analyses of oil palm have been mainly carried out on yield components or on traits involved in oil and fruit quality (Billote et al., 2010). High heritabilities have been found for quantitative traits related to bunch components and many QTL associated with these traits have been detected (Rance *et al.*, 2001). Moreover, several characters controlled by a single gene have been reported, such as shell thickness, leaflet lamina development (Corley and

Tinker, 2003), and, more recently, oil deterioration (Morcillo *et al.*, 2013). Nevertheless, except for coarsely defined traits related to characteristics such as leaf area, rachis length, or stem height (Rance *et al.*, 2001; Corley and Tinker, 2003), no detailed analysis combining genotypic variability and architectural traits of oil palm has been performed up to now.

The modelling approach presented here couples mixed-effect models with a 3D architectural model based on oil palm. The major architectural traits that are likely to govern light interception (leaf and leaflet geometries) were studied and analysed in terms of their variability between and within progenies. Observations were performed on 60 individuals among five progenies of different genetic origins. Linear and nonlinear allometric relationships were designed for modelling the selected traits and combined with mixed-effect models to explore the significance of intra- and inter-progeny effects. The trait variabilities estimated by these models were finally used to parameterize the reconstruction of 3D mock-ups representative of the variability observed in the field between individuals and between progenies.

Material and methods

Architectural description

The description of the geometry of plant components and their topological arrangement was carried out at two scales of organization: plant scale and leaf scale (Table 1). At the plant scale, attributes related to stem (height H and basal diameter D) and crown (number of leaves and phyllotaxis φ) were defined. At the leaf scale, petiole, rachis, and leaflet geometry was characterized as well as the spatial organization of leaflets along the rachis.

Leaves were topologically positioned along the stem depending on their insertion rank, where leaf rank 1 corresponded to the youngest leaf displaying fully unfolded leaflets (Corley and Tinker, 2003). Three types of attributes were considered to account for leaf geometry: (i) dimensional attributes (rachis length L_{rac} and petiole length Lp); (ii) structural attributes (number of leaflets $NbLft$); and (iii) attributes related to leaf orientation and angle along the rachis. Leaf curvature, deviation, and twist were described respectively by functions of the elevation angle (δ), azimuth angle (Δ), and twist angle (θ) along consecutive segments of rachis (see Fig. 1A). Leaflets were characterized by their dimensions (length L and width w) and their insertion angles on rachis (α and ρ; see Fig. 1B).

Model description

In this experiment, plants had not yet reached the mature stage so leaf size was still increasing along the stem with plant age. The methodology used to describe organ geometry and their changes within the plant topology was based on positional information (Prusinkiewicz *et al.*, 2001). We assumed that, over the considered developmental stage (3–4 years after planting), the allometric relationships governing the shape of the leaf and leaflets were invariant and that only their dimensions evolved with plant age. Conversely, the ratio of petiole length to rachis length ($ratio_L$), the relative position of point B on the rachis ($PosB_{rel}$), and the gradients of leaflet geometry (shape and angles) along the rachis were assumed to be identical for all the leaves of a given individual, at least for the studied plant ages.

Modelling morphogenetic gradients Linear, logistic, and polynomial functions were used to model geometric gradients of plant components according to temporal or spatial variables (Tables 2 and 3). As far as possible, functions were designed parsimoniously (low number of parameters), with parameters related to the observable geometrical properties and minimizing mean square error and bias between observed and simulated values.

Table 1 *Symbols and abbreviations*

Field observations	
Plant scale	
$\sum leaves$	Plant age expressed as the number of leaves emitted since planting date
H_C*	Height of rachis point C from the ground (cm)
H_A*	Height of rachis point A from the ground (cm)
H	Stem height (cm)
D	Stem basal diameter (cm)
φ	Phyllotaxis (degrees)
Leaf scale	
Rk	Leaf rank: spatial position of the leaf on stem (leaf rank 1 at stem top)
L_{rac}	Rachis length (cm)
L_p	Petiole length (cm)
$NbLft$	Number of leaflets per leaf
Pos	Metric position on rachis
δ_C	Declination at point C: angle from the vertical axis to rachis axis at petiole tip (degrees)
δ_A	Declination at point A: angle from the vertical axis to rachis axis at rachis tip (degrees)
δ	Rachis curvature: evolution of declination along rachis (degrees)
Δ	Rachis deviation: projection angle of rachis in an horizontal plan (degrees)
θ	Rachis twist: rotation angle of rachis from the horizontal plan (degrees)
$Area$*	Leaf area (m^2)
$RkLft$	Leaflet rank: rank of leaflet along the rachis (leaflet rank 1 at point C)
L_B	Leaflet length at point B (cm)
W_B	Leaflet maximum width at point B (cm)
L	Leaflet length (cm)
W	Leaflet maximum width (cm)
w	Leaflet width (cm)
α	Leaflet axial insertion: azimuth angle of leaflet midrib projected on the local rachis plan (degrees)
ρ	Leaflet radial insertion: elevation angle of leaflet midrib projected on the local rachis plan (degrees)
$Area_{Lft}$*	Leaflet area (cm^2)
Calculated variables	
$ratio_L$	Ratio of petiole length to rachis length
$FreqLft$	Ratio of leaflets number to rachis length
Pos_{rel}	Relative metric position on rachis
$PosB_{rel}$	Relative metric position of point B on rachis
$PosLft_{rel}$	Relative metric position on leaflet midrib
$RankLft_{rel}$	Relative leaflet rank
L_{rel}	Relative leaflet length (relative to all leaflets on rachis)
W_{rel}	Relative leaflet maximum width (relative to all leaflets on rachis)
w_{rel}	Relative leaflet width (relative to all positions along leaflet midrib)

* variables used to assess model reconstruction

At the plant scale, two variables were introduced to account for morphogenetic gradients of leaves in the crown: the number of leaves emitted from planting date ($\sum leaves$) and the leaf rank (Rk) (Table 1). The evolution of rachis length (L_{rac}) over time was estimated as a linear function of $\sum leaves$ (*eq3* in Table 2 and Table 3). Rk was used to model the evolution of rachis declination at point C along the stem [δ_C (*eq6*)].

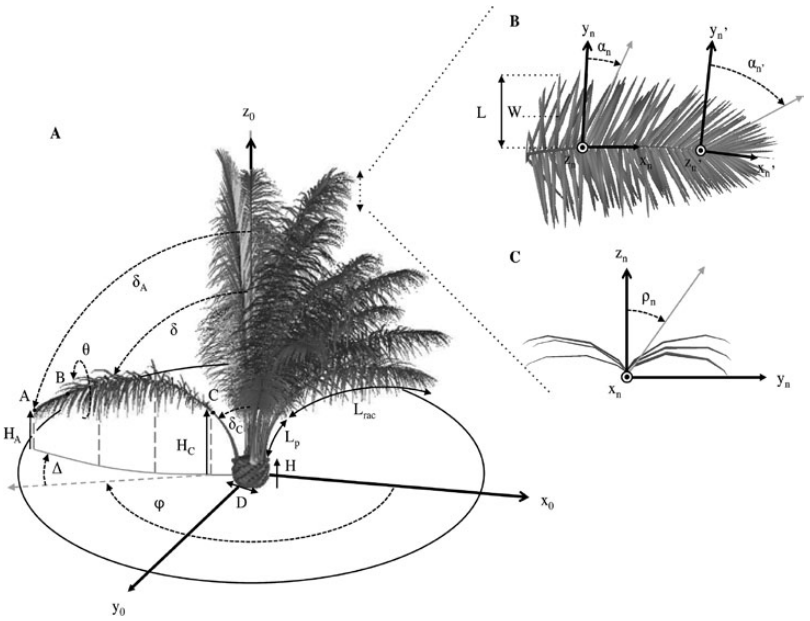

Fig. 1 Geometric variables for assessing and generating 3D oil palm architecture. (**A**) Variables at the plant and leaf scale. Elevation angles (δ) are measured from the vertical reference z_C. Rachis azimuth (Δ) is measured through the projected points along the rachis on (x_C,y_C) plane. Phyllotaxis (φ) is measured as the azimuth angle from one leaf insertion relatively to the following one. Rachis twist (θ) is measured as the rotation angle of the rachis local plane from a vertical plane. (**B**) Detailed top view of a leaf in horizontal plane. Leaflet lengths (L) and widths (W) are measured in a sample of 10 leaflets per leaf. Axial insertion (azimuth angles α_n) is measured with reference to local rachis planes (x_n,y_n). (**C**) Detailed front view of a leaf in a transverse plane to the rachis axis. Radial insertion (elevation angle ρ_n) is measured with reference to local rachis planes (z_n, y_n). Definitions and symbols are given in Table 1.

At the leaf scale, the relative metric position on the rachis (Pos_{rel} = Pos/L_{rac}) was used to describe the evolution of the rachis segment angles [elevation δ (*eq6*, *eq7*, and *eq8*); azimuth Δ (*eq9*); and twist θ (*eq10*)]. Similarly, geometrical attributes [length (*eq14*), width (*eq15*), and insertion angle (*eq19* and *eq20*)] of leaflets were determined according to their relative position along the rachis (Pos_{rel}). Finally, the relative metric position of the leaflet midrib [$PosLft_{rel}$ (*eq11*)] was introduced for modelling leaflet shape (evolution of width, *eq18*).

Modelling organ dimensions Once organ geometry was modelled by allometric functions describing relative proportions (variables with subscript '*rel*'), 'scaling' functions were applied to estimate their absolute value (variables expressed as a function of $\sum leaves$ or L_{rac}; Table 2). The number of leaflets borne by the rachis was predicted by a logistic function of rachis length (*eq5*; Table 2).

Leaflet dimensions were estimated from linear relationships between rachis length and leaflet dimension at point B (*eq12* and *eq13*). Absolute dimensions of leaflets (L, W) along the rachis were estimated using their relative values (L_{rel}, W_{rel}), that is, relative to the longest (largest) leaflets on the rachis, and rescaling them using the absolute values L_B and W_B (*eq14* and *15*).

Model calibration

Plant material and growing conditions Measurements were performed at an experimental plantation of the SMART Research Institute (SMARTRI, Smart Tbk.) located in South Sumatra province, Indonesia (2° 59′ 27.99″ S, 104° 45′ 24.24″ E). The trial was set up in 2010, ~18 months after seedling germination. The genetic material studied was composed of 25 progenies of tenera hybrids selected by the PalmElit Company and SMARTRI using several criteria: production of fresh fruit bunches, oil yield, stem growth, precocity of production, and parent origins. The experimental design was a Fisher block design of five blocks subdivided into 25 elementary parcels, each parcel including 25 trees of the same progeny (see Supplementary Fig. S1). The planting density was 136 plants ha^{-1} in a 9.5 m equilateral triangular pattern whatever the progeny. For this study, we selected five progenies (hereafter referred to as DA1, DL7, DS, DU, and DY4) in view of their different morphologies, diversity of origins (Asian and African origins), and their architecture. All progenies were known to have a good production performance (see Supplementary Table S1). The site is characterized by a tropical humid climate, considered as optimal conditions for oil palm cultivation.

Data collection Architectural measurements were performed on plants located in the same experimental block in order to reduce sources of heterogeneity (see Supplementary Fig. S1). Every 6 months from December 2010 to November 2014, coarse-scale measurements (rachis length, stem basal diameter, and number of leaflets) were made for each individual (5 progenies × 25 plants × 6 dates). The number of observations per progeny was dependent on the type of measurements (see Supplementary Table S2). Numerous detailed measurements were subsequently collected 39 months after planting (MAP; April 2014) and used to define the allometric relationships. A second set of data was collected at 47 MAP (November 2014) for a larger number of individuals (between 6 and 12 plants per progeny) for a more detailed assessment of trait variability among progenies. For allometric relationships related to ontogenetic gradients, that is, dependent on plant age, model calibration was performed on data collected from 6 to 47 MAP.

Estimated variables Analyses of data collected at 39 MAP showed that sampling 10 leaflets was sufficient to simulate accurately the leaf area and the leaf shape (see Supplementary Fig. S1). Similarly, the marked leaf symmetry observed meant that we could limit measurements to only one side of the leaf (see Supplementary Figs S2 and S3).

Each leaf was labelled as soon as it was fully open, thus enabling us to count the leaves emitted per plant since planting date ($\sum leaves$). Leaf area ($Area$) was estimated by dividing the leaf into 10 equal sections along the rachis. On each section, the number of leaflets was counted on both sides and a median leaflet was chosen for which length and width of segments (in five regular intervals along the leaflet midrib) were measured to estimate the entire individual leaflet area ($Area_{Lft}$). This approach considered the leaflet as a sum of trapezes. For each rachis section, total leaflet area was approximated by multiplying the individual leaflet area by the number of leaflets on the corresponding section. Leaf area was finally obtained as the combined sum area of the 10 sections (Talliez and Ballo Koffi, 1992).

Leaf curvature and deviation along the rachis were estimated by measuring distances between control points along the rachis

Table 2 *Allometric relationships (functions f_1 to f_{10} refer to the functions presented in Table 3). See Table 1 for variable meanings*

Predicted variable	Explanative variable	Equation	Parameters meaning
Stem scale			
H (cm)	$\sum leaves$	(*eq 1*) $H = f_3(\sum leaves)$	h_0: stem height at planting date hg: growth rate factor
D (cm)	L_{rac}	(*eq 2*) $D = f_2(L_{rac})$	D_{max}: maximum basal diameter D_{slp}: slope factor at inflexion point $L_{D\ infl}$: rachis length at inflexion points
Leaf scale			
L_{rac} (cm)	$\sum leaves$	(*eq 3*) $L_{rac} = f_1(\sum leaves)$	$L_{rac\ int}$: intercept $L_{rac\ slp}$: slope
L_p (cm)	L_{rac} (*cm*)	(*eq 4*) $L_p = ratio_L . L_{rac}$	$ratio_L$: ratio of petiole length to rachis length
$NbLft$	L_{rac} (*cm*)	(*eq 5*) $NbLft = f_2(L_{rac})$	Nb_{max}: maximum number of leaflets per leaf Nb_{slp}: slope factor at inflexion point $L_{Nb\ infl}$: rachis length at inflexion points
δ_C (°)	Rk	(*eq 6*) $\delta_C = f_1(Rk)$	$\delta_{C\ int}$: intercept $\delta_{C\ slp}$: slope
δ_A (°)	δ_C (°)	(*eq 7*) $\delta_A = f_2(\delta_C)$	$\delta_{A\ max}$: maximum declination at point A $\delta_{A\ slp}$: slope factor at inflexion point $\delta_{A\ infl}$: δ_C angle at inflexion points
δ (°)	Pos_{rel}	(*eq 8*) $\delta = \delta_C + f_4(Pos_{rel})\ (\delta_A - \delta_C)$	Δ_{sf}: evolution of curvature along the rachis
Δ (°)	Pos_{rel}	(*eq 9*) $\Delta = \Delta_a . f_4(Pos_{rel})$	Δ_a: deviation angle at point A Δ_{sf}: evolution of deviation along the rachis
θ (°)	Pos_{rel}	(*eq 10*) $\theta = f_5(Pos_{rel})$	θ_a: twist angle at point A θ_s: evolution of twist along the rachis
$PosLft_{rel}$	$RankLft_{rel}$	(*eq 11*) $PosLft_{rel} = f_4(RankLft_{rel})$	d_{Lft}: evolution of inter-leaflet distance along the rachis
L_B (cm)	Pos_{rel}	(*eq 12*) $L_B = f_1(L_{rac})$	$L_{B\ int}$: intercept $L_{B\ slp}$: slope
W_B (cm)	Pos_{rel}	(*eq 13*) $W_B = f_1(L_{rac})$	$W_{B\ int}$: intercept $W_{B\ slp}$: slope
L_{rel}	Pos_{rel}	(*eq 14*) $L_{rel} = f6(Pos_{rel})$	l_c: L_{rel} at point C p_L: relative position of the longest leaflet on rachis l_a: L_{rel} at point A
W_{rel}	Pos_{rel}	(*eq 15*) $W_{rel} = f_7(Pos_{rel})$	w_c: W_{rel} at point C p_W: relative position of the largest leaflet on rachis w_a: W_{rel} at point A
L (cm)	Pos_{rel}	(*eq 16*) $L = L_{rel}\ L_B / f_6(PosBrel)$	
W (cm)	Pos_{rel}	(*eq 17*) $W = W_{rel} . W_B / f_7(PosBrel)$	$PosB_{rel}$: relative position of point B on rachis
$Wrel$	Pos_{rel}, $PosLft_{rel}$	(*eq 18*) $w_{rel} = f_8(Pos_{rel}, PosLft_{rel})$	p_w: relative position of maximum width on leaflet sl: leaflet shape factor
α (°)	Pos_{rel}	(*eq 19*) $\alpha = f_9(Pos_{rel})$	α_c: leaflet axial insertion angle at point C α_s: decreasing factor of axial angle along rachis α_a: leaflet axial insertion angle at point A
ρ (°)	Pos_{rel}	(*eq 20*) $\rho = f_{10}(Pos_{rel})$	ρ_c: leaflet radial insertion angle at point C $\rho_{0.5}$: radial insertion angle on middle rachis length

and their projections on a horizontal plane (10 points per leaf). Projection distances were used afterwards to estimate deviation and elevation angles of leaves along the stem.

Analyses of inter-individual variability and differences among progenies A first analysis was aimed at assessing architectural variability between and within progenies at 47 MAP (Fig. 1). For the variables not related to rachis length ($\sum leaves$, Φ, H, and ∂_C), one-way ANOVA were performed with progeny effect. Conversely, when variables were correlated with rachis length, analyses of covariance (ANCOVA) were performed without interaction but considering a genotype and a rachis length effect. For all the variables, Tukey's tests were used for *post hoc* comparisons. The homoscedasticity of variables and normality of model residuals were verified using Levene's and Shapiro–Wilks' tests respectively.

For evaluating the variability of a trait within a progeny and comparing it to the variability between progenies, we defined an inter-family broad-sense heritability (h^2) as the ratio of progeny variance to total variance ($h^2 = \sigma^2_{progeny}/\sigma^2_{total}$) (Gallais, 1990). This index was calculated for each phenotypic trait at 47 MAP, using the restricted maximum likelihood method (Corbeil and Searle, 1976) to estimate progeny and residual variances (the total variance being given by the sum of both progeny and residual variances). In a second step, the study focused on variables affected by ontogenetic and morphogenetic gradients. Because these variables change over time and space, statistical analyses were performed directly on the allometric relationships to test for differences among and within progenies. Allometric relationships were adjusted on different data sets: all data gathered (null model), data sorted per progeny (progeny model), and data sorted per plant (individual model). A likelihood ratio test was then carried out using a Chi-squared test (χ^2) to compare models (null, progeny, and individual) and to assess inter-progeny and intra-progeny effects. Likelihoods of progeny models were calculated in reference to the total variance, whereas likelihoods of individual models were calculated in reference to the intra-progeny variance for each progeny.

Table 3 *Functions used in the allometric relationships*

$$f_1(x) = int + x \times slp$$

$$f_2(x) = \frac{max}{1 + e^{4 \times slp(x_{infl} - x)}}$$

$$f_3(x) = h_0 e^{h_g \times x}$$

$$f_4(x) = \frac{(1+sf)x^2}{1 + sf \times x^2}$$

$$f_5(x) = \theta_a x^{3\theta_s}$$

$$f_6(x) = \begin{cases} l_c + 2\frac{(1-l_c)}{p_L}x + \frac{(l_c-1)}{{p_L}^2}x^2, & \text{if } x \le p_L \\ 1 + \frac{(l_a-1)}{(1-p_L)^2}(x-p_L)^2, & \text{if } x > p_L \end{cases} \quad \text{and } p_L \in]0,1[$$

$$f_7(x) = \begin{cases} w_c + \frac{(1-w_c)}{p_W}x & \text{if } x \le p_W \\ \frac{1 - w_a \times p_W + (w_a - 1)x}{1 - p_W}, & \text{if } x > p_W \end{cases} \quad \text{and } p_W \in]0,1[$$

$$f_8(u,v) = \frac{v^{p-1}(1-v)^{q-1}}{p_w}$$

where:

p and q are functions of p_w and sl

$p_w = f_1(u)$

$sl = f_1(u)$

$$f_9(x) = \sqrt{{\alpha_c}^2 + 2\alpha_c \alpha_s x + \left[{\alpha_a}^2 - \alpha_c^2 + 2\alpha_c \alpha_a\right]x^3}$$

$$f_{10}(x) = \begin{cases} \rho_c + 4x(\rho_{0.5} - \rho_c)(1-x), & \text{if } x \le 0.5 \\ 4bx(x-1), & \text{if } x > 0.5 \end{cases}$$

The number of estimated parameters varied depending on the considered function and the significance of progeny and individual effects. If both progeny and individual effects were significant, model parameters were then estimated by performing hierarchical mixed-effects models, considering the individual plant (intra-progeny) effect as a random effect nested within a fixed progeny effect, expressed as a matrix (Pinheiro and Bates, 2000) as:

$$y_{ijk} = f\left(x_{ijk}; \varphi_{ij}\right) + \varepsilon_{ijk} \quad \varepsilon_{ijk} \sim N(0, \sigma^2)$$

$$\varphi_{ij} = A_{ij}\beta_i + B_{ij}b_{ij}, b_{ij} \sim N\left(0, \psi_i\right), \; b_{ij} \sim N(0, \psi_i)$$

where f represents one of the allometric relationships presented (Table 3), y_{ijk} labels the k^{th} observation of the j^{th} individual of the i^{th} progeny, x_{ijk} is the covariate vector related to this observation, and ε_{ijk} represents model residuals (assumed to be independent and identically distributed). The vector φ_{ij} represents the model parameters associated with the j^{th} individual of the i^{th} progeny, β_i is the vector of fixed effects related to the i^{th} progeny, and b_{ij} is the random effect vector associated with the j^{th} individual of progeny i. In other words, b_{ij} represents the deviation of the φ_{ij} from the mean parameter β_i due to the j^{th} individual. A_{ij} and B_{ij} are incidence matrices and ψ_i is the variance-covariance matrix associated with the progeny i. Consequently, for each function used to predict trait values, the progeny effect is related to the mean parameter (β_i) of the model whereas the individual effect defines variance parameters (ψ_i) of the model. In cases where only the inter-progeny effect was significant, only mean parameters and model residuals were estimated for each progeny.

The inter-progeny coefficient of variation (CV) was calculated by dividing the SD of the mean values of each progeny by the overall mean. The intra-progeny CV was calculated for each progeny as the SD of individual parameters estimated by the mixed-effect model ($\sqrt{\psi_i}$) divided by the corresponding mean value.

Model simulation and validations

Simulations of palm mock-up A dedicated oil palm simulation model (VPalm) was developed using the basis of a former simulator of coconut palms (Dauzat and Eroy, 1997). VPalm was written using object-oriented programming in Java language as an application of the AMAPstudio software suite (Griffon and de Coligny, 2014). The VPalm simulator enabled us to rebuild the topological structure of the palm through decomposition into elementary components organized along a multiscale tree graph (Godin and Caraglio, 1998). The simulator was designed for integrating the allometric relationships (Table 2) needed to render the plant topology and its 3D geometry. Each individual palm was reconstructed from an input file generated to account for the progeny parameterization as well as individual variability. The random sampling procedure of R (R Development Core Team, 2015) was used to generate random individual parameters by combining estimated mean parameters associated with progeny effect (β_i) with variance-covariance matrices associated with individual effect (ψ_i) when significant. Even if significant, individual effects were not considered if the explanatory variables of the allometric relationship were estimated using individual effects (like *NbLft* for instance) to avoid any over-parameterization. In other words, we assumed that the variance component linked to the explanatory variable (e.g. rachis length) was sufficiently spread into the response and consequently did not require the estimation of individual variance components. Twenty-five random VPalm parameters files were created in this way to generate 25 virtual individuals for each progeny that were subsequently laid out to reproduce the experimental parcels at 47 MAP.

Assessing model reconstruction Different variables were extracted from the 3D simulated mock-ups with the Xplo software of AMAPstudio to compare their value with field observations in terms of root mean square error (RMSE), normalized RMSE (NRMSE), and bias, defined as follows:

$$RMSE = \sqrt{\sum^n_{i=1}(s_i - m_i)^2} / \sqrt{n} \quad NRMSE = RMSE / (\sum^n_{i=1} m_i / n)$$
$$Bias = \sum^n_{i=1}(s_i - m_i) / n$$

with s_i and m_i the i^{th} simulated and measured values and n the number of observations.

The accuracy of model prediction was evaluated for variables related to leaf and leaflet geometry (rachis and petiole lengths, leaflet length and width, leaf and leaflet angles). Inspecting the potential errors resulting from the successive assembly of allometric relationships was crucial. As an example, the area of leaflets along the rachis ($Area_{Lft}$) combined several allometric relationships (*eq3*, *eq5*, and *eq11–18* in Table 2) needed to reproduce accurately morphogenetic gradients. Similarly, the height of the rachis tip (H_A; Fig. 1) depended on many intermediate variables (stem height, leaf length, and leaf curvature) and we therefore checked the simulated values against measurements for different leaf ranks. Finally, the simulated variances computed after running 25 random simulations were compared to the observed variances using Fisher's test.

All statistical analyses presented above were performed with R software and the parameters of mixed-effect models were estimated using the 'nlme' package of R.

Results

Progeny effect at 47 months after planting

At the plant scale, progeny effect was highly significant ($P < 0.001$) for the number of leaves emitted since planting

date ($\sum leaves$) and stem basal diameter (D). However, no effect was found for phyllotaxis nor stem height ($P > 0.05$; Table 4). Stem diameter was significantly smaller for the progeny DU, which also emitted a lower number of leaves from planting date. Important and significant variability was observed in the number of emitted leaves between progenies (103–121 leaves between plants), with DA1 and DS developing significantly more leaves than other progenies.

At the leaf scale, difference between progenies were significant ($P < 0.05$) for all variables. Leaf area and the ratio of petiole length to rachis length showed the most important variability between progenies and the highest inter-family heritabilities (0.63 for the ratio and 0.65 for leaf area, Table 4). DA1 developed leaves with a petiole accounting for a third of the total leaf length ($ratio_L = 0.32$) whereas DS displayed the smallest petioles ($ratio_L = 0.23$). The highest density of leaflets ($FreqLft$) was observed for progeny DS (0.79 leaflet cm^{-1}) and progeny DL7 displayed the lowest leaf area. Finally, ∂_C displayed large intra-progeny variability (CV = 0.15), leading to a weak inter-family heritability estimated for this trait ($h^2 = 0.12$).

For leaflet dimensions (L_B and W_B), progenies exhibited significant differences and high heritabilities were observed ($h^2 > 0.4$). Progeny DU had the largest leaflets whereas the longest ones were observed for progeny DA1.

Progeny and individual effects on allometry

Likelihood ratio tests between nested models (null model, progeny model, individual model) highlighted the significance ($P < 0.001$) of progeny effects for all the studied variables. For instance, at the plant scale, the growth rate parameter h_g was significantly different between progenies ($h_g = 0.021$ for DA1 and $h_g = 0.025$ for the other progenies). At the leaf scale, the tendency to increase leaf length during plant development ($Lrac_{slp}$) displayed low variability between progenies (Fig. 2A; CV = 0.06). The variability in the declination of the leaf at point C (δ_C) with leaf rank indicated that trends in leaf bending along the stem varied by progeny (Fig. 2B). The progeny DU presented a steep increase in δ_C with leaf rank ($\delta_{Cslp} = 1.67°$ rank^{-1}) whereas DS displayed a slower increase in δ_C ($\delta_{Cslp} = 1.33°$ rank^{-1}). For leaflet shape, the relative position of the longest leaflet (p_{wL}) and the relative position of the largest leaflet (p_w) presented low variabilities between progenies compared to leaflet length and width at rachis extremities (l_c, l_a, w_c, and w_a) (Fig. 2C, D). Parameter values per progeny for all the allometric relationships are summed up in Supplementary Table S3.

Likelihood ratio tests revealed significant individual effects for all variables except for stem basal diameter (Table 5). Some variables ($Lrac$, δ_C, L_{rel}, and W_{rel}) showed a highly significant individual effect for all progenies, contrary to other variables ($NbLft$, L_B, W_B) for which intra-progeny differences were only significant for some progenies. The highest intra-progeny variability was detected for the leaflet relative length at point C (lc), which exhibited a CV varying from 0.21 to 0.92 within progenies.

Assessment of model reconstruction

Assessment of mean prediction per progeny As expected, when data were used directly to calibrate the model (Fig. 3A), the predictions were close to observations with a NRMSE < 0.08 and a low bias. Simulated petiole length (Lp) was slightly overestimated (bias = 5.79 cm), probably because the ratio of petiole to rachis length was calibrated using older leaves that were more accessible for measurements than those used for model validation.

Table 4 *Mean and standard deviation (in parentheses) observed per progeny and estimated heritability (broad sense) for the variables (data collected 47 MAP). Leaf rank refers to the ranks on which observations were done. Significance levels of progeny effects correspond to the P-value of ANOVA and ANCOVA*

		Progeny					
Variables	**Leaf rank**	**DA1**	**DL7**	**DS**	**DU**	**DY4**	**Heritability**
Plant scale							
$\sum leaves$	-	116 (3) a	111 (2) b	115 (2) a	108 (4) c	109 (3) bc	0.58 ***
φ (°)	-	136.7 (0.5)	137.1 (1)	136.9 (1)	137.0 (0.7)	136.6 (0.4)	<0.01 n.s.
H (cm)	-	90 (9)	84 (13)	86 (9)	78 (13)	84 (20)	0.02 n.s.
D (cm)	-	75 (6) a	78 (7) a	74 (6) a	63 (4) b	80 (7) a	0.53 ***
Leaf scale							
L_{rac} (cm)	32 ± 6	340 (28) b	326 (18) b	340 (20) c	343 (23) ab	352 (34) a	0.21 ***
$ratio_L$	32 ± 6	0.32 (0.03) a	0.25 (0.02) b	0.23 (0.02) c	0.27 (0.02) b	0.25 (0.03) b	0.63 ***†
L_p (cm)	32 ± 6	96 (18) a	73 (16) b	71 (15) c	76 (21) bc	87 (13) ab	0.29 ***†
$NbLft$	34 ± 6	247 (7) bc	243 (10) c	263 (16) a	244 (16) c	255 (8) ab	0.26 ***†
$FreqLft$	34 ± 6	0.72 (0.03) c	0.76 (0.05) ab	0.79 (0.07) a	0.73 (0.05) b	0.73 (0.06) b	0.28 ***†
δ_c (°)	17 ± 2	34 (5) a	38 (4) a	35 (7) a	39 (5) a	37 (5) a	0.12 *
L_B (cm)	31 ± 6	79 (6) a	69 (4) c	77 (5) ab	75 (3) b	76 (5) b	0.41 ***†
W_B (cm)	31 ± 6	4.3 (0.4) bc	4.1 (0.3) c	4.4 (0.3) b	5.0 (0.3) a	4.4 (0.4) b	0.44 ***†
$Area$ (m^2)	30 ± 4	3.7 (0.2) a	2.7 (0.2) b	3.3 (0.4) a	3.5 (0.1) a	3.6 (0.3) a	0.65 ***†

Progeny with different letters are significantly different (Tukey's test $P < 0.05$). For variable abbreviations see Table 1. †Variable on which an ANCOVA is performed with progeny factor and rachis length as covariable. n.s., non-significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

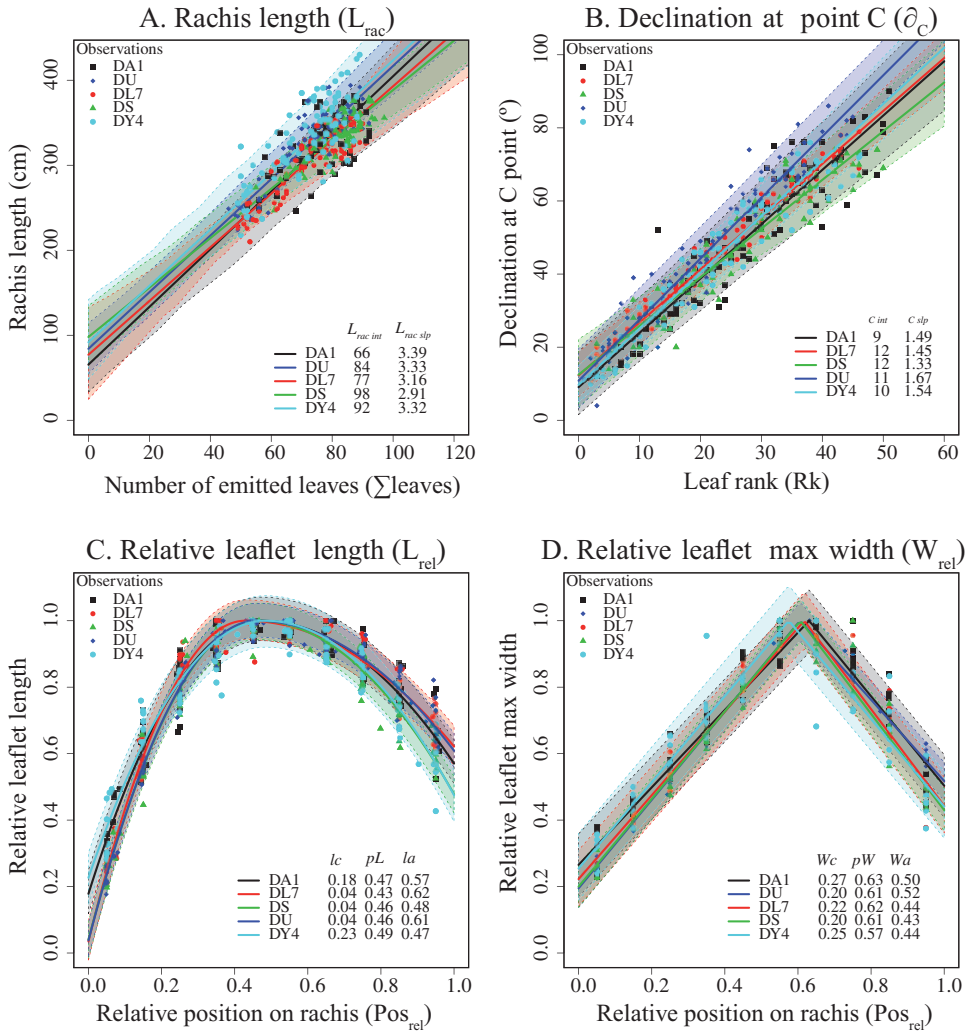

Fig. 2 Comparison of fitted curve per progeny (solid lines) for the rachis length (L_{rac}) (**A**), the declination at point C (δ_C) (**B**), the relative leaflet length (L_{rel}) (**C**), and the relative leaflets width (W_{rel}) (**D**). Limits represent the distribution of individual effects (95% confidence intervals estimated from 100 simulations for each progeny). Mean values of parameters are presented at the lower right of each graphic. See Table 1 for variable abbreviations, Table 2 for parameter abbreviations, and Table 3 for the equation used for fittings.

Regarding the variables simulated from a combination of various allometric relationships (Fig. 3B), greater discrepancies were noted, with greater NRMSE values ranging from 0.09 to 0.19. The most important differences between observations and simulations were observed for rachis heights at points A and C. Leaf area at rank 17 was overestimated on average (bias = 0.23 m^2), mainly due to progeny DU showing important dissimilarities with observations (NRMSE = 0.16 for this progeny).

Simulation of morphogenetic gradients within the canopy The average predictions of leaflet area along the rachis were in accordance with observations, with an RMSE varying from 19 to 27 cm^2 and with a bias <16 cm^2 (Fig. 4A).

Predictions of leaf areas according to their position on the stem were accurate (NRMSE ≤ 0.16) and with low bias except for progeny DU, which displayed, on average, a larger simulated leaf area than observed (bias = 0.57; Fig. 4B).

Regarding the development of rachis height at point C (H_C) and point A (H_A) with leaf rank, H_A was slightly underestimated on average (bias = −13.52 cm) with an NRMSE < 0.2 (Fig. 3B). Similarly, the decrease in H_C (Fig. 4D) was correctly simulated, particularly for progenies DA1 and DY4, but more errors were detected for progeny DL7 and DS (bias > 10 cm).

Assessment of variance prediction per progeny 3D mock-ups of each studied progeny (Fig. 5) revealed the capacity of the modelling approach to simulate the architectural genotypic characteristics described above. The quality of variance prediction was assessed for each trait by analysing the ratio of simulated SD from 25 mock-ups to the observed SD (Table 6). No significant difference was reported between observed and simulated SD for rachis length and the declination at point C. Likewise, no difference was noticed for H_A and H_C (ratio varying from 0.83 to 1.85).

Slight differences were observed for the number of leaflets for which the predicted SD was higher than that observed for progenies DA1 and DY4 (ratio = 1.63). Conversely, simulated variances were lower than observed for petiole length and leaflet width (Fig. 3A); however, these differences were non-significant for all progenies.

Table 5 *Coefficient of variation per progeny and within progenies of variables and parameters associated with allometric relationships. Significance levels of progeny correspond to the P-value of the likelihood ratio tests between null and progeny models. Significance levels of individual effect correspond to the P-value of the likelihood ratio tests between progeny and individual models*

			Individual effect				
Variables	**Parameters**	**Progeny effect**	**DA1**	**DL7**	**DS**	**DU**	**DY4**
Plant scale							
H		***	*	*	n.s.	*	n.s.
	h_0 [a]	0	0	0	0	0	0
	h_g	0.07	0.07	0.05	0.03	0.05	0.08
D		***	n.s.	n.s.	n.s.	n.s.	n.s.
	D_{max}	0.05	0.06	0.10	0.06	0.19	0.11
	D_{slp}	0.18	0.17	0.50	0.34	0.21	0.31
	$L_{D\ infl}$	0.04	0.07	0.07	0.07	0.13	0.13
Leaf scale							
L_{rac}		***	***	***	***	***	***
	$L_{rac\ int}$	0.15	0.78	0.65	0.72	0.39	0.52
	$L_{rac\ slp}$	0.06	0.20	0.25	0.26	0.15	0.24
$NbLft$		***	**	n.s.	n.s.	n.s.	n.s.
	Nb_{max}	0.04	0.16	0.07	0.13	0.18	0.08
	Nb_{slp}	0.21	0.42	0.30	0.34	0.25	0.23
	$L_{Nb\ infl}$	0.01	0.15	0.05	0.13	0.18	0.07
δ_c		***	***	***	***	***	***
	$\delta_{c\ int}$	0.13	0.42	0.29	0.41	0.45	0.25
	$\delta_{c\ slp}$	0.09	0.09	0.08	0.15	0.09	0.10
L_B		***	***	*	**	n.s.	**
	$L_{B\ int}$	0.16	0.48	0.36	0.31	0.58	0.20
	$L_{B\ slp}$	0.11	0.21	0.25	0.28	0.17	0.33
W_B		***	n.s.	*	n.s.	n.s.	*
	$W_{B\ int}$	0.18	0.16	0.33	0.18	0.40	0.16
	$W_{B\ slp}$	0.31	0.23	0.34	0.64	0.34	0.68
L_{rel}		***	***	***	***	***	***
	l_c	0.84	0.21	0.67	0.92	0.92	0.51
	p_L	0.05	0.16	0.11	0.07	0.04	0.13
	l_a	0.13	0.14	0.12	0.13	0.14	0.13
W_{rel}		***	***	***	**	***	***
	w_c	0.18	0.09	0.13	0.17	0.13	0.22
	p_W	0.02	0.05	0.06	0.03	0.03	0.07
	w_a	0.09	0.15	0.16	0.08	0.12	0.14

[a]Stem height at planting date (h_0) was fixed at 5 cm for all progenies. n.s., non-significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). See Table 1 for variable abbreviations and Table 2 for parameter abbreviations.

Finally, the simulated leaf area SD was higher than that observed for all progenies except for progeny DS. This difference was only significant for progeny DU, in which the simulated SD was more than three times higher than the observed SD.

Discussion

Genetic control of plant architecture

The present study highlighted significant progeny effect for all studied architectural traits of young plants except for phyllotaxis and stem height (Table 4). These results are in accordance with a study by Billote et al. (2010) on adult oil palms. These authors did not detect a significant difference in stem height among 15 crosses, whereas high genetic variability was highlighted for leaf and leaflet dimensions (rachis length, number of leaflets, leaflet length, and leaflet width).

In other species, high heritabilities of plant height and diameter have been previously observed in several dicotyledons, such as *Populus*, *Eucalyptus*, and apple trees (*Malus domesticus*) (Bradshaw and Stettler, 1995; Byrne *et al.*, 1997; Wu and Stettler, 1998; Osorio et al., 2000; Segura *et al.*, 2006), and monocotyledons, such as rice (*Oryza sativa*) and maize (Hung *et al.*, 2012; Yang and Hwa, 2008). In these studies, heritability values were computed as the ratio of genotypic variance to phenotypic variance. In our study, the same estimation of heritability was not possible because progenies were grown without any replicate of each genotype. An inter-family broad-sense heritability (ratio of progeny variance to phenotypic variance) was thus presented as an index to estimate the stability of the architectural traits within families. This inter-family broad-sense heritability estimated at a given plant age was high for stem diameter but close to zero for stem height. Stem height was, however, significantly different between progenies when taking into account the number

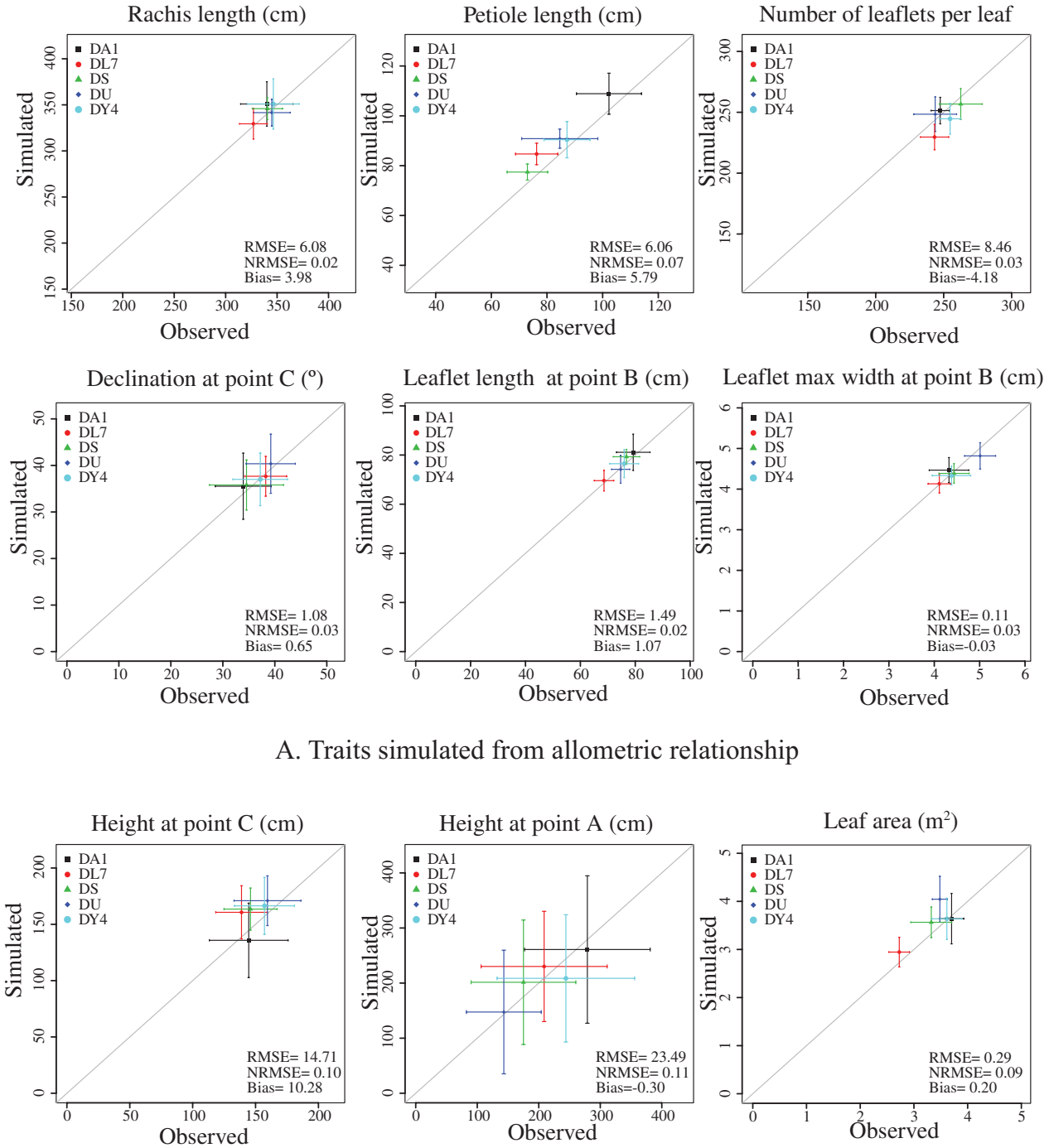

Fig. 3 Comparison between measured and simulated variables for each progeny 47 MAP (points represent mean values, crossed bars represent the range of standard deviation around mean value). Simulated variables are extracted from 25 mock-ups generated by VPalm. Declination angle at point C are represented for leaves located at rank 17 ± 2. For the other variables, the ranks of simulated leaves correspond to the ranks of observed leaves (30 ± 5 in average). Solid line represents the 1:1 line. This figure is available in colour at *JXB* online.

of leaves produced since planting date (Table 5), suggesting potential differences in internode lengths.

At the leaf scale, the main difference among progenies was the length of the petiole relative to leaf length ($ratio_L$), and this ratio was found to be the most stable trait within progenies (h^2 = 0.63) together with leaf area (h^2 = 0.65). Interestingly, $ratio_L$ was much higher for the progeny DA1, which was the only family selected from an Asian pedigree. Such results indicate the importance of genetic control on leaf morphology, as has been observed in other species for leaf length and width (Frary *et al.*, 2004; Hung *et al.*, 2012) and leaf area (Byrne *et al.*, 1997; Wu and Stettler, 1998). Conversely, the low heritability found for the declination at point C contrasts with previous studies that showed high heritability of leaf angle in a maize

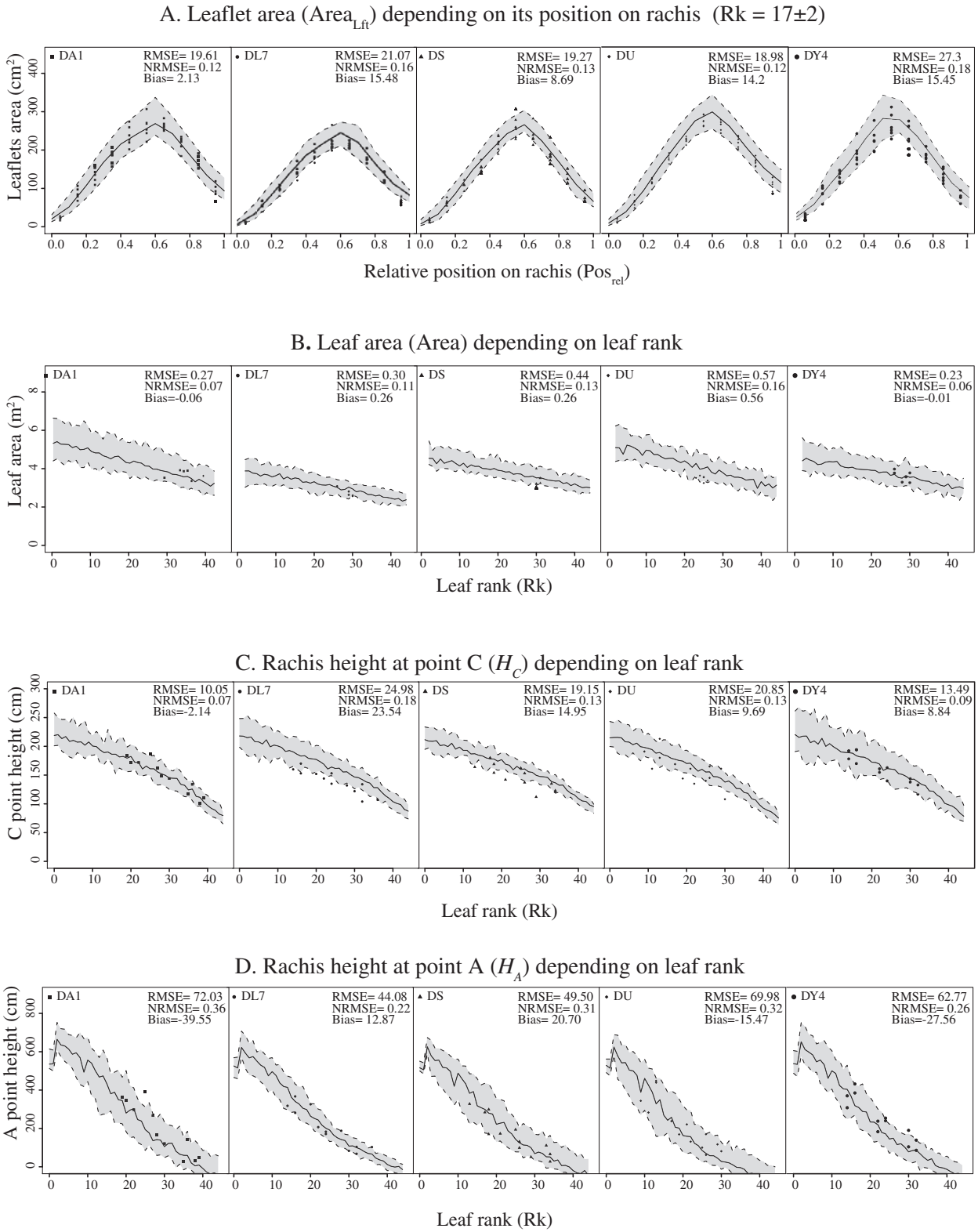

Fig. 4 Evaluation of model reconstruction for five progenies. Mean (continuous lines) and 95% confidence interval (envelopes) of predictions are calculated from 25 plants simulated with VPalm for each progeny. Points represent observed values.

population (Hung *et al.*, 2012) and genetic control of leaf curvature in *Arabidopsis thaliana* (Serrano-Cartagena *et al.*, 1999) and rice based on analyses of mutants (Yang and Hwa, 2008).

In the present study, the significant intra-progeny variability may be explained by genetic segregation and/or by soil and resource heterogeneity within the field (Welham *et al.*, 2002). Indeed, even if all plants were grown in the same block, the environmental variability between parcels could be confused with an inter-progeny genetic effect on the architecture due to the experimental design. The absence of a linked pedigree between

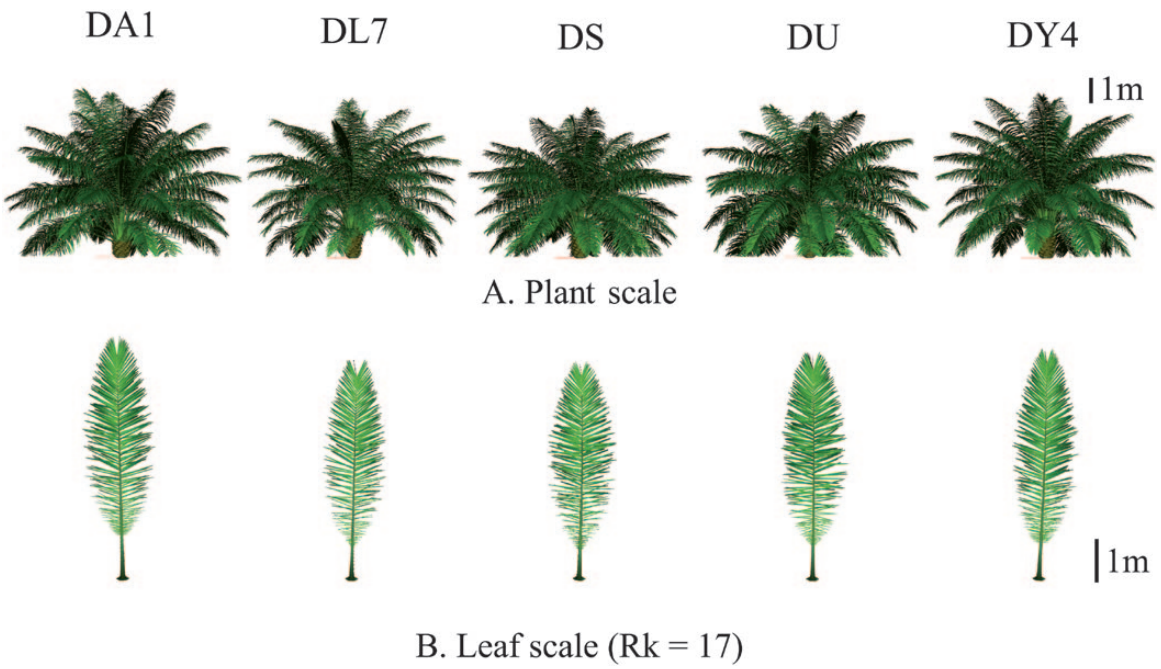

Fig. 5 Three-dimensional mock-ups simulated with VPalm. (**A**) Representation of an average plant for each progeny (generated from parameters mean values per progeny). (**B**) Top view of the 17th expanded leaf taken from the plant in panel (A). This figure is available in colour at *JXB* online.

Table 6 *Comparison between observed and simulated variances 47 MAP. The table shows the ratio of simulated standard deviation to observed standard deviation. Leaf rank refers to the ranks on which observations were done. Significance levels correspond to the P-value of the Fisher test between observed and simulated variance*

		Progeny				
Variables	**Leaf rank**	**DA1**	**DL7**	**DS**	**DU**	**DY4**
L_{rac}	32±6	0.95 n.s.	1.23 n.s.	0.81 n.s.	0.81 n.s.	1.08 n.s.
L_p	32±6	0.70 n.s.	0.57 *	0.44 **	0.28 ***	0.88 n.s.
NbLft	34±6	1.63 *	1.01 n.s.	0.81 n.s.	0.90 n.s.	1.62 *
δ_C	17±2	1.32 n.s.	1.06 n.s.	0.76 n.s.	1.34 n.s.	1.07 n.s.
L_B	31±6	1.20 n.s.	1.18 n.s.	0.63 ***	1.72 **	1.08 n.s.
W_B	31±6	0.71 **	0.91 n.s.	0.72 **	0.95 n.s.	0.54 ***
H_C	24±8	1.06 n.s.	1.15 n.s.	0.89 n.s.	0.83 n.s.	1.06 n.s.
H_A	24±8	1.31 n.s.	0.98 n.s.	1.33 n.s.	1.85 n.s.	1.03 n.s.
Area	30±4	2.29 n.s.	1.61 n.s.	0.83 n.s.	3.38 *	1.53 n.s.

n.s., non-significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. For variable abbreviations see Table 1.

the studied families, and the lack of genetic information as well as information on soil characteristics, meant that we were not able to separate genetic effect from environmental effect. In the future, the integration of architectural phenotypic data in an experimental design involving crosses with known pedigree or clones could lead to a better depiction of the genetic control (QTL analysis) of architectural traits as previously performed for production variables in oil palm (Tisné *et al.*, 2015).

Using allometry to analyse genotypic variability

The use of distinct response curves has been proposed as a way to account for the genetic variability of responses to environmental conditions in plant models (Tardieu, 2003). Likewise, allometric relationships have been used for modelling the architecture of different genotypes (Casella and Sinoquet, 2003; Rey *et al.*, 2008). Our allometry-based approach was particularly appropriate for oil palm because it displays a very simple branching pattern. Indeed, the regular succession of phytomers in a single axis allowed us to study and describe the whole-plant architecture solely with allometric relationships based on leaf position on the stem or leaflet position on the rachis. However, allometric rules may not be sufficient to describe plants that exhibit a complex branching pattern (Lopez *et al.*, 2008; Costes *et al.*, 2008). Nonetheless, the principle of coupling mathematical functions describing the relationships between plant variables (allometric relationships or response curves to environmental variables) with a mixed-effect model,

as proposed here, remains relevant to any modelling approach that aims to describe genotypic behaviours.

Mixed-effect models have mainly been used in descriptive modelling approaches that deal with genotype–environment interactions (Smith *et al.*, 2005) or, more recently, to enhance the predictive capacity of agronomic and forest growth models (Hall and Bailey, 2001; Nothdurft *et al.*, 2006; Baey *et al.*, 2013; Le Bec *et al.*, 2015). Characterizing genetic behaviour through mixed-effects models is nevertheless possible when data are available for a large number of individuals. 3D plant reconstructions based on allometric relationships were thus preferred to digitizing because of the significant time saving.

In this study, the allometric relationships selected had the benefit of using model parameters linked to geometrical and topological properties. Consequently, parameters could be directly measured in future studies (using these already defined allometric relationships) to avoid having to make exhaustive measurements to estimate parameters from curve fitting. This trade-off between model accuracy and sampling effort is fully justified in cases of quantitative genetics and studies on plant architecture that require a large number of individuals to be phenotyped.

Another advantage of using an allometric approach is the possibility of characterizing contrasting profiles of ontogenetic and morphogenetic gradients between progenies (Table 5 and Fig. 2). Hence, not only could we compare plant architecture at a given time, we could also examine the temporal variability of architectural traits. This allowed us to detect features such as differential stem growth that were not identified at 47 MAP.

Progeny effects were, however, estimated trait by trait, without considering correlations between traits. Correlations between traits could be considered in further studies on the genetic determinism of plant architecture as a whole. From this perspective, methods discriminating classes of architecture from similarity indices between structures (Segura *et al.*, 2008b; Kawamura *et al.*, 2013) could be relevant. However, the classes mainly reflect the variation in the object sizes or number of components. In addition, in the absence of a genetic interaction between architectural components, such an approach may lead to a loss of important phenotypic information for breeding. Alternatively, a system of equations representing the trait dynamics or co-variations of trajectories, as proposed by Wu *et al.* (2011) could be used to study the genetic determinants of developmental processes of plant architecture.

Model simulation and accuracy of 3D reconstruction

In most modelling approaches, the general assessment of plant reconstruction relies on quantitative comparisons between means observed and simulated values for geometrical (Sonohat *et al.*, 2006) or topological descriptors (Costes *et al.*, 2008), or for more integrative features related to ecophysiological variables such as light interception (Casella and Sinoquet, 2003; Louarn *et al.*, 2008). In the current study, the quality of plant reconstruction was evaluated both in terms of mean and variance. The overall comparison showed the accuracy of the model reconstruction because it reproduced the main differences in architectural traits between progenies (Fig. 3 and Table 6). Nevertheless, model simulation accuracy tended to decrease when considering integrative variables simulated by a set of equations (e.g. leaf area for progeny DU).

Contrary to the variables simulated only through direct allometric relationships (L_{rac} or L), assessing the general quality of the 3D mock-ups generated by VPalm was not straightforward. An initial validation of the quality of the simulated 3D mock-ups was performed for the height of the rachis extremity but further investigations need to be carried out for the intra-canopy structure of plants, for instance, by using hemispherical photographs (Louarn *et al.*, 2008) or terrestrial LiDAR (Côté *et al.*, 2009).

Exploration of genotypic performance using 3D reconstruction

The architectural dissimilarities reported here between progenies, such as the number of leaflets per leaf, leaf curvature, and leaf shape, confer different spatial arrangements of leaves that likely influence light capture efficiency at the plant and leaf scale (Takenaka 1994; Takenaka *et al.*, 2001; Falster and Westoby 2003; Dauzat *et al.*, 2008). The combination of the reconstruction model proposed here with a radiative balance model would enable us to address the influence of plant architecture on light interception, considering not only inter-progeny variability but also intra-progeny variability. One originality of our approach was to integrate differences in ontogenetic gradients among plants, which increases the potential to generate plant architecture at different stages over a plant's lifetime. However, the conservation of constant allometric relationships related to plant components and morphologies during development is questionable and would need further investigation (Niklas, 1995).

The integration of inter-progeny and inter-individual variability in plant architecture is a first step towards investigating the impact of the architecture of oil palms on their performance in the field. Forthcoming work will be dedicated to sensitivity analyses coupling our structural model with a radiative balance model to identify the key architectural traits involved in light interception efficiency. Further research could also include a transition from the static model proposed here to a dynamic model with a trade-off between 3D architecture, light interception, photosynthesis, and plant growth over time (Vos *et al.*, 2010). Such prospects would involve coupling this architectural model with a dedicated plant growth simulator (Pallas *et al.*, 2013), allowing the simulation of retroactions between plant functioning, growth, and production.

Supplementary data

Supplementary data are available at *JXB* online.

Table S1. Progenies description

Table S2. Monitoring of data collection.

Table S3. Mean and SD of parameters used in allometric relationships for the five studied progenies.

Figure S1. Experimental design.

Figure S2. Leaflet length adjustment with different sample size from observed data.

Figure S3. Comparison of the number of leaflets on each side of rachis.

Figure S4. Length of leaflets along the rachis measured on each side of leaf.

Acknowledgements

This work was financially supported by the SMART Research Institute (SMARTRI, Smart Tbk.). The authors make special thanks to the SMARTRI staff for the logistic and technical assistance, in particular to Mr Doni A. Raharjo, Mr Fadhli, Mr Bengki, and Mr Muklis for their much valued help in the field data acquisition. We also greatly thank Nick Rowe for his help in English revision.

References

Baey C, Didier A, Sébastien L, Maupas F, Cournède PH. 2013. Modelling the inter-individual variability of organogenesis in sugar beet populations using a hierarchical segmented model. Ecological Modelling **263,** 56–63.

Barthélémy D, Caraglio Y. 2007. Plant architecture: a dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. Annals of Botany **99,** 375–407.

Ben Sadok I, Moutier N, Garcia G, Dosba F, Grati-Kamoun N, Rebai A, Khadari B, Costes E. 2013. Genetic determinism of the vegetative and reproductive traits in an f1 olive tree progeny. Tree Genetics and Genomes **9,** 205–221.

Billotte N, Jourjon MF, Marseillac N *et al*. 2010. QTL detection by multi-parent linkage mapping in oil palm (*Elaeis guineensis* Jacq.). Theoretical and Applied Genetics **120,** 1673–1687.

Bradshaw HJ, Stettler R. 1995. Molecular genetics of growth and development in *Populus*. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. Genetics **139,** 963–973.

Byrne M, Murrell JC, Kriedemann JVOP, Williams ER, Moran GF. 1997. Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. Theoretical and Applied Genetics **94,** 674–681.

Casadebaig P, Guilioni L, Lecoeur J, Christophe A, Champolivier L, Debaeke P. 2011. SUNFLO, a model to simulate genotype-specific performance of the sunflower crop in contrasting environments. Agricultural and Forest Meteorology **151,** 163–178.

Casella E, Sinoquet H. 2003. A method for describing the canopy architecture of coppice poplar with allometric relationships. Tree Physiology **23,** 1153–1170.

Chenu K, Chapman SC, Tardieu F, McLean G, Welcker C, Hammer GL. 2009. Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize: a "gene-to-phenotype" modeling approach. Genetics **183,** 1507–1523.

Corbeil RR, Searle SR. 1976. Restricted maximum likelihood (REML) estimation of variance components in the mixed model. Technometrics **18,** 31–38.

Corley R, Tinker P. 2003. The oil palm. Fourth edition. Oxford, Blackwell Science Ltd.

Costes E, Lauri PE, Laurens F, Moutier N, Belouin A, Delort F, Legave JM, Regnard JL. 2004. Morphological and architectural traits on fruit trees which could be relevant for genetic studies: a review. Acta Horticulturae **663,** 349–355.

Costes E, Smith C, Renton M, Guédon Y, Prusinkiewicz P, Godin C. 2008. MappleT: simulation of apple tree development using mixed stochastic and biomechanical models. Functional Plant Biology **35,** 936–950.

Côté JF, Widlowski JL, Fournier RA, Verstraete MM. 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial LiDAR. Remote Sensing and Environment **113,** 1067–1081.

Dauzat J, Eroy MN. 1997. Simulating light regime and intercrop yields in coconut based farming systems. European Journal of Agronomy **7,** 63–74.

Dauzat J, Clouvel P, Luquet D, Martin P. 2008. Using virtual plants to analyse the light-foraging efficiency of a low-density cotton crop. Annals of Botany **101,** 1153–1166.

Falster DS, Westoby M. 2003. Leaf size and angle vary widely across species: what consequences for light interception? New Phytologist **158,** 509–525.

Frary A, Fritz LA, Tanksley SD. 2004. A comparative study of the genetic bases of natural variation in tomato leaf, sepal, and petal morphology. Theoretical and Applied Genetics **109,** 523–533.

Gallais A. 1990. Théorie de la sélection en amélioration des plantes. Paris, Masson.

Godin C, Caraglio Y. 1998. A multiscale model of plant topological structures. Journal of Theoretical Biology **191,** 1–46.

Godin C, Costes E, Sinoquet H. 1999. A method for describing plant architecture which integrates topology and geometry. Annals of Botany **84,** 343–357.

Griffon S, de Coligny F. 2014. Amapstudio: an editing and simulation software suite for plants architecture modelling. Ecological Modelling **290,** 3–10.

Gu J, Yin X, Zhang C, Wang H, Struik PC. 2014. Linking ecophysiological modelling with quantitative genetics to support marker-assisted crop design for improved yields of rice (*Oryza sativa*) under drought stress. Annals of Botany **114,** 499–511.

Hackenberg J, Morhart C, Sheppard J, Spiecker H, Disney M. 2014. Highly accurate tree models derived from terrestrial laser scan data: a method description. Forests **5,** 1069–1105.

Hall DB, Bailey RL. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. Forest Science **47,** 311–321.

Halle F, Oldeman R. 1970. Essai sur l'architecture et la dynamique de croissance des arbres tropicaux. Monographies de botanique et de biologie végétale édition. Paris, Masson et Cie.

Hammer GL, van Oosterom E, McLean G, Chapman SC, Broad I, Harland P, Muchow RC. 2010. Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. Journal of Experimental Botany **61,** 2185–2202.

Hung HY, Browne C, Guill K *et al*. 2012. The relationship between parental genetic or phenotypic divergence and progeny variation in the maize nested association mapping population. Heredity **108,** 490–499.

Kawamura K, Oyant LHS, Foucher F, Thouroude T, Loustau S. 2013. Kernel methods for phenotyping complex plant architecture. Journal of Theoretical Biology **342,** 83–92.

Le Bec J, Courbaud B, Le Moguédec G, Pélissier R. 2015. Characterizing tropical tree species growth strategies: learning from inter-individual variability and scale invariance. PloS One **10**(3), e0117028.

Lecoeur J, Poiré-Lassus R, Christophe A, Pallas B, Casadebaig P, Debaeke P, Vear F, Guilioni L. 2011. Quantifying physiological determinants of genetic variation for yield potential in sunflower. SUNFLO: a model-based analysis. Functional Plant Biology, **38,** 246–259.

Letort V, Mahe P, Cournède PH, de Reffye P, Courtois B. 2008. Quantitative genetics and functional–structural plant growth models: simulation of quantitative trait loci detection for model parameters and application to potential yield optimization. Annals of Botany **101,** 1243–1254.

Li C, Li Y, Shi Y, Song Y, Zhang D, Buckler ES, Zhang Z, Wang T, Li Y. 2015. Genetic control of the leaf angle and leaf orientation value as revealed by ultra-high density maps in three connected maize populations. PloS One **10**(3), e0121624.

Lopez G, Favreau RR, Smith C, Costes E, Prusinkiewicz P, DeJong TM. 2008. Integrating simulation of architectural development and source–sink behaviour of peach trees by incorporating Markov chains and physiological organ function submodels into l-peach. Functional Plant Biology **35,** 761–771.

Louarn G, Lecoeur J, Lebon E. 2008. A three-dimensional statistical reconstruction model of grapevine *Vitis vinifera* simulating canopy structure variability within and between cultivar/training system pairs. Annals of Botany **101,** 1167–1184.

Morcillo F, Cros D, Billotte N *et al*. 2013. Improving palm oil quality through identification and mapping of the lipase gene causing oil deterioration. Nature Communications **4,** 2160.

Niinemets U. 2007. Photosynthesis and resource distribution through plant canopies. Plant, Cell and Environment **30,** 1052–1071.

Niklas KJ. 1995. Size-dependent allometry of tree height, diameter and trunk taper. Annals of Botany **75,** 217–227.

Nothdurft A, Kublin E, Lappi J. 2006. A non-linear hierarchical mixed model to describe tree height growth. European Journal of Forest Research **125,** 281–289.

Osorio LF, White TL, Huber DA. 2001. Age trends of heritabilities and genotype-by-environment interactions for growth traits and wood density from clonal trials of *Eucalyptus grandis* hill ex maiden. Silvae Genetica **50,** 30–37.

Pallas B, Clément-Vidal A, Rebolledo MC, Soulié JC, Luquet D. 2013. Using plant growth modeling to analyze C source-sink relations under drought: inter and intraspecific comparison. Frontiers in Plant Science **4,** 437.

Parveau CE, Chopard J, Dauzat J, Courbaud B, Auclair D. 2008. Modelling foliage characteristics in 3D tree crowns: influence on light interception and leaf irradiance. Trees **22,** 87–104.

Phattaralerphong J, Sinoquet H. 2005. A method for 3D reconstruction of tree crown volume from photographs: assessment with 3D-digitized plants. Tree Physiology **25,** 1229– 1249.

Pinheiro JC, Bates M. 2000. Mixed-effects models in S and S-PLUS. Springer Science & Business Media, New York.

Plomion C, Durel CE, O'Malley DM. 1996. Genetic dissection of height in maritime pine seedlings raised under accelerated growth conditions. Theoretical and Applied Genetics **93,** 849–858.

Prusinkiewicz P, Muendermann L, Karwowski R, Lane B. 2001. The use of positional information in the modeling of plants. In Proceedings of SIGGRAPH 2001, 12–17 August 2001, Los Angeles, California. The Association for Computing Machinery, California.

R Development Core Team 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Rance K, Mayes S, Price Z, Jack P, and Corley R. 2001. Quantitative trait loci for yield components in oil palm (*Elaeis guineensis* Jacq.). Theoretical and Applied Genetics **103,** 1302–1310.

Rees A. 1964. The apical organization and phyllotaxis of the oil palm. Annals of Botany **28,** 57–69.

Rey H, Dauzat J, Chenu K, Barczi JF, Dosio G, Guillermo A, Lecoeur J. 2008. Using a 3-D virtual sunflower to simulate light capture at organ, plant and plot levels: Contribution of organ interception, impact of heliotropism and analysis of genotypic differences. Annals of Botany **101,** 1139–1151.

Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F. 2003. Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. Plant Physiology **131,** 664–675.

Sakamoto T, Matsuoka M. 2004. Generating high-yielding varieties by genetic manipulation of plant architecture. Current Opinion in Biotechnology **15,** 144–147.

Sarlikioti V, de Visser PHB, Buck-Sorlin GH, Marcelis LFM. 2011. How plant architecture affects light absorption and photosynthesis in tomato: towards an ideotype for plant architecture using a functional–structural plant model. Annals of Botany **108,** 1065– 1073.

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

Segura V, Cilas C, Costes E. 2008a. Dissecting apple tree architecture into genetic, ontogenetic and environmental effects: mixed linear modelling of repeated spatial and temporal measures. New Phytologist **178,** 302–314.

Segura V, Ouangraoua A, Ferraro P, Costes E. 2008b. Comparison of tree architecture using a tree edit distance: application to 2-year-old apple hybrids. Euphytica **161,** 155–164.

Serrano-Cartagena J, Robles P, Ponce MR, Micol JL. 1999. Genetic analysis of leaf form mutants from the *Arabidopsis* information service collection. Molecular and General Genetics **261,** 725–739.

Sinoquet H, Rivet P, Godin C. 1997. Assessment of the three-dimensional architecture of walnut trees using digitising. Silva Fennica **31,** 265–273.

Smith A, Cullis B, Thompson R. 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. Journal of Agricultural Science, **143,** 449– 462.

Sonohat G, Sinoquet H, Kulandaivelu V, Combes D, Lescourret F. 2006. Three-dimensional reconstruction of partially 3d-digitized peach tree canopies. Tree Physiology **26,** 337–351.

Takenaka A. 1994. Effects of leaf blade narrowness and petiole length on the light capture efficiency of a shoot. Ecological Research **9,** 109–114.

Takenaka A, Takahashi K, Kohyamas T. 2001. Optimal leaf display and biomass partitioning for efficient light capture in an understorey palm, *Licuala arbuscula*. Functional Ecology **15,** 660–668.

Talliez B, Ballo Koffi C. 1992. A method for measuring oil palm leaf area. Oléagineux **47,** 537–545.

Tardieu F. 2003. Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. Trends in Plant Science **81,** 9–14.

Tisné S, Denis M, Cros D, Pomiès V, Riou V, Syahputra I, Omoré A, Durand-Gasselin T, Bouvet JM, Cochard B. 2015. Mixed model approach for IDB-based QTL mapping in a complex oil palm pedigree. BMC Genomics **16,** 798.

Vos J, Evers JB, Buck-Sorlin GH, Andrieu B, Chelle M, de Visser PHB. 2010. Functional–structural plant modelling: a new versatile tool in crop science. Journal of Experimental Botany **61,** 2101–2115.

Wang Y, Li J. 2005. The plant architecture of rice *Oryza sativa*. Plant Molecular Biology **59,** 75–84.

Welham CVJ, Turkington R, Sayre C. 2002. Morphological plasticity of white clover (*Trifolium repens l*.) in response to spatial and temporal resource heterogeneity. Oecologia **130,** 231–238.

Willaume M, Lauri PE, Sinoquet H. 2004. Light interception in apple trees influenced by canopy architecture manipulation. Trees **18,** 705–713.

Wu R, Stettler R. 1998. Quantitative genetics of growth and development in *Populus*. III. Phenotypic plasticity of crown structure and function. Heredity **81,** 299–310.

Wu R, Cao J, Huang Z, Wang Z, Gai J, Vallejos E. 2011. Systems mapping: how to improve the genetic mapping of complex traits through design principles of biological systems. BMC System Bilology **5,** 84.

Yang XC, Hwa CM. 2008. Genetic modification of plant architecture and variety improvement in rice. Heredity **101,** 396– 404.

Yin X, Kropff MJ, Stam P. 1999. The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. Heredity **82,** 415–421.