Multiscale digital *Arabidopsis* predicts individual organ and whole-organism growth

Yin Hoon Chew[a], Bénédicte Wenden[b], Anna Flis[c], Virginie Mengin[c], Jasper Taylor[d], Christopher L. Davey[e], Christopher Tindal[a], Howard Thomas[e], Helen J. Ougham[e], Philippe de Reffye[f], Mark Stitt[c], Mathew Williams[g], Robert Muetzelfeldt[d], Karen J. Halliday[a], and Andrew J. Millar[a,1]

[a]SynthSys and School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JD, United Kingdom; [b]Institut National de la Recherche Agronomique and Université Bordeaux, Unité Mixte de Recherche 1332 de Biologie du Fruit et Pathologie, F-33140 Villenave d'Ornon, France; [c]Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany; [d]Simulistics Ltd., Loanhead EH20 9PA, United Kingdom; [e]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 2FG, United Kingdom; [f]Cirad-Amis, Unité Mixte de Recherche, Association pour le Maintien d'une Agriculture Paysanne, F-34398 Montpellier Cedex 5, France; and [g]School of GeoSciences, University of Edinburgh, Edinburgh EH9 3JN, United Kingdom

Understanding how dynamic molecular networks affect whole-organism physiology, analogous to mapping genotype to phenotype, remains a key challenge in biology. Quantitative models that represent processes at multiple scales and link understanding from several research domains can help to tackle this problem. Such integrated models are more common in crop science and ecophysiology than in the research communities that elucidate molecular networks. Several laboratories have modeled particular aspects of growth in *Arabidopsis thaliana*, but it was unclear whether these existing models could productively be combined. We test this approach by constructing a multiscale model of *Arabidopsis* rosette growth. Four existing models were integrated with minimal parameter modification (leaf water content and one flowering parameter used measured data). The resulting framework model links genetic regulation and biochemical dynamics to events at the organ and whole-plant levels, helping to understand the combined effects of endogenous and environmental regulators on *Arabidopsis* growth. The framework model was validated and tested with metabolic, physiological, and biomass data from two laboratories, for five photoperiods, three accessions, and a transgenic line, highlighting the plasticity of plant growth strategies. The model was extended to include stochastic development. Model simulations gave insight into the developmental control of leaf production and provided a quantitative explanation for the pleiotropic developmental phenotype caused by overexpression of miR156, which was an open question. Modular, multiscale models, assembling knowledge from systems biology to ecophysiology, will help to understand and to engineer plant behavior from the genome to the field.

plant growth model | digital organism | crop modeling | ecology

Our goal is to understand the physiological effects of metabolic and regulatory networks that are now being elucidated at the molecular level. Such networks control the traits, such as drought resistance, that are important both in agriculture and in ecosystem responses to climate change. Molecular genetic approaches, often in model organisms, have uncovered the operating principles and mechanisms for a growing number of physiologically relevant cases. For example, environmental factors such as CO_2 concentration, temperature, and light flux can display coordinated diurnal and seasonal fluctuations (1, 2). For annual plants like the laboratory model species *Arabidopsis thaliana*, matching the timing of flowering to the favorable season, and thus the associated environment, increases reproductive success (3). This synchronization is achieved by changing gene expression and protein abundance at the molecular level. *Arabidopsis FLOWERING LOCUS T* (*FT*) is an example of such an integrator gene that induces flowering in response to environmental signals (4). *FT* is highly expressed in long (summer) days due to a combination of light and circadian clock regulation (5–8). Such responses collectively enable individual plants to survive in variable conditions. Plants adapt their resource allocation

processes to the environmental conditions to optimize growth and biomass accumulation (9). Plants also adjust their architecture to compete for light and nutrient resources (10, 11). Given the multiplicity and interactions of such responses, however, it can be difficult to determine how much a particular molecular change contributes to the effect on the whole plant. To understand physiology and to facilitate predictive biology from the molecular level, there is a well-recognized need for quantitative models that cross biological scales and link understanding from several scientific domains (12–14).

There already exist mathematical models describing various plant processes and their interactions with the environment (13). These models include varying levels of mechanistic detail, starting from simple statistical relationships, and they usually comprise two scales at most (15). Broader, molecular-based models are well advanced in only a few domains of plant science, such as photosynthesis research (16, 17) and root development (18). If the models can be assembled and updated in a modular fashion, then larger, multiscale models might be developed in

Significance

Plants respond to environmental change by triggering biochemical and developmental networks across multiple scales. Multiscale models that link genetic input to the whole-plant scale and beyond might therefore improve biological understanding and yield prediction. We report a modular approach to build such models, validated by a framework model of *Arabidopsis thaliana* comprising four existing mathematical models. Our model brings together gene dynamics, carbon partitioning, organ growth, shoot architecture, and development in response to environmental signals. It predicted the biomass of each leaf in independent data, demonstrated flexible control of photosynthesis across photoperiods, and predicted the pleiotropic phenotype of a developmentally misregulated transgenic line. Systems biology, crop science, and ecology might thus be linked productively in a community-based approach to modeling.

SYSTEMS BIOLOGY

a distributed, community approach. Experts in each domain could model a particular aspect of biology in detail and reuse the previously assembled models to represent other aspects, where coarse granularity might be sufficient. This approach follows that developed in other contexts, for example in other areas of biology (19–21) and in the Earth system modeling community. In that case, submodels for the atmosphere, ocean, ice sheets, and the land surface are coupled; their interactions and dynamics are then evaluated against independent observations (22).

Our work was motivated by the multiple challenges of coupling growth models to the molecular level. Computational problems are expected, due to logical or technical incompatibilities among models (23), but historical factors are also common, including the inaccessibility of executable forms or reference data for some published models. These challenges have been overcome to varying extents in modeling frameworks for crop science (24, 25) and for animal physiology (26) but rarely in plant molecular genetics (27) and molecular systems biology (28). In addition, one key biological issue is whether the data that were used in the construction (or calibration) of the constituent models are quantitatively compatible, so the models can be combined with minimal modification. Recalibrating the parameter values of models at many different scales is potentially laborious, requiring coordinated data acquisition by researchers from multiple disciplines. The quantitative compatibility of the constituent models is therefore a key question.

To develop a quantitative model that links multiple, interacting processes from metabolism to development, we implemented the modular modeling approach for *Arabidopsis thaliana*. The *Arabidopsis* research community has abundant molecular and physiological data, but the quantitative variation observed among laboratories is significant (29). We identified four independently developed models that characterize different aspects of plant biology and then combined and extended them to form a framework model (FM) for vegetative growth in *Arabidopsis* (Fig. 1). Key aspects of the models showed very good quantitative agreement. The FM was validated against metabolic, physiological, and bio-

mass data for multiple genetic backgrounds and tested for several growth conditions. Numerical simulations using FM enabled us to understand the physiological relevance of an observed developmental process. Finally, we compared additional data with model predictions to quantitatively explain a developmental phenotype, supporting one of two proposed mechanisms.

Model Description

Components of the FM. The circadian clock in *Arabidopsis* is one example of a pervasive, molecular regulator where we have substantial understanding of its molecular mechanisms (30). Photosynthetic carbon metabolism, vegetative growth, and flowering time are among the many biological processes controlled by the clock. A multiscale model will be required to understand how the circadian timing (or any other pervasive control) of each of these interacting processes contributes quantitatively to the growth of the whole plant under varying environmental conditions. We therefore integrated four models, each of which originated in a different laboratory:

i) A carbon dynamic model (CDM) that considers the subcellular processes of photosynthesis and sugar-starch partitioning, as well as carbon (C) allocation between the leaf area and the roots (31). It is assumed that a fixed proportion (12.5%) of C assimilated through photosynthesis is partitioned into starch, with the possibility to accumulate more starch if the remainder of the photosynthate (in the form of sugar) is not used for growth and respiration. At night, starch is degraded at a linear rate, adjusted to the night length, to sustain growth until dawn (32, 33). The rate of starch degradation is set such that 84% of that accumulated in the light period is degraded by dawn. The CDM was a discrete-time model with a 6-s time step, constructed using data of Columbia (Col) WT plants grown under 8-h light (L):16-h dark (D) conditions.

ii) A functional-structural plant model (FSPM) that describes individual organ growth and how each organ (leaf) contributes to the above-ground structure for light interception

Fig. 1. Overview of the FM. The FM takes environmental data as input (black) to four existing *Arabidopsis* models (blue shadowed boxes), which are (*A*) a carbon dynamic model (CDM) that describes carbon assimilation and resource partitioning (31); (*B*) a functional-structural plant model (FSPM) of individual organ growth that determines the rosette structure (green) and the area for light interception (34); (*C*) a photothermal model (PTM) that predicts flowering time (1); and (*D*) a photoperiodism model (PPM), which is a gene dynamic model of the circadian clock and the photoperiod response pathway (6). On integration, several original components were discarded (gray), whereas new connections were created (red).

(34). Each of these factors is represented by effective mathematical functions in the model, without mechanistic detail, but in a very concise form that was sufficient to represent *Arabidopsis* shoot growth and structure (34). It was parameterized using data of Col WT plants grown under 12-h L:12-h D conditions. Only a subset of the large original model (34) was applicable to our study. The relevant subset of parameter values and developmental structures was rewritten into a conventional, dynamic form that was compatible with the other submodels, as a discrete-time model with an hourly time step.

iii) A photothermal model (PTM) that predicts the timing of flowering, based on temperature integrated over time (thermal time) (1). In *Arabidopsis*, flowering time is governed by the photoperiod pathway that enables plants to sense daylength (35), the vernalization pathway that promotes flowering in the spring after a long chilling period over winter (36), and warm ambient temperature (37). Each of these factors is represented by effective mathematical functions in the model, without mechanistic detail, and it was parameterized using field data of various genotypes in the Col and Landsberg *erecta* (Ler) backgrounds (38). The model was formulated as a discrete-time model with an hourly time step.

iv) A photoperiodism model (PPM), which is a gene dynamic model of the circadian clock (39) and the photoperiod pathway (6). This model was a conventional ordinary differential equation (ODE) model, usually solved with an adaptive time step of minutes or less. The model was parameterized using data from Col and Ler WT plants grown under 16-h L:8-h D and 8-h L:16-h D conditions.

Model Integration Process. To link the four models, we first identified the essential variable(s) from each that could act as the connection points. New links and scaling factors were introduced, whereas redundant model components were replaced (Fig. 1 and *SI Appendix*). Unit conversions were required for compatibility, and two parameter values (*viii* and *ix* below) were measured from our experiments. The 124 other parameter values were taken from the original models. A summary of the integration process is as follows:

i) The model's time step was standardized to 1 h for all except for the PPM, which is solved at shorter, variable time steps. Our model therefore takes hourly meteorological data as input, similar to many crop and ecosystem models (Fig. 2 *A–C*) and thereby resolves diel behavior.

ii) The simple root-to-shoot carbon allocation ratio in the CDM was directly replaced with the dynamic pattern of demand from individual organs, calculated by the FSPM (Fig. 2 *E* and *F*).

iii) To facilitate the replacement step *ii* above, biomass measures considering only carbon in the CDM were converted to total dry mass using published leaf and root carbon content data (40–42), because not all biomass is carbon.

iv) The simple "big leaf" rosette area for photosynthesis in the CDM was directly replaced by the projected area of the rosette structure from the FSPM.

v) The sugar supply calculated by the CDM, from fine-grained processes such as photosynthesis, respiration, and sugar-starch partitioning, was directly provided to the FSPM as the sugar supply for growth. This variable replaced the empirical light use efficiency (LUE) component, which was previously estimated from experimental data through model inversion (34).

vi) Seedling emergence (43) and flowering time were represented explicitly, in terms of thermal times to emergence and flowering from the PTM. These variables were not previously considered in the CDM and the FSPM.

vii) The simple, piecewise-linear function for photoperiod response in the PTM was replaced by the continuous flowering function driven by the integrated expression of the flowering time gene *FT* in the PPM (6).

viii) The modified photothermal unit (MPTU) threshold in the PTM (threshold, Fig. 2*H*) was determined using the time of flowering measured in our experiments.

ix) Water content was measured from our experiments to facilitate simulation of fresh biomass, because this is a simpler and more widely available measurement than the dry mass used in both the CDM and the FSPM.

All of the modeling work and analysis were conducted in *MATLAB* (Mathworks) (*SI Appendix*). The Plant Systems Modeling (PlaSMo) online model repository (www.plasmo.ed.ac.uk) was developed as a shared portal to disseminate relevant models from systems biology and eco-physiology. The component models and the FM are publicly accessible from PlaSMo in *MATLAB* and Simile formats with the following identifiers: CDM, PLM_2; FSPM, PLM_75; PTM, PLM_73; PPM, PLM_74; and FM, PLM_76. Simile provides a visual modeling environment with a graphical user interface, plotting tools, and an animated display of simulated plant growth (Movie S1) (44).

Results

Model Validation and Testing. We first examined the performance of the FM in representing the growth of Col, which was the common *Arabidopsis* accession used to create the original models. As the model's flowering time was calibrated to the data, we focus here on vegetative growth. WT Col plants were grown in 12-h L:12-h D cycles close to 22 °C, because these conditions most closely matched the conditions used for the original models, except for the CDM that was tested using an 8-h photoperiod (31). Highly discriminating data sets were collected for the biomass of the total shoot and individual leaves and for the area of individual leaves at multiple time points after sowing. Using the original parameter values for each submodel, the FM overestimated growth (*SI Appendix*, Fig. S1). However, the literature shows that *Arabidopsis* grown in an 8-h photoperiod have altered photosynthetic physiology compared with our reference 12-h photoperiod. Specifically, the ratio of maximum electron transport to the maximum rate of carboxylation ($J_{max}:V_{cmax}$) decreases as photoperiod increases (45–47) (*SI Appendix*, Table S7). The CDM's original value for $J_{max}:V_{cmax}$ has only been tested in an 8-h photoperiod (31). Substituting the value measured in a 12-h photoperiod was sufficient for the FM to fit the Col biomass data (Fig. 2 *I–K*). The R^2 between measured and modeled values of fresh biomass, dry biomass, and area of the rosette were 0.98, 0.99, and 0.98, respectively, with normalized root mean square error (nRMSE) less than 10% (*SI Appendix*, Table S8). The median values of R^2 and nRMSE for all of the data, including individual leaf predictions, were 0.91 and 24.7%, respectively. The dynamic operation of the model in Simile is illustrated in Movie S1.

The FM was next tested by comparison with growth data from other *Arabidopsis* accessions: Ler and Feira-0 (Fei). Accession-specific parameters were measured for the seedling emergence and flowering times, as described above for Col, and for the changing rate of leaf production. Fei was expected to show a higher leaf appearance rate (48), and indeed it showed a larger leaf number compared with Ler at the same time points (*SI Appendix*, Fig. S2A). However, leaf appearance rate in Fei matched the Col rate when plotted against thermal time after seedling emergence (*SI Appendix*, Fig. S2B). We infer that the principal difference of Fei from Col is actually in the time to emergence, as Fei emerged at half the thermal time for Col (*SI Appendix*, Table S6). With only these changes, the model's match to data of Ler and Fei plants was as good as for Col (Fig. 3 *A, B, D*, and *E*), with median R^2 (and nRMSE) of 0.94 (16%) and 0.95 (17.3%), re-
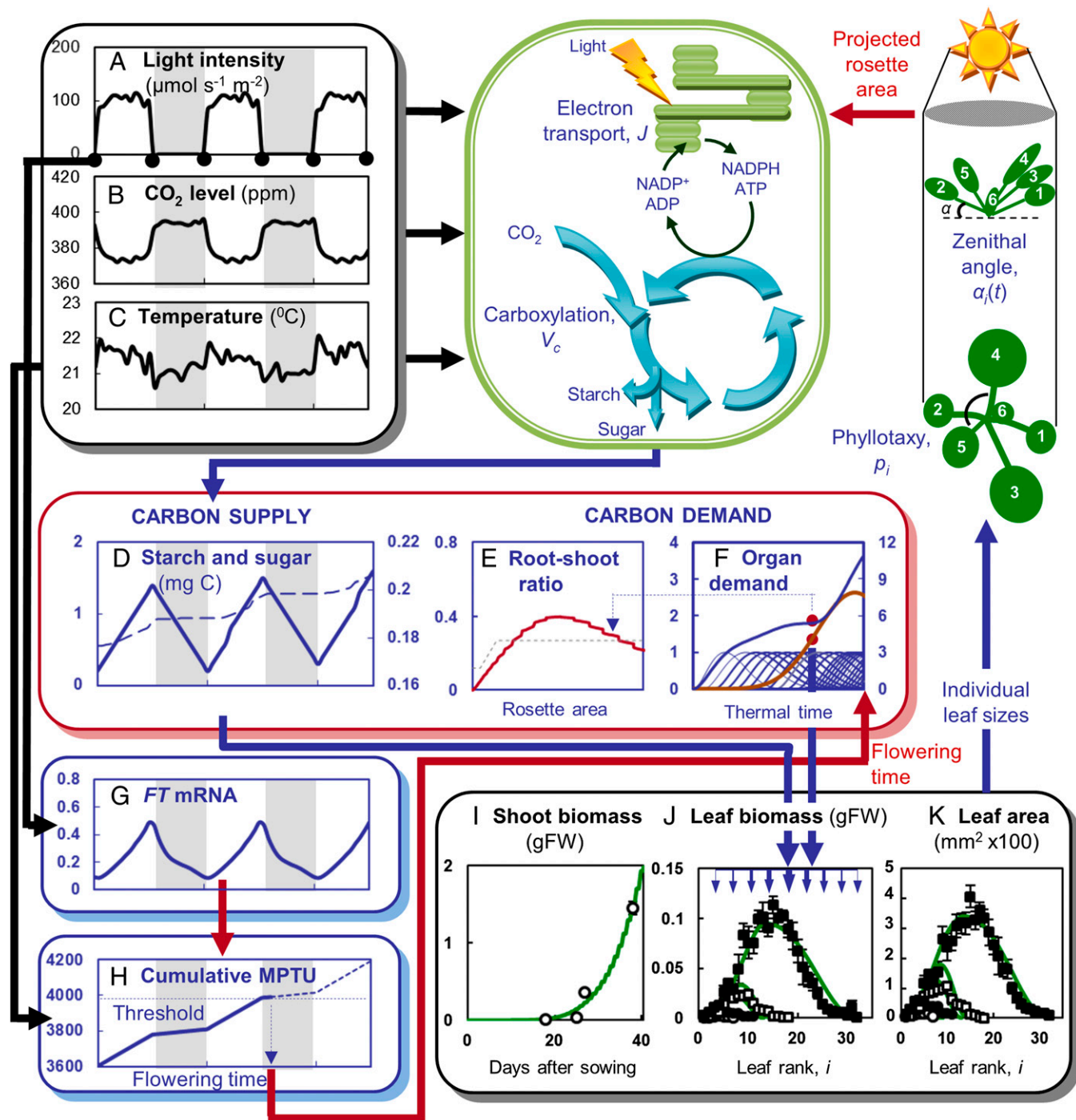
Fig. 2. The FM's workflow predicts whole-plant and individual organ growth data. Input data required are hourly light intensity (*A*), CO_2 level (*B*), and temperature (*C*), illustrated for simulated three 12-h light (open):12-h dark (shaded area) cycles. Carbon supply (*D*) is used as sugar (dashed line) or stored as starch (solid line). Carbon is allocated at each hourly time step according to the organ demand (*E* and *F*). The simulated pattern of demand from individual leaves (*F*, thin blue lines, left axis) is used to calculate the sum of demand (dots) from leaves (thick blue line, right axis) and roots (brown line, left axis). The root-to-shoot allocation ratio (*E*), calculated dynamically from the FSPM (red line), is similar to the piecewise-linear function prescribed in the CDM (31) (gray dashed line), which it replaces. Times of dawn and dusk (dots in *A*) affect the level of flowering gene *FT* mRNA (*G*) simulated by the PPM, which in turn controls the accumulation of modified photothermal units (MPTU; *H*). Once the accumulated photothermal units reach the threshold for flowering (dashed lines), rosette growth is terminated in the FSPM (red arrow). Model outputs include biomass of the shoot (*I*) and individual leaves (*J*). Simulations for the Col WT (green lines) closely match experimental data for (*I*) total shoot biomass, (*J*) leaf biomass, and (*K*) leaf area at 18 (○), 25 (●), 27 (□), and 38 (■) d after sowing. Leaves are ranked according to the order of appearance. The integrated model uses simulated sizes of individual leaves (*K*) to calculate the projected rosette area for photosynthesis (red arrow), considering the spiral leaf arrangement (phyllotaxy) and upward (zenithal) angle. Experimental conditions: ~21.3 °C; 12:12-h light/dark cycle; light intensity, 110 $\mu mol \cdot m^{-2} \cdot s^{-1}$; mean daytime CO_2 level, 375 ppm. The error bars show the SEs of five plants. The color code links to the model components in Fig. 1.
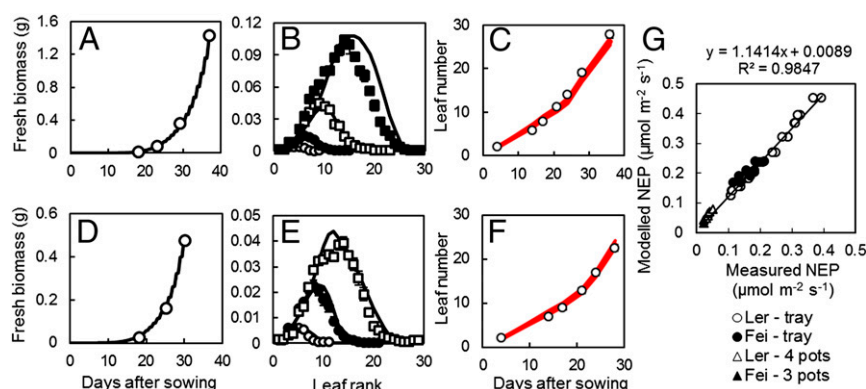
Fig. 3. The FM predicts plant growth and gas exchange data for different accessions. Model simulations (solid lines) and experimental data (symbols) of total shoot biomass, individual leaf biomass, and leaf number for Ler (*A–C*) and Fei (*D–F*) are shown. Time points of measurement in *B* are 18 (○), 23 (●), 29 (□), and 37 (■) d after sowing (DAS). Time points of measurement in *E* are 18 (○), 25 (●), and 30 (□) DAS. The thickness of the red lines in *C* and *F* represents a region with 1 SD above and below the mean values from the stochastic simulations of leaf number for 2,400 model runs. The plot of modeled and measured NEP of CO_2 is illustrated in *G*. NEP was measured for plants grown either as a small population on a tray or in individual pots. Experimental conditions: 22 °C; 12:12-h light/dark cycle; light intensity = 130 $\mu mol \cdot m^{-2} \cdot s^{-1}$; average daytime CO_2 concentration = 375 ppm. Error bars in *A, B, D,* and *E* show the SEs of $n = 10$ plants for total shoot biomass and $n = 5$ plants for individual leaf biomass. Error bars in *C* and *F* (smaller than the symbols) represent the SD of 24 plants.

spectively (*SI Appendix*, Table S8). The measured water content was found to be 92%, 91%, and 88% for Col, Ler, and Fei, respectively, which were used in the simulations. We also tested the use of a standardized water content of 91%. This standardization caused slight overestimation of fresh biomass for Fei, but less significant effects for Col (*SI Appendix*, Fig. S3).

We additionally tested the applicability of our multiscale FM to ecosystem studies by comparing model simulations to measured trace gas exchange data (*SI Appendix*). We measured net ecosystem production (NEP) of CO_2 for a population of *Arabidopsis* plants in an experimental setup typically used for ecological research (49, 50) (*SI Appendix*, Fig. S4). The model accurately predicted measured gas exchange from 26 d after sowing until flowering time ($R^2 = 0.98$; Fig. 3*G*). Our results therefore suggested that the robustness of photosynthetic physiology contributed to the compatibility of the independently developed models.

To determine which processes most affected the simulated biomass and flowering time, we conducted a sensitivity analysis, perturbing each parameter in turn by 5%. Perturbations that increased (or decreased) flowering time always increased (decreased) biomass at flowering (*SI Appendix*, Figs. S8 and S9 and Table S11), because of the longer (shorter) duration of biomass accumulation in the rosette. Flowering time was controlled by parameters of the PPM, by the overall flowering threshold and by the baseline *FLC* repression, as expected in our nonvernalizing conditions (*SI Appendix*). Vegetative growth was also assessed at a fixed time, 36.5 d after sowing. Of the 12 parameters that most affected fresh biomass at this time point, 2 parameters directly controlled the water and carbon content of the modeled biomass. Each of these parameters represents a complex physiological process. Eight parameters represented photosynthetic processes and two were related to leaf structure (specific leaf area), underlining the importance of these traits in predicting growth rate.

Model Extension: Photosynthetic Adaptation and Flexible Starch Metabolism Explain the Photoperiodic Regulation of *Arabidopsis* Growth Development. *Arabidopsis* can adapt to a wide range of photoperiods by adjusting photosynthetic capacity (45–47) and carbon allocation (9, 32). In particular, starch accumulation is faster and starch degradation is slower in short photoperiods. A large, independent study (51) allowed us to test the model predictions in 4-, 6-, 8-, 12-, and 18-h photoperiods (Fig. 4). Changing photoperiod is known to alter biochemical parameters of photosynthesis in the plant that were fixed in the CDM. We therefore

substituted the literature values for the $J_{max}:V_{cmax}$ ratio measured in the appropriate photoperiod conditions, assuming upper and lower limits (*SI Appendix*, Table S7). The simulations also replicated the relevant environmental conditions (*SI Appendix*, Section 3.11) (51).

Carbon assimilation and respiration rates were slightly underestimated (10.7% and 6% lower in the 12-h photoperiod, for example) on the simulated day corresponding to the day of measurement (Fig. 4*A*). The resulting net carbon fixation allowed the model to reproduce the full amount of starch accumulation by the end of the 12-h photoperiod (Fig. 4*B*), but starch levels were underestimated (by 10–26%) in shorter photoperiods (Fig. 4*B*). The model closely matched the starch levels remaining at the end of the night (Fig. 4*B*). However, in short photoperiods, the lower amount of starch accumulation in the light meant that the amount of starch mobilized per night was underestimated in the model. Additionally, part of the mobilized starch was used to maintain a higher sucrose level than observed in the data (*SI Appendix*, Fig. S5A), where sucrose levels decreased progressively as the photoperiod was shortened. These two factors resulted in lower growth per night in the model than in the data (Fig. 4*C*). The model more closely matched the observed growth increment in the 12-h photoperiod (Fig. 4*C*), where the simulated starch and sucrose levels matched observations (Fig. 4*B* and *SI Appendix*, Fig. S5A). Integrated over the life of the plant, the lower growth at night led the FM to underestimate total rosette biomass for short photoperiods (Fig. 4*D*). This result indicates that further parameters in addition to $J_{max}:V_{cmax}$ are important for modeling growth, especially in the extreme 4-h photoperiod. In contrast, the FM accurately predicted the biomass in the 12-h photoperiod protocol, to within the experimental error (Fig. 4*D*). These results confirm that the FM can closely match the data from independent laboratories in the reference conditions, but the simple CDM did not fully account for the changing starch and sugar dynamics in short photoperiods.

Between a 4- and 12-h photoperiod, biomass increased strongly, and the relative growth rate (RGR, milligrams fresh weight produced per day per unit existing fresh weight) increased almost linearly with light fluence (51). This response would be expected if the conversion efficiency of carbon into biomass was constant. This linear relation between daily light fluence and growth was lost in long photoperiods. Whereas light fluence increased by 50% between the 12- and 18-h photoperiod, RGR increased by only 18% (51). Observed changes in the 18-h photoperiod included higher starch levels at dawn and a reduction in specific leaf area
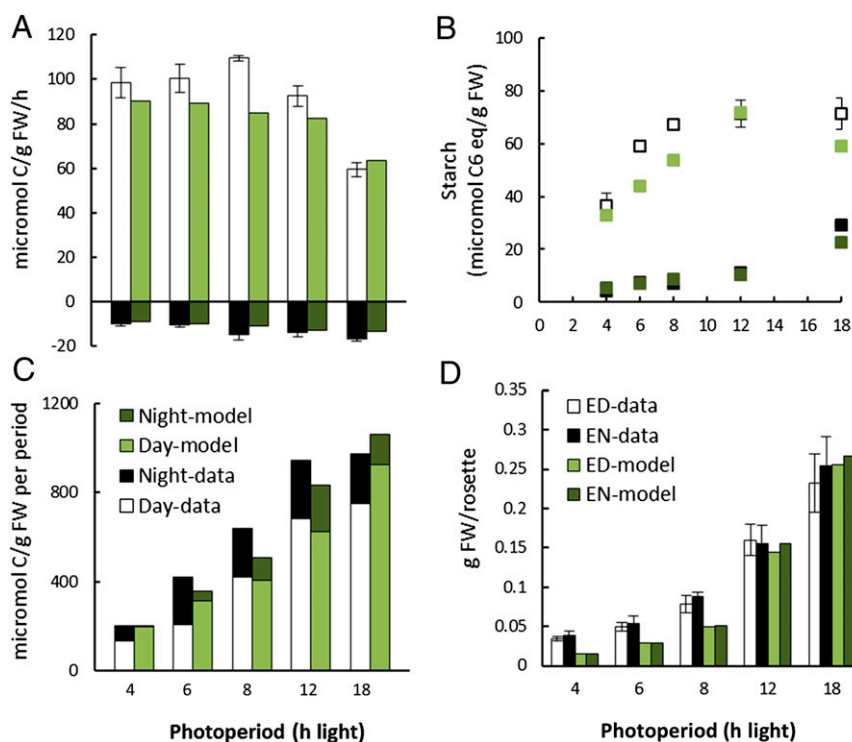
Fig. 4. Testing the FM under different photoperiods. Experimental data (black and white) (51) are compared with model simulations (light and dark green) in the photoperiods indicated, for (*A*) carbon assimilation and respiration rates; (*B*) starch levels; (*C*) amount of growth per day or night period; and (*D*) rosette fresh weight at the end of day (ED; white and light green) and end of night (EN; black and dark green). Error bars show the SD of five plants.

(i.e., increased leaf thickness) (51). Both of these are expected to reduce growth rates; incomplete starch mobilization will sequester carbon from growth, whereas increased leaf thickness will mean that less leaf area is generated per unit fixed carbon, which will decrease future light absorption and photosynthesis. Including the slower night-time starch breakdown (to 60% of initial starch rather than 84%) and measured 15–25% increase in leaf thickness in the FM, in addition to substituting J_{max}:V_{cmax} with the published value for the 14-h photoperiod, reproduced the observed biomass (Fig. 4*D*). This result was also recapitulated by extrapolating J_{max}:V_{cmax} below the published value for 14-h photoperiods and reducing starch breakdown, but without considering the increase in leaf thickness (*SI Appendix*, Fig. S5*B*). Thus, these three factors are sufficient to account quantitatively for the altered growth rate under long photoperiods, although the balance among them remains to be determined experimentally.

Model-Guided Understanding: Stochasticity and Tradeoffs in Development. To explore the model's potential, we extended the FM to include stochastic development at the organ (leaf) level, adopting a probabilistic organ initiation concept used for describing nonsymmetrical branching in plant architecture (52). Leaves are considered to appear at a regular interval (or growth cycle) with a simple, binomial probability that was estimated at 0.97 from our experimental data on Ler and Fei (*SI Appendix*, Section 3.12 and Fig. S6). Thus, leaves appear on most growth cycles but not all, reflecting variation in the processes of organ initiation and expansion. This stochastic model explained the variance of leaf number in our samples at every time point (P > 0.05; Fig. 3 *C* and *F* and *SI Appendix*, Table S9) while accounting for 11.3% (Ler) and 12.7% (Fei) of the variance in biomass measured at flowering time. The SDs in the timing of leaf appearance (phyllochron) from our simulations [in degree-days: 2.35 (Ler) and 1.86 (Fei)] were, however, lower than the SD in leaf initiation (plastochron) reported in Col (12.72; *Discussion*) (53).

Besides interplant variation, both leaf initiation and leaf appearance rates increase with plant age in *Arabidopsis* (48, 53). The model reproduces this using a piecewise-linear rate, with a phase transition point at 355 degree-days, around half the vegetative period (Fig. 3 *C* and *F* and *SI Appendix*, Fig. S2). We explored the significance of this developmental timing, by simulating earlier or later transition points (*SI Appendix*, Fig. S7*A*). To distinguish the effect of the varying rate, we included controls that generated the same final leaf number at a constant rate. Model simulations with a transition point earlier than the reference, hence a longer interval of rapid leaf production, generated biomass as low as 46.4% of the reference value (Fig. 5*A*). Most leaves were small: median and third-quartile leaf areas fell to 32.6% and 33.5% of the reference value (Fig. 5*C*). The high leaf number and smaller size resulted in self-shading that reduced biomass. The varying leaf production rate generally resulted in a larger fraction of functional (photosynthesising) leaves at flowering time than in the controls (Fig. 5*B*) and, for transition points at 100–400 degree-days, in a greater proportion of large leaves (third quartile area above control; Fig. 5*C*) that partly escaped shading, resulting in higher biomass than in the controls (Fig. 5*A*). Simulations with a later transition point, hence a longer interval of slow leaf production, increased biomass (6% increase from transitions at 500–650 degree-days; Fig. 5*A*). The associated controls increased biomass up to 10.9%. The plant's observed behavior, represented by the reference transition point, seemed suboptimal. However, the later transition points reduced the percentage of functional leaves at flowering from 88.9% to 81.8% (Fig. 5*B*). Median leaf area increased by 21.3% with a transition point at 600 degree-days, similar to total biomass, but a few leaves grew very large (third quartile area increased by 73% of the reference, but was only 78.6% of the maximum area; Fig. 5*C*). Thus, the higher biomass of these simulated plants depended on a smaller number of larger leaves. In contrast, near the reference transition point (300–400 degree-days), the third quartile leaf area was up to 93.8% of the
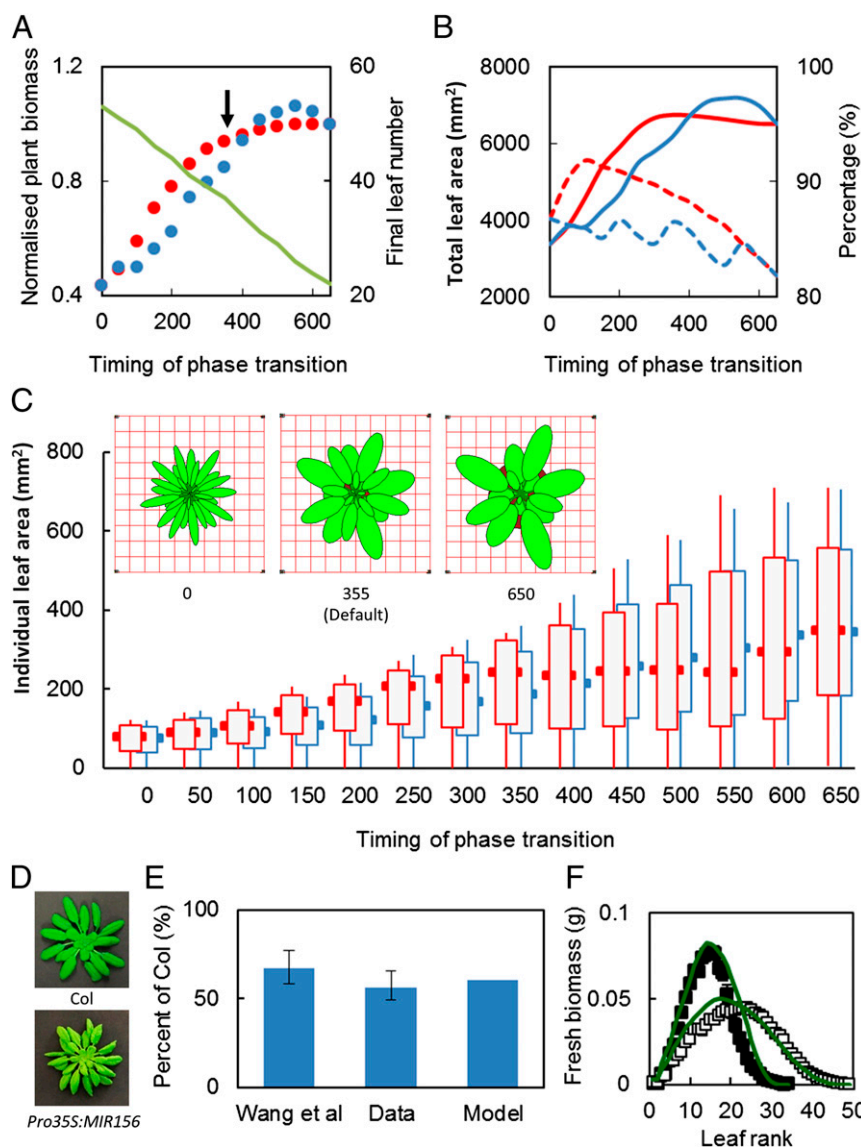
Fig. 5. Leaf production rate balances biomass and leaf area for photosynthesis. Simulation results with time-varying leaf production rates (red) and the associated controls with constant rates (blue) are shown for (*A*) plant biomass (symbols, left axis) and final leaf number at flowering (green line, right axis); biomass is normalized to the maximum value achievable with the varying leaf production rate, which corresponds to a phase transition to the higher, mature rate at 550 degree-days after sowing. (*B*) Total functional (photosynthesising) leaf area (solid lines, left axis) and percentage of functional leaves (dashed lines, right axis). (*C*) Boxplots showing the size distribution of functional leaves. Results shown include the minimum and maximum values (whiskers), the first and third quartiles (boxes), and the median values (outer markers). *Inset* in *C* illustrates the images of simulated rosettes from the Simile animation tool, for three transition points as indicated under each image. The arrow in *A* indicates the default (reference) phase transition point in our model. The timing of the phase transition (*x* axes) are expressed in thermal time after plant emergence. (*D*) Rosette images of 37-d-old Col WT (*Upper*) and the greater number of smaller leaves in *Pro35S:MIR156* (*Lower*). (*E*) Area of the largest leaf in *Pro35S:MIR156*, relative to Col WT (100%), in the data of Wang et al. (54), our experimental data, and model simulation. Error bars show the SD of five plants in our study. Leaf area in Wang et al. was calculated from published leaf length and width, assuming an elliptical shape. (*F*) Model simulations (green lines) and experimental data (symbols) of individual leaf biomass in Col (■) and *Pro35S:MIR156* (□) at 37 DAS. Experimental conditions: ~20.7 °C; 12:12-h light/dark cycle; light intensity = 100 μmol·m^{-2}·s^{-1}; average daytime CO_2 concentration = 405 ppm. Error bars show the SEs of five plants.

maximum size, indicating that the proportion of large leaves was high. Taken together, our analysis suggested that increasing the leaf production rate at midvegetative stage incurs a slightly lower total biomass, relative to a later transition point, but reduces the plant's reliance on a few, large leaves.

Model-Guided Understanding of a Developmental Phenotype. The FM predicted how much rapid leaf production will reduce leaf size (Fig. 5*C*). This relationship has been described as a dual effect in plants overexpressing microRNA156 (*Pro35S:MIR156*), which have a short plastochron relative to WT plants (54). To

test whether our model could reproduce the behavior of these developmentally altered plants and explain the dual effect, we grew *Pro35S:MIR156* plants alongside Col WT for 37 d. Consistent with the previous study, *Pro35S:MIR156* plants had a higher leaf production rate and smaller leaves compared with the WT (Fig. 5*D* and *SI Appendix*, Fig. S7*B*). The size of the largest leaf in *Pro35S:MIR156* was only 57% of that in WT (Fig. 5*E*). To test whether the leaf production rate alone was sufficient to explain this phenotype, we simulated the growth of the WT and *Pro35S:MIR156* for 37 d, fixing the leaf production rate in the model to the measured rates in each ge-

notype. With only this change, our model not only replicated the observed size of the largest leaf to within the experimental error (Fig. 5E) but also closely matched the distribution of size ($R^2 = 0.90$; nRMSE = 12.9%) and biomass ($R^2 = 0.92$; nRMSE = 13.3%) for all of the individual leaves in the mutant, including their smaller size relative to the WT (Fig. 5F and *SI Appendix*, Table S10). As an additional test, we repeated the simulations with the model's simpler, piecewise linear leaf production rate, using the default values for Col and refitting the piecewise function to the data for *Pro35S: MIR156* (*SI Appendix*, Fig. S7 *B* and *C*). The model slightly underestimated leaf number in this experiment, causing an increase in the largest leaf size in both genotypes; nonetheless, the simulated mutant's largest leaf reached only 65% of the WT value, within the experimental range (Fig. 5E). Our results indicated that the observed, higher leaf production rate in *Pro35S:MIR156* plants was sufficient to predict the observed, smaller final size of each leaf, given the normal photosynthetic function and carbon partitioning among organs in the FM.

Discussion

We present methods, examples, and validation for one approach to developing a multiscale, whole-plant model of *A. thaliana*, inspired by crop science, by integrating existing models from different laboratories. The resulting FM closely matched data at multiple levels, acquired by two of our groups in different countries, allowing deeper analysis of experimental results and conceptual growth strategies. Our study suggests that a distributed, community-wide effort could successfully extend and refine the FM by integrating further, focused models into the larger framework.

Our approach stems from the recognition of potential synergies among diverse plant modelers (55), which encouraged us to integrate models from different domains using the modular approach. Ideally the integration process would not have altered the models at all, but this is unrealistic unless the models were originally designed for composition. In practice, unit conversions were required to make the models logically compatible and the FSPM was more substantially rewritten, as our aim was more limited than its original scope. Four redundant components were replaced by new connections. Only two parameter values were calibrated to our experimental data (discussed below). Another measurable parameter, the $J_{max}:V_{cmax}$ ratio that describes photosynthetic physiology, was modified using values from the literature for the 12-h photoperiod of our validation experiments. These changes were sufficient for the FM to match our experimental data (Fig. 2), confirming that the models were mutually compatible despite their different origins.

One general concern in mathematical modeling is overfitting, which becomes more significant in models of high complexity. This concern was part of our motivation to maintain the parameter values from the original models, which were already constrained to the most relevant data, instead of reoptimizing them to fit our data. In cases where unit conversions and scaling factors were required or in condition-specific scenarios, e.g., different photoperiods (see above), we adopted values directly from the literature. Although each of the four model components were calibrated and/or optimized with different techniques using independent datasets, the resulting FM matched our experimental data from two different laboratories. This broad predictive performance is generally not displayed in overfitted models.

We conducted a sensitivity analysis to examine the behavior of the FM. This analysis identified 7 of 18 photosynthetic parameters that are highly sensitive, although this number was likely underestimated because parameters related to ribulose-1,5-biphosphate carboxylase/oxygenase were redundant under our light-limiting conditions. In particular, the response of electron transport to temperature appears to have large effects on simulated biomass under our conditions. Indeed, the temperature response of key

parameters in the Farquhar model has been the focus of other studies, with model accuracy decreasing when temperatures deviated from the 25 °C condition where the model was originally parameterized (56, 57). These studies proposed different temperature response functions to improve the estimation of photosynthetic parameters, and they can be readily incorporated into the FM in future. We also identified many parameters with large effects on the simulated flowering time and thus biomass at flowering. Our results are consistent with the analysis of many crop models, which revealed high uncertainties in yield predictions at elevated CO_2 and increasing temperature, partly due to these models' simulated phenology and partly caused by the complex interactions between processes such as growth and leaf area (58). Together, our work and that of others highlight the need for improved systems understanding and mathematical representation to predict plant behavior accurately, for example, in projected, future climates.

The norms of the *Arabidopsis* research community were obviously beneficial, as each model had independently used the standard, Col accession. Nonetheless, significant variability among laboratories was recently reported even in standardized *Arabidopsis* studies (29), so compatibility of the models was not assured. The FM accurately predicted CO_2 exchange at the population level, as well as biomass and area of both total and individual leaves at various time points during rosette growth, for plants of three accessions grown under 12-h photoperiods (Fig. 3 and *SI Appendix*, Table S8). Accurate biomass and area predictions depended on simulating the temperature and lighting regimes and the CO_2 levels of each experiment and required the joint operation of the CDM and FSPM (Fig. 1; see discussion of miR156). Five or fewer accession-specific parameters were modified based on our data to obtain these results out of a total 126 parameters. These revealed limited variation in water content [88%–92%, in agreement with a previous study (59)], which had only a small effect on the fresh biomass predictions. If calibration is necessary, water content is easily measured. Variation in seedling emergence was discovered (early in Fei), because Fei was selected for its increased leaf number in a previous study (48); its early emergence was sufficient to explain our data without altered leaf appearance rate (phyllochron). Phyllochron can also easily be determined through observation or automated imaging systems (60, 61) should calibration be required (as in Fig. 5F). Flowering time variation among laboratories and accessions is common, indeed the original PTM had four accession-specific parameters (1, 38). Until the sources of variation can be identified, therefore, the flowering threshold (at least) should be calibrated to each laboratory's data, to test further regulation by the PPM (6) and PTM (1).

The FM also reproduced the measured biomass of plants grown in 12-h photoperiods under slightly different conditions (Fig. 4D), as part of a large, independent dataset testing multiple photoperiods (51). However, in shorter photoperiods, the model underestimated starch accumulation in the light and hence the rate of starch breakdown at night, as well as growth at night and total biomass under these conditions (Fig. 4). These discrepancies highlight how much the plant's carbon metabolism adapts to different photoperiod conditions. The CDM assumes a fixed relation between photosynthesis and starch accumulation in the day, a fixed proportion of starch mobilization at night, and a fixed minimum sugar level. Although this changes the model's absolute starch dynamics to some extent under different photoperiods, the responses measured in many plants are even more plastic (9). First, measured starch synthesis is faster in short than in long photoperiods, which contributed to the model's underestimating the starch level at dusk in short photoperiods (Fig. 4B), consequently underestimating the rate of starch degradation to sugars at night. Sugar dynamics are also flexible; the measured sucrose level was lower at dawn than dusk and was lower at both times under short photoperiods compared with long photo-

periods (*SI Appendix*, Fig. S5*A*). Together, these effects caused the model to underestimate growth at night in short photoperiods (Fig. 4*C*). This error might be compounded, for example, if carbon conversion efficiency was underestimated in the model or maintenance costs were overestimated but these processes were not directly measured. Second, experimental data show that the assumption of almost complete (84%) starch mobilization at night is not always applicable, for example, in long photoperiods when growth is probably sink limited. Indeed, a basic problem of many models is that they assume only source limitation (13, 14). Reducing starch breakdown to the measured level, along with a further change in one (extrapolated J_{max}:V_{cmax} ratio) or two (J_{max}:V_{cmax} ratio and measured leaf thickness) parameters, matched the data (Fig. 4*D*). Among many possible extensions, the CDM might in future be supplemented with more detail on the plant's starch dynamics, carbon partitioning, and the relationship of sucrose to growth rate (62–65).

In addition to biochemical regulation, we illustrate the potential of the FM to understand the effect of developmental programs on growth and the final rosette form. First, we introduced stochastic leaf production that reproduced the varying leaf number observed in *Arabidopsis* rosettes (Fig. 3). However, this developmental variation accounted for rather little (~12%) of the observed variation in rosette biomass. Our simulations of phyllochron (time to leaf appearance) for Ler and Fei had lower SDs compared with the SD of the plastochron (time to leaf initiation) reported in Col (53). However, leaf initiation is a developmental process, whereas leaf appearance also involves growth: variation in growth might thus compensate for variation in development, reducing the observed variance in phyllochron. A field study of sorghum varieties also found a lower deviation in phyllochron compared with plastochron (66), although the two measures were tightly related.

Second, varying the age dependence of the phyllochron (Fig. 5) suggested a tradeoff in the developmentally regulated rate of leaf production and helped us to understand its origins. The measured leaf initiation rate was initially slow and then increased. Constantly rapid leaf production reduced the simulated biomass, because the many, small leaves quickly shaded each other. Conversely, constantly slow leaf production gave a slight advantage in simulated biomass but produced few leaves, many of which were older than in the reference model. Given the risks to leaf function from predation and other damage, this suggested that the plant's strategy maintains almost maximal biomass production, without relying on an aging leaf population. Compared with the biomass-maximizing, slow-production strategy, this developmental program distributes carbon investment (and thus leaf size) more widely, a feature characteristic of bet hedging strategies that could be tested in ecological studies (67, 68).

Third, our model reproduced the smaller leaf size phenotype of the developmentally misregulated *Pro35S:MIR156* transgenic line by modifying only the model's leaf production rate (Fig. 5). Two possibilities were proposed to explain this dual effect of miR156 in the original study: (*i*) the existence of a compensatory mechanism whereby plastochron length and leaf size affect each other reciprocally to reduce changes to the overall plant biomass; or (*ii*) a common regulator that controls each of the two traits (54). The combined operation of the CDM and FSPM in the FM provides a parsimonious explanation for the dual effect. High leaf production rate requires carbon resources to be shared among more leaves (Fig. 2*F*), leading to a decrease in individual leaf growth. Using this mechanism alone, the FM matched the mutant leaf size distribution as accurately as it did the WT (Fig. 5*F*). Partitioning of a given amount of carbon among a larger number of leaves is a sufficient compensatory mechanism (54), although more complex models are of course possible. No common regulator is required to explain the observed relationship between leaf production and organ size. Similar quantitative analysis using the FM might contribute to link further research on developmental regulators (such as those targeted by miR156) and sucrose signaling (69–71) to whole-plant phenotypes and might extend to applications that modulate organ size, for example, in pruning (72, 73).

Our results on miR156 again validated the FM, particularly the benefit derived by coupling the CDM and FSPM. The FSPM did not predict growth rate based on the measured experimental conditions but rather used model inversion to learn the light use efficiency from observed plant growth data. This aggregate parameter is not directly measurable, as it combines photosynthesis, sugar-starch partitioning, respiration, and the daily allowable growth rate, which are all separately represented in the CDM. The CDM predicted sugar production and partitioning to starch based on the experimental temperature, light:dark, and CO_2 conditions but considered the rosette as one big leaf, whereas the FSPM provided information on the demand and growth of individual organs. We could only predict the biomass and detailed rosette structure in particular experimental conditions by combining these models in the FM. The FM not only explained the relationship of organ number and size in *Pro35S:MIR156* plants but also predicts that the measurable parameters of carbon utilization are unaffected in this line.

In conclusion, quantitative dynamic models are valuable both to understand and to engineer organismal growth and physiology, from the level of molecular and biochemical processes. The FM and the approach used to build it provide a flexible context to expand the detail and scope of component models, for example, to whole-cell models (28), and also to study the dynamic interactions among multiple processes. These resources will be particularly important to understand the pervasive effects of environmental stresses or pleiotropic biological regulators, such as the circadian clock. Finally, multiscale digital plant models might contribute to link systems biologists with ecophysiology and crop science, where significant synergies may be gained.

1. Chew YH, et al. (2012) An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytol* 194(3):654–665.
2. Chapman HW, Gleason LS, Loomis WE (1954) The carbon dioxide content of field air. *Plant Physiol* 29(6):500–503.
3. Franklin KA (2009) Light and temperature signal crosstalk in plant development. *Curr Opin Plant Biol* 12(1):63–68.
4. Michaels SD, Himelblau E, Kim SY, Schomburg FM, Amasino RM (2005) Integration of flowering signals in winter-annual Arabidopsis. *Plant Physiol* 137(1):149–156.
5. Corbesier L, et al. (2007) FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* 316(5827):1030–1033.
6. Salazar JD, et al. (2009) Prediction of photoperiodic regulators from quantitative gene circuit models. *Cell* 139(6):1170–1179.
7. Song YH, Smith RW, To BJ, Millar AJ, Imaizumi T (2012) FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science* 336(6084):1045–1049.
8. Valverde F, et al. (2004) Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303(5660):1003–1006.
9. Smith AM, Stitt M (2007) Coordination of carbon supply and plant growth. *Plant Cell Environ* 30(9):1126–1149.
10. Ruberti I, et al. (2012) Plant adaptation to dynamically changing environment: The shade avoidance response. *Biotechnol Adv* 30(5):1047–1058.
11. López-Bucio J, Cruz-Ramírez A, Herrera-Estrella L (2003) The role of nutrient availability in regulating root architecture. *Curr Opin Plant Biol* 6(3):280–287.
12. Yin X, Struik PC (2010) Modelling the crop: From system dynamics to systems biology. *J Exp Bot* 61(8):2171–2183.

13. Poorter H, Anten NPR, Marcelis LFM (2013) Physiological mechanisms in plant growth models: Do we need a supra-cellular systems biology approach? *Plant Cell Environ* 36(9):1673–1690.
14. Boote KJ, Jones JW, White JW, Asseng S, Lizaso JI (2013) Putting mechanisms into crop production models. *Plant Cell Environ* 36(9):1658–1672.
15. Lucas M, Laplaze L, Bennett MJ (2011) Plant systems biology: Network matters. *Plant Cell Environ* 34(4):535–553.
16. Bernacchi CJ, et al. (2013) Modelling C₃ photosynthesis from the chloroplast to the ecosystem. *Plant Cell Environ* 36(9):1641–1657.
17. Nedbal L, Cerveny J, Schmidt H (2009) *Photosynthesis In Silico*, eds Laisk A, Nedbal L, Govindjee (Springer, Dordrecht, The Netherlands), Vol 29, pp 17–29.
18. Deinum EE, Geurts R, Bisseling T, Mulder BM (2012) Modeling a cortical auxin maximum for nodulation: Different signatures of potential strategies. *Front Plant Sci* 3:96.
19. Hunter PJ, Crampin EJ, Nielsen PMF (2008) Bioinformatics, multiscale modeling and the IUPS Physiome Project. *Brief Bioinform* 9(4):333–343.
20. Villa F (2001) Integrating modelling architecture: A declarative framework for multi-paradigm, multi-scale ecological modelling. *Ecol Modell* 137(1):23–42.
21. Root TL, Schneider SH (1993) Can large-scale climatic models be linked with multiscale ecological studies. *Conserv Biol* 7(2):256–270.
22. Collins WD, et al. (2006) The Community Climate System Model version 3 (CCSM3). *J Clim* 19(11):2122–2143.
23. Overstreet CM, Nance RE, Balci O (2002) Issues in enhancing model reuse. *Proceedings of the International Conference on Grand Challenges for Modeling and Simulation*, January 2002, Texas, pp 27–31.
24. Brisson N, et al. (2003) An overview of the crop model STICS. *Eur J Agron* 18(3-4):309–332.
25. Bergez JE, et al. (2013) An open platform to build, evaluate and simulate integrated models of farming and agro-ecosystems. *Environ Model Softw* 39:39–49.
26. Hunter PJ, Viceconti M (2009) The VPH-Physiome Project: Standards and tools for multiscale modeling in clinical applications. *IEEE Rev Biomed Eng* 2:40–53.
27. Pradal C, Dufour-Kowalski S, Boudon F, Fournier C, Godin C (2008) OpenAlea: A visual programming and component-based software platform for plant modelling. *Funct Plant Biol* 35(9-10):751–760.
28. Karr JR, et al. (2012) A whole-cell computational model predicts phenotype from genotype. *Cell* 150(2):389–401.
29. Massonnet C, et al. (2010) Probing the reproducibility of leaf growth and molecular phenotypes: A comparison of three Arabidopsis accessions cultivated in ten laboratories. *Plant Physiol* 152(4):2142–2157.
30. Nakamichi N (2011) Molecular mechanisms underlying the Arabidopsis circadian clock. *Plant Cell Physiol* 52(10):1709–1718.
31. Rasse DP, Tocquin P (2006) Leaf carbohydrate controls over Arabidopsis growth and response to elevated CO2: An experimentally based model. *New Phytol* 172(3):500–513.
32. Gibon Y, et al. (2009) Adjustment of growth, starch turnover, protein content and central metabolism to a decrease of the carbon supply when Arabidopsis is grown in very short photoperiods. *Plant Cell Environ* 32(7):859–874.
33. Gibon Y, et al. (2004) Adjustment of diurnal starch turnover to short days: Depletion of sugar during the night leads to a temporary inhibition of carbohydrate utilization, accumulation of sugars and post-translational activation of ADP-glucose pyrophosphorylase in the following light period. *Plant J* 39(6):847–862.
34. Christophe A, et al. (2008) A model-based analysis of the dynamics of carbon balance at the whole-plant level in Arabidopsis thaliana. *Funct Plant Biol* 35(11):1147–1162.
35. Yanovsky MJ, Kay SA (2002) Molecular basis of seasonal time measurement in Arabidopsis. *Nature* 419(6904):308–312.
36. Sung S, Amasino RM (2004) Vernalization and epigenetics: How plants remember winter. *Curr Opin Plant Biol* 7(1):4–10.
37. Kumar SV, et al. (2012) Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature* 484(7393):242–245.
38. Wilczek AM, et al. (2009) Effects of genetic perturbation on seasonal life history plasticity. *Science* 323(5916):930–934.
39. Locke JC, et al. (2005) Extension of a genetic network model by iterative experimentation and mathematical analysis. *Mol Syst Biol* 1:0013.
40. Gorsuch PA, Pandey S, Atkin OK (2010) Thermal de-acclimation: How permanent are leaf phenotypes when cold-acclimated plants experience warming? *Plant Cell Environ* 33(7):1124–1137.
41. Gorsuch PA, Pandey S, Atkin OK (2010) Temporal heterogeneity of cold acclimation phenotypes in Arabidopsis leaves. *Plant Cell Environ* 33(2):244–258.
42. Kumar S, Udawatta RP, Anderson SH (2010) Root length density and carbon content of agroforestry and grass buffers under grazed pasture systems in a Hapludalf. *Agrofor Syst* 80(1):85–96.
43. Boyes DC, et al. (2001) Growth stage-based phenotypic analysis of Arabidopsis: A model for high throughput functional genomics in plants. *Plant Cell* 13(7):1499–1510.
44. Muetzelfeldt R, Massheder J (2003) The Simile visual modelling environment. *Eur J Agron* 18(3-4):345–358.
45. Pons TL (2012) Interaction of temperature and irradiance effects on photosynthetic acclimation in two accessions of Arabidopsis thaliana. *Photosynth Res* 113(1-3):207–219.
46. Flexas J, et al. (2007) Mesophyll conductance to CO2 in Arabidopsis thaliana. *New Phytol* 175(3):501–511.
47. Bunce JA (2008) Acclimation of photosynthesis to temperature in Arabidopsis thaliana and Brassica oleracea. *Photosynthetica* 46(4):517–524.
48. Méndez-Vigo B, de Andrés MT, Ramiro M, Martínez-Zapater JM, Alonso-Blanco C (2010) Temporal analysis of natural variation for the rate of leaf production and its relationship with flowering initiation in Arabidopsis thaliana. *J Exp Bot* 61(6):1611–1623.
49. Li YL, et al. (2008) Patterns in CO2 gas exchange capacity of grassland ecosystems in the Alps. *Agric Meteorol* 148(1):51–68.
50. Williams M, Street LE, van Wijk MT, Shaver GR (2006) Identifying differences in carbon exchange among arctic ecosystem types. *Ecosystems (NY)* 9(2):288–304.
51. Sulpice R, et al. (2014) Arabidopsis coordinates the diurnal regulation of carbon allocation and growth across a wide range of photoperiods. *Mol Plant* 7(1):137–155.
52. Wang F, et al. (2011) A stochastic model of tree architecture and biomass partitioning: Application to Mongolian Scots pines. *Ann Bot (Lond)* 107(5):781–792.
53. Guenot B, et al. (2012) Pin1-independent leaf initiation in Arabidopsis. *Plant Physiol* 159(4):1501–1510.
54. Wang JW, Schwab R, Czech B, Mica E, Weigel D (2008) Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in Arabidopsis thaliana. *Plant Cell* 20(5):1231–1243.
55. Thomas H (2008) Systems biology and the biology of systems: How, if at all, are they related? *New Phytol* 177(1):11–15.
56. Bernacchi CJ, Singsaas EL, Pimentel C, Portis AR, Long SP (2001) Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant Cell Environ* 24(2):253–259.
57. Medlyn BE, Loustau D, Delzon S (2002) Temperature response of parameters of a biochemically based model of photosynthesis. I. Seasonal changes in mature maritime pine (Pinus pinaster Ait.). *Plant Cell Environ* 25(9):1155–1165.
58. Asseng S, et al. (2013) Uncertainty in simulating wheat yields under climate change. *Nature Climate Change* 3(9):827–832.
59. Cross JM, et al. (2006) Variation of enzyme activities and metabolite levels in 24 Arabidopsis accessions growing in carbon-limited conditions. *Plant Physiol* 142(4):1574–1588.
60. Tisné S, et al. (2013) Phenoscope: An automated large-scale phenotyping platform offering high spatial homogeneity. *Plant J* 74(3):534–544.
61. Granier C, et al. (2006) PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytol* 169(3):623–635.
62. Scialdone A, et al. (2013) Arabidopsis plants perform arithmetic division to prevent starvation at night. *Elife (Cambridge)* 2:e00669.
63. Seaton DD, Ebenhöh O, Millar AJ, Pokhilko A (2014) Regulatory principles and experimental approaches to the circadian control of starch turnover. *J R Soc Interface* 11(91):20130979.
64. Pokhilko A, Flis A, Sulpice R, Stitt M, Ebenhöh O (2014) Adjustment of carbon fluxes to light conditions regulates the daily turnover of starch in plants: A computational model. *Mol Biosyst* 10(3):613–627.
65. Pal SK, et al. (2013) Diurnal changes of polysome loading track sucrose content in the rosette of wild-type arabidopsis and the starchless *pgm* mutant. *Plant Physiol* 162(3):1246–1265.
66. Clerget B, Dingkuhn M, Gozé E, Rattunde HFW, Ney B (2008) Variability of phyllochron, plastochron and rate of increase in height in photoperiod-sensitive sorghum varieties. *Ann Bot (Lond)* 101(4):579–594.
67. Simons AM (2009) Fluctuating natural selection accounts for the evolution of diversification bet hedging. *Proc R Soc B Biol Sci* 276(1664):1987–1992.
68. Simons AM (2011) Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proc R Soc B Biol Sci* 278(1712):1601–1609.
69. Wahl V, et al. (2013) Regulation of flowering by trehalose-6-phosphate signaling in Arabidopsis thaliana. *Science* 339(6120):704–707.
70. Yang L, Xu M, Koo Y, He J, Poethig RS (2013) Sugar promotes vegetative phase change in *Arabidopsis thaliana* by repressing the expression of *MIR156A* and *MIR156C*. *Elife (Cambridge)* 2:e00260.
71. Yu S, et al. (2013) Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. *Elife (Cambridge)* 2:e00269.
72. Baldet P, et al. (2006) The expression of cell proliferation-related genes in early developing flowers is affected by a fruit load reduction in tomato plants. *J Exp Bot* 57(4):961–970.
73. Dash M, Johnson LK, Malladi A (2013) Reduction of fruit load affects early fruit growth in apple by enhancing carbohydrate availability, altering the expression of cell production-related genes, and increasing cell production. *J Am Soc Hortic Sci* 138:253–262.

Supporting Information

Chew et al. 10.1073/pnas.1410238111
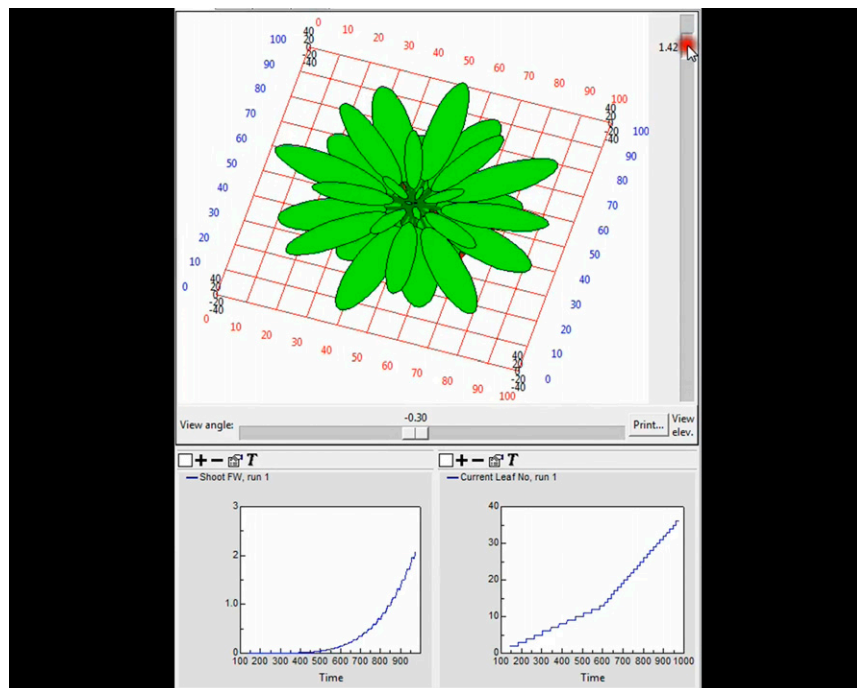
Movie S1. The Simile visual animation tool for the framework model. (*Upper*) Animation of an *Arabidopsis* rosette simulated using the input growth conditions of 21.3 °C, 12:12-h light/dark cycle, 110 μmol·m^{-2}·s^{-1}, and an average daytime CO_2 level of 375 ppm. (*Lower Left*) Plot of shoot fresh weight. (*Lower Right*) Plot of rosette leaf number.

[Movie S1](#)

Other Supporting Information Files

[SI Appendix (PDF)](#)

Table of Contents

1   Experimental procedures

1.1   Plant materials and growth conditions

Arabidopsis thaliana from the Columbia (Col), Landsberg *erecta* (Ler) and Feira (Fei) (kindly provided by Dr. Ronan Sulpice from Max Planck Institute for Molecular Plant Physiology, Potsdam-Golm, Germany) accessions were used in this study. *Pro35S:MIR156* seeds were obtained from the European Arabidopsis Stock Centre (NASC). Seeds were first sown on half strength Murashige and Skoog (MS) solution and stratified in darkness at 4 °C for 5 days before being exposed to white light at the desired photoperiod and temperature. Four-day-old seedlings were then transferred to trays of 6 x 4 pots containing Levington seed and modular compost (plus sand). Light was provided by Polylux XL_R 835 (GE Lighting) triphosphor white fluorescent tubes.

1.2   Leaf number and biomass assay

The total number of leaves (including the cotyledons) was recorded every 3-4 days from germination until the flower buds appeared. Only leaves exceeding 1 mm^2 in size (estimated with the naked eye) were considered in the total leaf count. Biomass was also measured at different time points throughout the vegetative stage. Plants were selected randomly from each accession and rosettes were dissected based on order of appearance (leaf rank) for the measurement of individual leaf biomass. Leaves were nipped around the edges and pressed flat to obtain actual area. For the assay at the final time point, only plants with visible flower buds were selected to ensure they were all at the same developmental stage. For dry biomass, dissected plants were dried in the oven at 80 °C for 7 days. Area analysis was conducted using ImageJ (56). Each image was first processed with colour thresholding to isolate the green region, which was next converted into binary format. The area was then determined using the Analyze Particles tool.

1.3   Gas exchange measurement

An EGM-4 Environmental Gas Monitor for CO_2 (PP Systems, US) was used for CO_2 flux measurement. A 50 cm x 35 cm x 20 cm Plexiglass chamber was connected to the EGM-4 with two butyl tubes for closed-loop measurement (Fig. S4). An airtight seal was created between shelf top and chamber by lining adjoining surfaces with neoprene and pressing the chamber tightly to the shelf during measurement. Each individual measurement was taken by placing a tray or pots of plants inside the chamber for approximately 60 seconds, during which the EGM-4 recorded CO_2 concentration (μmol mol^{-1} or ppm) every 4.6 seconds. To ensure good circulation inside the chamber, two small fans were installed at opposite corners. CO_2 flux was calculated from the slope of CO_2 concentration plotted against time as follows:

$$F_c = \frac{dC}{dt} \cdot \rho V \qquad\qquad (1.3.1)$$

$$\rho = \frac{P}{RT} \qquad\qquad (1.3.2)$$

where

F_c	=	net CO_2 flux (μmol s^{-1})
ρ	=	air density (mol m^{-3})
V	=	air volume in the chamber (m^3)
dC/dt	=	slope of CO_2 concentration against time (μmol mol^{-1} s^{-1})
P	=	pressure (Pa)
R	=	ideal gas constant, 8.314 (m^3 Pa mol^{-1} K^{-1})
T	=	temperature (K)

CO_2 flux was measured every day from 26 days after sowing until flower buds first became visible to the naked eye. We included two negative controls: 1) a tray of soil without any plants and; 2) a tray of soil that was constantly covered with black opaque plastic. We also covered the soil surface of the pots with black opaque plastic, leaving only a small hole in the middle for the plants. CO_2 enrichment of the atmosphere in the growth chambers due to the experimenters' breathing proved a particular issue in preliminary studies; this was corrected using a breath-scrubbing device. Hourly CO_2 concentration at leaf level was also monitored by connecting the EGM-4 to a computer for automated data logging. The average hourly CO_2 level was used as input to the model.

Measured NEP:

A sample measurement for a tray of 5-week-old Ler is shown in Fig. S17. NEP was determined from the slope of regression using equation 3.3.1. No significant readings could be detected for either of the negative controls, suggesting that soil respiration or any other gas exchange activities on the soil surface was very low. All significant readings were therefore assumed to be due to the gas exchange activities of the Arabidopsis plants.

Modelled NEP:

Modelled NEP per plant was converted from g C m^{-2} h^{-1} to μmol C m^{-2} s^{-1} using the following:

$$NEP = [\text{Carbon assimilation} - R_{above} - R_{below} - R_{ml} - R_{mr}] \times \frac{1}{3600} \times \frac{1}{12} \times 10^6 \qquad (1.3.3)$$

The modelled NEP was multiplied by the number of plants on a tray or the number of pots measured for comparison with the measured values.

2 Model Implementation

This section describes the original format of the four existing models and how they were implemented in MATLAB, the modelling environment used in this study. Steps were taken to ensure reproducibility of simulated values from the original models. In addition, we also implemented the models in Simile. MATLAB is a generic programming language that offers a library of commonly used functions, with a large community of users. On the other hand, Simile has an intuitive graphic user interface (GUI) specifically for ecological and biological modelling. Implementing the models in both MATLAB and Simile therefore offers programmers and skilled modellers the flexibility to conduct in-depth model analysis in MATLAB, while allowing beginners and non-modellers to do basic simulation with a visual display in Simile (see Supplementary Video).

2.1   The Four Component Models

As described in the main text, there are four component models that make up the Framework Model (FM). They are: 1) a Carbon Dynamic Model (CDM); 2) a Functional-Structural Plant Model (FSPM); 3) a Photothermal Model (PTM) and; 4) a Photoperiodism Model (PPM). These models were each implemented in a different format. Since linking them in their original formats would have been technically challenging, we re-implemented each one in a single modelling environment, MATLAB. The Simile versions were developed based on the MATLAB codes, except for the CDM. These are described in the following.

The original CDM was developed in Fortran 95 (1). However, this model was already re-implemented in Simile based on the Fortran source code as part of the BBSRC-funded PlaSMo project (BBR), with the aim of fully reproducing the behaviour of the original model. During the Simile re-implementation, the original Fortran was modified slightly to facilitate inspection of internal values, run using GNU Fortran 95, and values for state variables were compared. Complete agreement between the two versions was obtained. The MATLAB version was developed based on the Simile implementation, again with complete agreement.

The FSPM was adapted from the GreenLab model (2), as described in Section 3.2 below. GreenLab allows for complex branching structures, but this complexity is not required for the simple, non-branching rosette architecture of Arabidopsis. In addition, the original MATLAB code for GreenLab was built in a structure that facilitates model inversion and computation of complex branching topology (3), and these parameter estimation aspects were also not required. Therefore, the FSPM was re-created in MATLAB based on the relevant concepts in GreenLab, but in a conventional, dynamic form that was compatible with the other component models in our study. Our simplified version was also developed to run at a time step of one hour instead of a growth cycle (see Section 3.2) that was used in the original model (2). Since the model requires optimised light-use efficiency (LUE) as an input, it cannot be used as a stand-alone model.

The PTM used here is the Chew et al. revised version of an Arabidopsis phenology model (4). Originally the Wilczek et al. model (5) was developed in Excel, but the revised version was built in MATLAB (4), which was used directly in the current study. The fourth component model, the PPM as described in Salazar et al., was already in MATLAB originally (6) and so it was also used directly.

2.2   The Framework Model (FM)

To integrate the four models, a main function file was created, which takes environmental conditions as input arguments and contains all the model initialisation equations (Section 3.9). The FSPM is located in this function file while the other three model components (which were written as function files) are called within this file so input and output variables can be passed around sequentially in each time step (see details in sections 3.6-3.9). As new connections were formed, all the redundant equations from the original models were removed from the MATLAB codes. Development of the Simile version of the FM was more straightforward as this modelling environment supports plug-and-play modularity, allowing a composite model (in this case the FM) to be created from two or more sub-models (the component models). Owing to differences in numerical solvers, floating-point representation and a rigid time-stepping scheme, the Simile version generated small numerical differences (< 2.5%) from the MATLAB version. However, the Simile version brings major advantages as mentioned earlier, so we regard these small numerical differences as acceptable. All the results shown in the main text (except for the simulated rosette images in Fig. 5C inset) and supplementary information (except for the Supplementary Video) are simulations from the MATLAB version.

5

3 Model description

In this section, each of the existing models is presented in their original stand-alone form, to provide a single, consistent description of the components. The integration process is described, including the standardisation of units and replacement/removal of model components. The operation of the resulting FM is described, including model initialisation and case-specific parameter values. Finally, the stochastic version of the FM is presented.

3.1 The Carbon Dynamic Model (CDM)

This model (Fig. S10) consists of two mechanistic components, i.e. carbon assimilation and sugar-starch metabolism (1). In the carbon assimilation component, leaf photosynthesis is determined using the Farquhar et al. (7) equations that consider CO_2 level, light intensity and temperature. This classical biochemical model considers the two rate-limiting factors in photosynthesis: the carboxylase activity of RuBisCO and electron transport (8). An important parameter in the former is the maximum rate of carboxylation (V_{cmax}), which can be determined by conducting a series of gas exchange measurements at different CO_2 levels under saturating irradiance. On the other hand, electron transport is light dependent with a potential rate of J_{max} that can also be determined experimentally. In the Rasse and Tocquin study (1), only V_{cmax} was measured, while J_{max} was set to 2.1 times V_{cmax}, the same ratio as calculated in Farquhar et al. (7).

In the sugar-starch partitioning component, a fixed proportion of the photoassimilate calculated from the assimilation component is turned into transitory starch at a baseline rate (9, 10), while the rest is converted into soluble sugars which are used for growth and respiration. Two types of respiration are considered in the model: maintenance respiration which depends on sugar concentration (1, 11), and growth respiration which is a fixed fraction of the growth demand. The remaining carbon available for growth is then allocated to leaves and roots based on the root-to-shoot allocation ratio, *RS*. As growth demand is limited to a maximum rate, any excess photoassimilate is converted into starch through an overflow mechanism (12, 13). At night time, starch is broken down into sugar to sustain growth. The model adopted a linear breakdown rate where a constant percentage of the end-of-day starch is consumed by the end of the night regardless of the night length. Such a behaviour has been observed in Arabidopsis (14, 15) and was found to be circadian regulated (16, 17). Translocation from the organs, i.e. leaves and roots, is also allowed to maintain a minimum sugar level. This element was created in the Rasse and Tocquin (1) model to reflect observations in an Arabidopsis starchless mutant in which sugar was consumed rapidly in the dark period but never reached zero, while genes involved in the breakdown of non-starch energy sources were up-regulated at the end of the night (18, 19). Model equations for each component are as follows:

Carbon assimilation:

The following equations represent the biochemical model of photosynthetic CO_2 assimilation in C_3 plants developed by Farquhar et al.(7). In general, this model considers two rate limiting factors in the Calvin-Benson cycle: (1) RuBisCO carboxylation; and (2) electron transport.

RuBisCO activity:

The parameter values of kinetic constants at leaf temperature, T_{leaf}, are first determined using values at the standard temperature 25° C as follows:

$$K_c(T_{leaf}) = K_c(25^0 C) \cdot \exp\left(\frac{H_c \cdot (T_{leaf} - 25)}{298 \cdot R \cdot (T_{leaf} + 273)} \right) \tag{3.1.1}$$

$$K_o(T_{leaf}) = K_o(25^0 C) \cdot \exp\left(\frac{H_o \cdot (T_{leaf} - 25)}{298 \cdot R \cdot (T_{leaf} + 273)} \right) \tag{3.1.2}$$

$$V_{c_{max}}(T_{leaf}) = V_{c_{max}}(25^0 C) \cdot \exp\left(\frac{H_V \cdot (T_{leaf} - 25)}{298 \cdot R \cdot (T_{leaf} + 273)} \right) \tag{3.1.3}$$

Rate of assimilation limited by RuBisCO, A_c, is then calculated as:

$$A_c = V_{c_{max}}(T_{leaf}) \cdot \frac{[CO_2]_i - \Gamma *}{[CO_2]_i + K_c(T_{leaf}) \cdot \left(1 + \frac{[O_2]}{K_o(T_{leaf})} \right)} \tag{3.1.4}$$

where $[O_2]$ is the O_2 partial pressure having a constant value of 20500 Pa, while $[CO_2]_i$ is the intercellular CO_2 partial pressure given by

$$[CO_2]_i = p_1 \cdot [CO_2] \tag{3.1.5}$$

The CO_2 compensation point in the absence of mitochondrial respiration, $\Gamma *$, is determined as

$$\Gamma * = p_2 + p_3 \cdot (T_{leaf} - 25) + p_4 \cdot (T_{leaf} - 25)^2 \tag{3.1.6}$$

Electron transport:

The potential rate of electron transport at leaf temperature, J_{max} (T_{leaf}), is determined using the following equation:

$$J_{max}(T_{leaf}) = J_{max}(25^0 C) \cdot \frac{\exp\left(\dfrac{H_J \cdot (T_{leaf} - 25)}{298 \cdot R \cdot (T_{leaf} + 273)}\right) \cdot \left(1 + \exp\left(\dfrac{298 \cdot p_5 - p_6}{298 \cdot R}\right)\right)}{1 + \exp\left(\dfrac{(T_{leaf} + 273) \cdot p_5 - p_6}{(T_{leaf} + 273) \cdot R}\right)}$$

(3.1.7)

J_{max} (25° C) can be measured experimentally, but in Rasse and Tocquin (1) and in the current study, this parameter is estimated as

$$J_{max}(25^0 C) = R_{JV} \cdot V_{c_{max}}(25^0 C)$$

(3.1.8)

The rate of electron transport, J, depends on the irradiance PAR (µmol photon m^{-2} s^{-1}) and is calculated by solving the following quadratic function:

$$p_7 \cdot J^2 - \left(\frac{PAR \cdot (1 - f_{spec})}{2} + J_{max}(T_{leaf})\right) \cdot J + \frac{PAR \cdot (1 - f_{spec})}{2} \cdot J_{max}(T_{leaf}) = 0$$

(3.1.9)

Rate of assimilation limited by electron transport, A_j, is determined as:

$$A_j = \frac{J \cdot ([CO_2]_i - \Gamma^*)}{4 \cdot ([CO_2]_i - 2 \cdot \Gamma^*)}$$

(3.1.10)

Net rate of carbon assimilation:

The net rate of leaf photosynthesis per unit area, A_{net} (µmol CO_2 m^{-2} s^{-1}), is given by the minimum of the two rate limiting factors described earlier. Therefore:

$$A_{net} = \text{Min}(A_c, A_j)$$

(3.1.11)

In the case where no solution can be found for equation 3.1.9, A_{net} takes the value of A_c.

The total amount of carbon assimilated per plant (in g C) in each hour, which is the time step used in this study, is calculated using conversion factors as follows:

Carbon assimilation (t) = 12 x 10^{-6} x 3600 x $A_{net}(t)$ x Rosette area $(t-1)$ (3.1.12)

The projected rosette area at the end of the previous time point is used as the total area available for light interception.

Carbon partitioning:

The following equations represent the mechanistic allocation and growth model for Arabidopsis (1). This model considers four carbon pools: (1) Starch carbon; (2) Sugar carbon; (3) Leaf carbon; and (4) Root carbon.

Starch-sugar partitioning:

At day time, a fixed baseline portion of carbon assimilates is stored as transitory starch and the rest is turned into sugar. Therefore:

Starch synthesis (t) = ST_{br} x Carbon assimilation (t) (3.1.13)

Partition to sugar (t) = Carbon assimilation (t) − Starch synthesis (t) (3.1.14)

At night time, when there is no photosynthesis, a fixed portion (ST_c) of the starch accumulated at the end of the day (EoD) is degraded at a linear rate to form sugar, as follows:

$$\text{Starch degradation } (t) = \frac{ST_c \times [\text{Starch carbon}]_{EoD}}{\text{Night length}}$$ (3.1.15)

where night length is measured in hours.

Maintenance respiration:

Carbon is constantly respired for plant maintenance, and leaf maintenance respiration per unit area, R_l, at 20° C has been found to have the following linear correlation with sugar content (1):

$$R_l(t, 20^0 C) = p_8 \cdot [\text{Sugar content per unit rosette area } (t-1)] + p_9$$ (3.1.16)

The amount of carbon respired for maintenance in each corresponding time step is estimated from the sugar content at the end of the previous time step. If sugar content is 0, R_l is set to 0.

The above-ground and below-ground maintenance respirations at leaf temperature are computed as:

$$R_l(T_{leaf}) = R_l(20^0\text{C}) \cdot \exp\left(\frac{H_r \cdot (T_{leaf} - 20)}{293 \cdot R \cdot (T_{leaf} + 273)}\right)$$
(3.1.17)

Total leaf maintenance respiration is calculated as:

$$R_{above}(t) = R_l(T_{leaf}) \times \text{Rosette area } (t\text{-}1)$$
(3.1.18)

Assuming that root maintenance respiration per unit carbon is similar to that in the leaves, the total root maintenance respiration is given by:

$$R_{below}(t) = R_{above}(t) \cdot \frac{\text{Root carbon } (t-1)}{\text{Leaf carbon } (t-1)}$$
(3.1.19)

Carbon available for growth:

The transient amount of carbon available for growth, Q_{trans}, is the excess from the sugar pool after balancing all the carbon fluxes in the corresponding time step as follows:

$$Q_{trans}(t) = \text{Sugar carbon } (t\text{-}1) + S(t) - R_{above}(t) - R_{below}(t)$$
(3.1.20)

where

$$S(t) = \begin{cases} \text{Partition to sugar}(t) & sunrise < t \le sunset \\ \text{Starch degradation}(t) & \text{otherwise} \end{cases}$$
(3.1.21)

However, a minimum sugar content of $SSU_{\min}$ (g sugar-C m^{-2}) has to be maintained. Therefore carbon availability, C_{avail}, is calculated as:

$$C_{avail}(t) = Q_{trans}(t) - SSU_{\min} \times \text{Rosette area } (t\text{-}1)$$
(3.1.22)

If C_{avail} has a negative value, no carbon is available for growth in that time step. Instead, carbon is translocated from the leaf and root carbon pools to the sugar pool to maintain its minimum level. As such:

Component	Value used		
	$C_{avail}(t) \geq 0$	$C_{avail}(t) < 0$	
Carbon available for growth, $Q_c(t)$	$C_{avail}(t)$	0	(3.1.23)
Translocation from leaf carbon, $TransL(t)$	0	$-C_{avail}(t) \cdot \dfrac{\text{Leaf carbon}\,(t-1)}{\text{Total}\,(t-1)}$	(3.1.24)
Translocation from root carbon, $TransR(t)$	0	$-C_{avail}(t) \cdot \dfrac{\text{Root carbon}\,(t-1)}{\text{Total}\,(t-1)}$	(3.1.25)

where

$$\text{Total}\,(t-1) = \text{Leaf carbon}\,(t-1) + \text{Root carbon}\,(t-1) \tag{3.1.26}$$

Organ growth demand:

The model distinguishes between carbon demand (sink strength) and actual growth. Leaf (or rosette) growth demand per time step, d_L (in g C per hour), is limited by a maximum daily relative growth rate, GR_{max} (g g^{-1} d^{-1}), as shown in the following:

$$d_L(t) = GR_{max} \times \text{Leaf carbon}\,(t-1) \times \frac{1}{24} \tag{3.1.27}$$

Note: In CDM, the whole rosette is treated as one big leaf without consideration of each individual leaf.

A fraction, α, of the total growth demand is used for growth respiration, therefore the required leaf growth respiration, d_{RL}, is:

$$d_{RL}(t) = \frac{d_L(t) \cdot \alpha}{1 - \alpha} \tag{3.1.28}$$

Root growth demand, d_R, is calculated from the leaf growth demand using the root-to-shoot allocation ratio, RS, as in the following:

$$RS(t) = \begin{cases} 0.12 & \text{Rosette area } (t-1) < 1.22e-4 \\ 0.0496 + 555 \cdot [\text{Rosette area } (t-1)] & 1.22e\text{-}4 < \text{Rosette area } (t-1) < 3.96e\text{-}4 \\ 0.27 & \text{otherwise} \end{cases}$$

(3.1.29)

$$d_R(t) = RS(t) \times d_L(t)$$ (3.1.30)

Note: In the current study, RS in equation 3.1.29 has been replaced by the ratio of root sink demand to shoot sink demand derived from the FSPM (equation 3.6.1).

Similarly, the required root growth respiration, d_{RR}, can be calculated as:

$$d_{RR}(t) = \frac{d_R(t) \cdot \alpha}{1 - \alpha}$$ (3.1.31)

Therefore, the total growth demand for carbon, D_c, is given by:

$$D_c(t) = d_L(t) + d_{RL}(t) + d_R(t) + d_{RR}(t)$$ (3.1.32)

Allocation:

The actual growth depends on whether the total growth demand is met by the carbon available for growth, as shown below:

	Value used		
Growth component	**Full demand is met**	**Full demand is not met**	
	$D_c(t) \leq Q_c(t)$	$D_c(t) > Q_c(t)$	
Leaf growth, $L(t)$	$d_L(t)$	$\frac{d_L(t)}{D_c(t)} \cdot Q_c(t)$	(3.1.33)
Leaf growth respiration, $R_{ml}(t)$	$d_{RL}(t)$	$\frac{d_{RL}(t)}{D_c(t)} \cdot Q_c(t)$	(3.1.34)
Root growth, $R(t)$	$d_R(t)$	$\frac{d_R(t)}{D_c(t)} \cdot Q_c(t)$	(3.1.35)

Root growth respiration, $R_{mr}(t)$	$d_{RR}(t)$	$\dfrac{d_{RR}(t)}{D_c(t)} \cdot Q_c(t)$	(3.1.36)

In the case where the full demand is met, any excess carbon is transferred to the starch pool in the model to represent starch production through an overflow mechanism, O_{sta}. This overflow starch production is constrained to the light period.

$$O_{sta}(t) = \begin{cases} Q_c(t) - D_c(t) & sunrise < t \le sunset \\ 0 & otherwise \end{cases}$$

(3.1.37)

Amount of carbon in each pool:

Leaf carbon (t) = Leaf carbon (t-1) + $L(t)$ − $TransL(t)$ (3.1.38)

Root carbon (t) = Root carbon (t-1) + $R(t)$ − $TransR(t)$ (3.1.39)

Starch carbon (t) = Starch carbon (t-1) + Starch synthesis (t) −
                Starch degradation (t) + $O_{sta}(t)$ (3.1.40)

Sugar carbon (t) = Sugar carbon (t-1) + $S(t)$ − $R_{above}(t)$ − $R_{below}(t)$ −
                $R_{ml}(t)$ − $R_{mr}(t)$ − $O_{sta}(t)$ − $L(t)$ − $R(t)$ + $TransL(t)$ +
                $TransR(t)$ (3.1.41)

All the parameter values are listed in Table S1. The parameter values, as used in the original study, were either extracted from the literature or determined through experimental measurements (1).

3.2   The Functional-Structural Plant Model (FSPM)

The GreenLab functional-structural plant growth concept for Arabidopsis (2) was utilised here (Fig. S11). The general GreenLab model adopts the concept of the elementary growth unit known as the phytomer or metamer. For an Arabidopsis rosette, each phytomer consists of a leaf with negligible internode (stem) elongation. Therefore, the GreenLab model was adapted to the simpler non-branching structure of Arabidopsis rosette, and we re-defined physiological age as the time since organ emergence to reflect the physiological transition in leaves from carbon sink to carbon source, as they mature. Plant age indicates the time since sowing (t) while the physiological age of each leaf (phytomer) is the thermal time elapsed since its

13

appearance (n). A growth cycle (numbered in j) is the period between the appearances of two successive phytomers (phyllochron).

The first growth cycle starts from the time of sowing until plant emergence, which is the stage when the cotyledons (the first phytomer) are fully opened (20). Upon plant emergence, the early vegetative stage follows, where one phytomer appears approximately every 30.3 degree days (°Cd) (2). Above 355 °Cd after plant emergence, a phytomer appears every 11.9 °Cd.

As the first phytomer consists of two leaves, the number of existing leaves after j growth cycles, $L(j)$, is therefore:

$$L(j) = y(j) + 1 \tag{3.2.1}$$

where $y(j)$ is the number of phytomers after j growth cycles.

Besides organogenesis, growth processes are also computed in each growth cycle and they can be generally divided into two groups: the source and the sink (Fig. S11). For the first growth cycle, seed biomass is assumed to be the only source (2).

From the second cycle onwards, biomass production Q (source) is determined by a simple function of light use efficiency (LUE), rosette area at the end of the previous cycle for light interception (S_{ros}) and the photosynthetic active radiation (PAR) as in the following:

$$Q(j) = LUE(j) \cdot S_{ros}(j-1) \cdot PAR(j) \tag{3.2.2}$$

Only the vertically projected rosette area is considered in the model, therefore S depends on the phyllotaxy and zenithal angle (from the horizontal surface) of each leaf. In Arabidopsis, both cotyledons appear opposite each other during plant emergence. The third leaf is initiated at an angle close to 90° from one of the cotyledons and the fourth leaf is almost 180° from the third leaf. For the fifth leaf onwards, leaves appear with a spiral phyllotaxy(21) of angle varying from 137.5° (22) to 138.2° (23). This range of angle is approximated by a phyllotaxy of 5/13 (138.46°), where after 5 rotations or 13 leaves, the fourteenth leaf appears on the same orientation as the first leaf. Therefore, when the rosette leaf number exceeds 15 (as the first three leaves do not follow the spiral phyllotaxy), only the 13 largest functional leaves are considered for light interception due to self-shading:

$$S_{ros}(j-1) = \underset{k \in \{1,...,L_j\}}{\text{Max}} \left(\sum_{i=k}^{k+12} S_i(j-1) \cdot \cos \alpha_i \right) \tag{3.2.3}$$

where S_i is the area of the leaf of rank i. The zenithal angle of each leaf, α_i, changes from 70° at leaf emergence to 10° at the end of leaf expansion (22), which is calculated using:

$$\alpha_i = \begin{cases} 10 & 1 \le i \le i_{max} \\ 10 + 60 \cdot \dfrac{i - i_{max}}{i_{current} - i_{max}} & i_{max} < i \le i_{current} \end{cases}$$

(3.2.4)

where i_{max} is the rank of the largest leaf while $i_{current}$ is the newest leaf in the growth cycle.

The biomass produced in each cycle j is next distributed to two types of sink organ (o), i.e. leaves (l) and roots (r). Each leaf is considered as an individual organ while the roots are regarded as a single entity. In GreenLab, the incremental growth of each organ depends on the trophic competition among all the organs existing at each time point. This is decided based on the demand of each organ (d_o), which is a function of sink strength (P_o) and sink variation (f_o) as follows:

$$d_o(n) = P_o f_o(n)$$

(3.2.5)

Sink strength is an organ-type specific parameter while sink variation is a beta-law (24) or bell-shaped curve that depends on the thermal time since organ emergence (or physiological age, n). This curve is normalised to its maximal value M as shown in the following:

$$f_o(n) = \frac{1}{M} \left(\frac{n + 0.5}{T_o} \right)^{a_o - 1} \left(1 - \frac{n + 0.5}{T_o} \right)^{b_o - 1}$$

(3.2.6)

where a_o and b_o are parameters for the organ-specific beta-law and T_o is the duration of organ expansion. The duration of leaf expansion was determined based on experimental observations (2). For the root system, the duration of expansion was set to be equal to the simulation period, which covered from plant emergence until the completion of flowering in the inflorescence (2).

The total sink demand (D) of each growth cycle is the sum of all the existing individual leaf (as illustrated in Fig. 2F) and root demands as follows:

$$D(j) = \sum_{i=1}^{L_j} d_i(n) + d_r(n)$$

(3.2.7)

The increase in dry biomass for each organ is therefore:

$$\Delta q_o(j,n) = \frac{d_o(n)}{D(j)} \cdot Q(j)$$

(3.2.8)

The dry weight of each organ (m_o) at the end of every cycle can then be calculated by:

$$m_o(j) = m_o(j-1) + \Delta q_o(j)$$

(3.2.9)

15

Next, the dry weight of each leaf (in g) is converted to leaf area (in m^2) for use in the next cycle using the specific leaf area (*SLA* in m^2 g^{-1}) that changes with thermal time since plant emergence, *TT* (2):

$$SLA(j) = 0.144 \exp(-0.002 TT(j)) \tag{3.2.10}$$

In the model, leaves are divided into three functional groups: (i) young leaves that function as both source and sink (from leaf emergence until the end of leaf expansion); (ii) mature leaves that function only as source (from the end of expansion till the end of its lifespan, T_s); (iii) senesced leaves that do not have any function. Only leaves in the first two groups are considered for light interception. All parameter values are listed in Table S2.

The GreenLab functional-structural model runs in steps of growth cycles. As this model is not linked to any phenology and assimilation models, input data from experimental measurement such as plant emergence and flowering time are required, while the *LUE* of each growth cycle has to be determined through model inversion as demonstrated in Christophe et al. (2). In our study, the FSPM is linked to the PTM where flowering time could be simulated based on environmental conditions and biomass for growth (*Q*) was replaced by the values derived from the CDM (see sections 3.6-3.8).

3.3   The Photothermal Model (PTM)

There are three main components in the Photothermal Model (PTM): photoperiod, thermal time, and vernalisation (Fig. S12). It calculates a modified photothermal unit (MPTU), which is a product of the main components, on an hourly basis.

The photoperiod component consists of a spline (piecewise-linear) function, with three sections divided by two critical day lengths. At day length (*DL*) below the critical short day length (*CSDL*), rate to bolting is at its minimum. As photoperiod lengthens, the rate increases linearly until the critical long day length (*CLDL*) is reached, where there is no further increase. The effect of photoperiod on rate to bolting at each time-point *t* is therefore:

$$Photoperiod(t) = \begin{cases} D_{SD} & DL(t) \le DCSL \\ D_{LD} & DL(t) \ge CLDL \\ D_{SD} + \dfrac{[DL(t) - CSDL] \cdot [D_{LD} - D_{SD}]}{CLDL - CSDL} & \text{otherwise} \end{cases} \tag{3.3.1}$$

where D_{SD} and D_{LD} are the minimum and maximum rates respectively. Here, *DL* (and thus the value of *Photoperiod*) is the same for each calendar day and is pre-determined by the sunrise and sunset times (in the 24-hour format) at the start of each day, i.e. hour zero. This photoperiodic response has been observed experimentally (5, 25-27) and could be reproduced theoretically using the clock gene circuit model (6).

Genotypes which are insensitive to photoperiod induction maintain a constant minimum rate.

The second component calculates the thermal time in degree-hour, as follows:

$$Thermal(t) = \begin{cases} P(t) \cdot [T(t) - T_b] & T(t) \geq T_b \\ 0 & \text{otherwise} \end{cases}$$

(3.3.2)

where

$$P(t) = \begin{cases} P_{day} & Sunrise \leq t \leq Sunset \\ P_{night} & \text{otherwise} \end{cases}$$

(3.3.3)

Only day temperatures above a base value, T_b, of 3 °C (28) are considered in the model. P serves as a gating function to account for the different sensitivity to day and night temperatures.

The effect of vernalisation is computed in the model by using first a beta function to represent vernalisation effectiveness (v_e) at different temperatures within a vernalising-temperature range:

$$v_e(t) = \exp(\kappa) \cdot [T(t) - T_{V\min}]^\omega \cdot [T_{V\max} - T(t)]^\xi$$

(3.3.4)

κ, ω and ξ are parameters in the beta function. The minimum ($T_{V\min}$) and maximum ($T_{V\max}$) vernalising-temperatures are fixed at -3.5 and 6 °C based on statistical analysis of laboratory and field data (5, 29).

Cumulative vernalisation hours (V_h) up to and including the hour that ends at t is computed using a summation function to determine the period of time exposed to the effective vernalising temperatures, as follows:

$$V_h(t) = \sum_{s=1}^{s=t} v_e(s) \times \Delta s$$

(3.3.5)

Δs is the time step used in the model, which is set to one hour. The extent of vernalisation, $Vern(t)$, is then determined using the following:

$$Vern(t) = \begin{cases} F_b + \dfrac{V_h(t) \cdot [1 - F_b]}{V_{sat}} & V_h(t) \leq V_{sat} \\ 1 & \text{otherwise} \end{cases}$$

(3.3.6)

F_b is a parameter representing baseline *FLC* repression. As the period of cold exposure increases, *FLC* becomes more repressed until a point when it is permanently inactivated. This is known as the saturation point, V_{sat}, which has a value of 960 hours or 40 days (29, 30). For vernalisation insensitive genotypes, *Vern*(*t*) remains at the F_b value regardless of the period of exposure to vernalising temperatures.

The MPTU is calculated by:

$$MPTU(t) = Photoperiod(t) \cdot Thermal(t) \cdot Vern(t)$$

(3.3.7)

A cumulative MPTU is determined at every time step (one hour) as follows:

$$CumMPTU(t) = \sum_{s=1}^{s=t} MPTU(t)$$

(3.3.8)

Flowering is predicted by the model when the hourly cumulative MPTU reaches a threshold value, T_h. This threshold feature is similar to that used in classical crop models. It is also consistent with molecular switch-like traits such as those shown by *FT*, *LEAFY* (*LFY*) and *APETALA1* (*AP1*) (31-33).

All the parameter values, as used in the original study (4), are listed in Table S3.

3.4   The Photoperiodism Model (PPM)

The Photoperiodism Model (PPM) consists of a system of ordinary differential equations (ODEs) representing the dynamics of core genes in the clock network and photoperiod pathway that are inter-regulated in feedback loops (Fig. S13). In Arabidopsis, the combined antagonistic and protagonistic actions of clock-regulated GI, FKF1 and CDF1 (34-37) collectively result in *CO* peaking before the end of LDs, thus promoting *FT* expression. In SDs, as *CO* mRNA peaks during the night and CO protein is unstable in the daytime, *FT* mRNA levels remain low (36, 38, 39). Such regulatory mechanism of *FT* provides a means for plants to time their life cycle according to the season. Previous work has combined the clock model with the photoperiodic regulation of flowering time by modelling the activation of *FT* through *CO* (6). In that study, a few model variants were explored by considering different mechanisms for the activation of *CO* and *FT* with or without FKF1. These model variants have managed to reproduce the different *FT* mRNA abundance profiles during LD and SD. A recent study (6, 37) has also incorporated into the model the CDF1 component and its interaction with FKF1, which are responsible for producing the first CO peak at the end of long days (34, 35). However, these models used experimental data for CDF1 and FKF1 as model inputs and thus could only describe flowering regulation at the few photoperiods where data were available. Future models that include these regulators as part of the modelled components would be required for studies under the natural range of photoperiod that varies daily. Therefore the current study utilised the simplest form of the Salazar et al. model (Model 3 that

does not include FKF1, Fig. S13). This simple model may not fully represent the exact biphasic dynamics of CO or the phase relations of all clock components; however it could provide good qualitative output of flowering regulation relative to light period (6), which was the aim of this study. The equations for this model are as follows:

$$\frac{d[\text{P}]}{dt} = 0.5 \cdot \left(1 - \Theta_{light}\right) - [\text{P}] \cdot \Theta_{light} - \frac{1.2 \cdot [\text{P}]}{1.2 + [\text{P}]} \tag{3.4.1}$$

$$\frac{d[LHY]}{dt} = \Theta_{light} \cdot v_1 \cdot [\text{P}] + \frac{v_2 \cdot [\text{X}_c]^{\mu_1}}{v_3 + [\text{X}_c]^{\mu_1}} - \frac{v_4 \cdot [LHY]}{v_5 + [LHY]} \tag{3.4.2}$$

$$\frac{d[\text{LHY}_c]}{dt} = v_6 \cdot [LHY] - v_7 \cdot [\text{LHY}_c] + v_8 \cdot [\text{LHY}_n] - \frac{v_9 \cdot [\text{LHY}_c]}{v_{10} + [\text{LHY}_c]} \tag{3.4.3}$$

$$\frac{d[\text{LHY}_n]}{dt} = v_7 \cdot [\text{LHY}_c] - v_8 \cdot [\text{LHY}_n] - \frac{v_{11} \cdot [\text{LHY}_n]}{v_{12} + [\text{LHY}_n]} \tag{3.4.4}$$

$$\frac{d[TOC1]}{dt} = \frac{v_{13} \cdot [\text{Y}_n]^{\mu_2}}{v_{14} + [\text{Y}_n]^{\mu_2}} \cdot \frac{v_{15}}{v_{16} + [\text{Y}_n]^{\mu_2}} - \frac{v_{17} \cdot [TOC1]}{v_{18} + [TOC1]} \tag{3.4.5}$$

$$\frac{d[\text{TOC1}_c]}{dt} = v_{19} \cdot [TOC1] - v_{20} \cdot [\text{TOC1}_c] + v_{21} \cdot [\text{TOC1}_n]$$
$$- \frac{\left(v_{22} \cdot \left(1 - \Theta_{light}\right) + v_{23}\right) \cdot [\text{TOC1}_c]}{v_{24} + [\text{TOC1}_c]} \tag{3.4.6}$$

$$\frac{d[\text{TOC1}_n]}{dt} = v_{20} \cdot [\text{TOC1}_c] - v_{21} \cdot [\text{TOC1}_n] - \frac{\left(v_{25} \cdot \left(1 - \Theta_{light}\right) + v_{26}\right) \cdot [\text{TOC1}_n]}{v_{27} + [\text{TOC1}_n]}$$

$$\tag{3.4.7}$$

$$\frac{d[X]}{dt} = \frac{v_{28} \cdot [\text{TOC1}_n]^{\mu_3}}{v_{29} + [\text{TOC1}_n]^{\mu_3}} - \frac{v_{30} \cdot [X]}{v_{31} + [X]} \tag{3.4.8}$$

$$\frac{d[\text{X}_c]}{dt} = v_{32} \cdot [X] - v_{33} \cdot [\text{X}_c] + v_{34} \cdot [\text{X}_n] - \frac{v_{35} \cdot [\text{X}_c]}{v_{36} + [\text{X}_c]} \tag{3.4.9}$$

$$\frac{d[\mathrm{X}_n]}{dt} = v_{33} \cdot [\mathrm{X}_c] - v_{34} \cdot [\mathrm{X}_n] - \frac{v_{37} \cdot [\mathrm{X}_n]}{v_{38} + [\mathrm{X}_n]} \tag{3.4.10}$$

$$\frac{d[Y]}{dt} = \left(\Theta_{light} \cdot v_{53} \cdot [\mathrm{P}] + \frac{\Theta_{light} \cdot v_{39} + v_{40}}{v_{41} + [\mathrm{TOCl}_n]^{\mu_4}} \right) \cdot \frac{v_{51}}{[\mathrm{LHY}_n]^{\mu_5} + v_{52}} - \frac{v_{42} \cdot [Y]}{v_{43} + [Y]} \tag{3.4.11}$$

$$\frac{d[\mathrm{Y}_c]}{dt} = v_{44} \cdot [Y] - v_{45} \cdot [\mathrm{Y}_c] + v_{46} \cdot [\mathrm{Y}_n] - \frac{v_{47} \cdot [\mathrm{Y}_c]}{v_{48} + [\mathrm{Y}_c]} \tag{3.4.12}$$

$$\frac{d[\mathrm{Y}_n]}{dt} = v_{45} \cdot [\mathrm{Y}_c] - v_{46} \cdot [\mathrm{Y}_n] - \frac{v_{49} \cdot [\mathrm{Y}_n]}{v_{50} + [\mathrm{Y}_n]} \tag{3.4.13}$$

Assuming that *CO* mRNA is the same as TOC1 nuclear protein, and *FT* mRNA is activated by CO protein:

$$\frac{d[\mathrm{CO}]}{dt} = v_{COm} \cdot [\mathrm{TOC1}_n] - \frac{(1 - \Theta_{light}) \cdot v_{COp} \cdot [\mathrm{CO}]}{k_{COp} + [\mathrm{CO}]} \tag{3.4.14}$$

$$\frac{d[FT]}{dt} = B_{CO} + \frac{V_{CO} \cdot [\mathrm{CO}]}{K_{CO} + [\mathrm{CO}]} - \frac{v_{FT} \cdot [FT]}{k_{FT} + [FT]} \tag{3.4.15}$$

$$\Theta_{light} = \frac{1}{4} \cdot (1 + \tanh(6 \cdot (t - sunrise))) \cdot (1 - \tanh(6 \cdot (t - sunset))) \tag{3.4.16}$$

The concentration (indicated by [ ]) of each component is written in italics to represent the corresponding mRNA, or in non-italics to represent the protein form. The subscripts *c* (cytoplasm) and *n* (nucleus) indicate the corresponding compartment. All the parameter values are listed in Table S4.

3.5 Standardisation of units

All the models utilised different time steps in the original work. The PTM runs on an hourly basis while the ordinary differential equations (ODEs) in the PPM are solved at variable time steps of seconds. For the CDM, a time step of 6 seconds was utilised in the original work (1). However, increasing this to one hour resulted in only a difference of 6.5% compared to its originally published final output of rosette area. Therefore an hourly step as used in the PTM was adopted except for the PPM, which

is still solved using a variable time step. As both the CDM and the FSPM were also modified to run every one hour, equations 3.2.1 to 3.2.10 were re-annotated by replacing growth cycle j with plant age t in hours. The thermal time unit (in degree days) for every hour can thus be calculated as:

$$Thermaltime(t) = \frac{T(t) - T_b}{24}$$

(3.5.1)

where the numerator is the temperature above a base value of T_b = 3 °C (as used previously in equation 3.3.2).

The different biomass units used in the CDM and the FSPM were maintained to ensure compatibility with measurement and optimisation carried out in the original studies (1, 2). However, unit conversion factors based on carbon content were introduced in the current work so that biomass information could be sent backwards and forwards between the two sub-models. For simplicity, an average value of 0.3398 g carbon per g leaf dry mass was used based on the reported total leaf carbon content per dry weight minus sugar and starch contents (40, 41). Root carbon content was fixed at 0.35 g per g dry weight (42, 43).

3.6 Connecting the Carbon Dynamic and Functional-Structural Models

In order to connect the CDM and the FSPM, a few connection points had to be identified and redundant equations were replaced or abolished. One redundant equation is 3.2.2 from the FSPM which determines the biomass production Q (source). This component was replaced with the CDM in the current study (Fig. S14). Both models are executed simultaneously with information passed back and forth in between. This combined model forms the vegetative growth component.

At each time step, the number of leaves (equation 3.2.1) is first calculated so the dry biomass sink demand for each of the organs can be computed (equation 3.2.5) and converted to carbon sink demand using the unit conversion factors. These values are then used to determine the carbon root-to-shoot allocation ratio (RS) by summing all the leaf demands as follows:

$$RS(t) = \frac{d_r(n)}{\sum_{i=1}^{L(t)} d_i(n)} \cdot \frac{\text{Root carbon content}}{\text{Leaf carbon content}}$$

(3.6.1)

where t is the plant age (or hours after sowing), n is the physiological age of each organ (thermal time after organ emergence) and $L(t)$ is the number of existing leaves at time t. The rosette area for light interception (equations 3.2.3 and 3.2.4) is also calculated and passed to the CDM (together with environmental input such as light intensity, temperature and CO_2 level) for determination of photosynthesis, respiration and carbon flow. The growth pool in the CDM (Figs. S10 and S14) is then allocated to the roots and shoots using the RS ratio in equation 3.6.1, which replaced the RS ratio (as a function of rosette area, equation 3.1.29) that was originally used (1). Once the carbon has been allocated, contents in the roots and shoots are re-converted to dry

mass unit. Distribution of the incremental shoot dry mass to different leaves is later determined using a modified version of equation 3.2.8 that was adjusted to represent trophic competition among the leaves, as shown in the following:

$$\Delta q_i(t,n) = \frac{d_i(n)}{\sum\limits_{i=1}^{L(t)} d_i(n)} \cdot \text{Total increment in shoot dry mass}\,(t) \qquad (3.6.2)$$

The new dry biomass for each leaf can thus be calculated using equation 3.2.9 and the individual leaf area determined from the specific leaf area, *SLA* (equation 3.2.10). As *SLA* decreases with thermal time, a constraint is applied to ensure that leaf area does not shrink in later time points, as follows:

$$S_i(t) = \text{Max}\{SLA(t) \cdot m_i(t), S_i(t-1)\} \qquad (3.6.3)$$

3.7 Linking the Photoperiodism and Photothermal Models

To link both models, the spline function in the photoperiod component of the PTM was replaced by the following calibration equation (equation 3.7.1). This function relates the area under the curve (integral over 24 hours) of *FT* mRNA produced by the PPM for each photoperiod, to the associated parameter value in the photoperiod component (equation 3.3.1). A sigmoid function as described in Salazar et al. (6) was used, where developmental rate reached its limits at both photoperiod extremes after certain critical day lengths (5, 44, 45).

$$Photoperiod = A + B\left[\frac{C^n}{C^n + (FTarea)^n}\right] \qquad (3.7.1)$$

All the constant values are listed in Table S5. These values were determined by entraining the PPM to different constant photoperiods. During each model entrainment, a unique solution was reached for every photoperiod where the concentration level of each component achieved a stable cycle of 24-hour period (Figs. S15a and S15b). In the natural environment, day length changes every day, and the concentration level of each component is carried forward to the following day. Therefore, for simulation of flowering time using the combined model, only the first day was entrained. The concentration levels of all components at the end of each day were then used as the initial values in the ODE simulation for the next day.

One feature of the original PTM is that the photoperiod component for every hour on the same day shares the same value, because day length is determined by the sunrise and sunset times of the associated day. This indirectly assigns the notion that plants know in advance the time of sunrise and sunset before the instant. While the clock may allow plants to anticipate dawn and dusk, the relevant molecular events are dynamically modulated by on-going environmental stimuli, which may either cause immediate responses or be stored as "memory" that affects later processes (46, 47).

An immediate response in this case would mean that photoperiodic induction of flowering changes depending on the hourly level of *FT*. We adopted the memory concept in this study, where the total amount (integral or area under the curve) of *FT* accumulated over the previous 24 hours from the hour at hand is considered. This is consistent with the observation that a single long day is insufficient for photoperiodic induction in young Arabidopsis, suggesting that information accumulates over several days (48). It could also embody the lag time required to transport FT protein from the leaves to the meristem (33). Our calibrated model could reproduce all the predicted values from the original non-calibrated model.

3.8   Framework Model

The combined PPM-PTM only predicts the flowering time of Arabidopsis without giving any information on plant size and biomass. On the other hand, the vegetative growth component in Section 3.6 has unlimited run-time as long as input data are provided, unless execution is terminated. In other words, the simulation process could run incessantly, though simulated rosette growth may eventually end when the carbon budget runs out due to self-shading, leaf senescence, and/or suboptimal growth conditions. The current study therefore connected both the PPM-PTM and vegetative growth models, with the former providing the time of flowering to the latter so that rosette growth can be terminated accordingly (Fig. S16). Four, hourly input data, i.e. temperature, light intensity, CO_2 level and sunset/sunrise time (which are also used for day length calculation), are necessary to run the model. These input data are sent to different component(s) in the model based on the specific requirement(s). The PPM-PTM starts execution instantaneously without any conditions, whereas the vegetative growth compartment has a conditional feature. The latter is only activated when the accumulated thermal time exceeds the threshold for seedling emergence. Once activated, this compartment begins its implementation using initial values from model initialisation (see Section 3.9). The circuit runs continuously in a loop until flowering time, as informed by the PPM-PTM, is reached. At this point, our Framework Model stops execution. Outputs from the Framework Model include time series data of sugar and starch contents, and the biomass and size of the whole plant and individual organs (Fig. 2).

3.9   Model initialisation

Neither the CDM nor the FSPM consider events or processes before plant emergence. In our study, the vegetative growth component remains inactive until plant emergence, TT_0. It is marked as the time point corresponding to 110 °Cd, which was the accumulated thermal time when the cotyledons were fully opened (Stage 1) (20). The thermal time post-plant emergence, *TT*, is therefore given by:

$$TT(t) = Thermaltime(t) - TT_0$$

(3.9.1)

As soon as TT_0 is reached, the vegetative growth component is triggered. Initial plant biomass is assumed to be the same as seed input (Table S2) (2), which is distributed to the sink organs, i.e. cotyledons, roots and hypocotyl, based on a constant fraction.

This fraction is determined by the sink demand function (equations 3.2.7 and 3.2.8), assuming that the period of root expansion begins on the day of sowing and the above-ground sink strength is similar to that of rosette leaves ($P_{above\text{-}ground}$ = 1) and at its maximum ($f_{above\text{-}ground}$ = 1). Above-ground biomass is shared between the cotyledons and hypocotyl by setting the initial cotyledon area as 1 mm^2 and the specific cotyledon area as the *SLA* at plant emergence. These give:

Hypocotyl biomass = Seed biomass – Root biomass – 2[Cotyledon area][$SLA_{TT=0}$]
(3.9.2)

The biomass of hypocotyl is set as a constant thereafter and the secondary growth of stem is assumed to be negligible. Therefore, any subsequent allocation to the shoot is distributed to leaves only.

Using the calculated biomass at emergence, the initial shoot and root carbon contents can next be generated with the biomass unit conversion factors. Initial sugar carbon and starch carbon are also determined as:

Sugar_carbon(emergence) = 2[Cotyledon area][Sugar content per unit area]
(3.9.3)

Starch_carbon(emergence) = [Initial starch:sugar ratio][Sugar_carbon(emergence)]
(3.9.4)

These initial ratio and specific sugar content are the same as those used during the initialisation of CDM (1). In the case where emergence time is modelled to occur in the middle of the night, the end-of-day starch level used to calculate starch degradation rate (equation 3.1.15) is determined pro rata based on the number of night hours that has passed out of the night length. This was included because starch degradation rate is regulated by the circadian clock, which is already functional at the germination stage (16, 49). All the parameter values required for the integration and initialisation of the Framework Model are listed in Table S6.

3.10   Parameter values

We maintained all the parameter values used in the original models, except for the following:

a) J_{max}:V_{cmax} ratio (or R_{JV} in equation 3.1.8) in the CDM. The reported values of this ratio in the literatures suggested a photoperiod-dependent tendency (Table S7), with an increasing ratio as photoperiod shortened. The ratio used in the original model (1) was 2.1, where the model was tested on plants grown under an 8h photoperiod. As the Col data were collected from plants grown under a 12h photoperiod, the current study adopted the J_{max}:V_{cmax} ratio measured under the same photoperiod, which corresponded to 1.7 (50);

b) the MPTU threshold (T_h) in the PTM. In the original model, this flowering threshold was determined by taking the mean of the MPTU totals for the set of line x field plantings that gave the minimum coefficient of variation (4, 5). Therefore, there

24

can be slight deviation between the observed and predicted flowering time. As the plants in our experiments flowered slightly later than the predicted time, flowering threshold was adjusted to extend the simulation of the plants' exponential growth until the last biomass assay (Table S3);

c) The duration of root system expansion (Table S2). As different accessions have different lifespans and rates of development (51), instead of using a fixed duration as in the GreenLab model, we scaled the duration relative to the flowering time of each accession. It was calculated from the original duration (2) as 1.3 fold the cumulative thermal time at the end of vegetative growth;

d) Leaf appearance rate for Ler and Fei. Consistent with its characteristic as a fast-growing ecotype, Fei was found to reach plant emergence early. This was incorporated by modifying the thermal time at plant emergence to that observed for Fei, which was enough for the model to reproduce its leaf appearance rate (Fig. S2). Ler on the other hand emerged at the expected time, but its leaf appearance rate at the adult stage was slightly lower than the rate derived from Col data that was used in the GreenLab functional-structural model (2). Therefore, the leaf appearance rate for Ler was also modified in our model to reflect this observation (Table S2).

3.11   Extension of the Framework Model to different photoperiods

To test the ability of the FM in simulating growth under photoperiods other than the 12h reference case study, we used measured data from the literature as benchmark to modify up to three photoperiod-dependent parameters in the model, as follows:

a) J_{max}:V_{cmax} ratio (or R_{JV} in equation 3.1.8). By assuming maximum and minimum limits (Table S7), we used a value of 2.1 for photoperiods of 8 hours and below, and 1.4 for 18h photoperiod;

b) Specific leaf area (SLA in equation 3.2.10). It was reduced by a factor of 0.88 for 18h photoperiod to increase leaf thickness (52);

c) Starch turnover (ST_c in equation 3.1.15). We reduced this to 0.6 for 18h photoperiod based on starch contents measured at dawn and dusk (52).

Together with the modifications above, we used as model inputs growth conditions such as photoperiod, light intensity, atmospheric CO_2 concentration and day/night temperatures similar to the experiments (52) to simulate plant growth. Results are shown in Fig. 4 in the main text. An alternative simulation was also conducted for 18h photoperiod by only modifying starch turnover and extrapolating J_{max}:V_{cmax} ratio to a value of 1.2 (Fig. S5b).

3.12   Introduction of stochasticity into the Framework Model

Stochasticity was introduced at the organ level in our model based on a concept adopted in branching systems (53, 54). In branching systems, each phytomer consists of a leaf, an internode (stem) and a bud, which could be apical and/or lateral. In the

stochastic version of the GreenLab functional-structural model (53, 54), a bud may or may not form a new phytomer at each growth cycle, causing the branching system of a plant to be non-symmetrical. The probability of forming a new phytomer is assumed to be constant and independent of the rate of phytomer appearance. This event is analogous to a binomial distribution, which is the distribution of the number of successes X in N number of trials, given that the probability of success p is independent of each trial. The mean and variance of X can be estimated as:

$$\text{Mean}(X) = Np \tag{3.12.1}$$

$$\text{Variance}(X) = Np(1\text{-}p) \tag{3.12.2}$$

Combining both equations gives:

$$\text{Variance}(X) = (1\text{-}p){\cdot}\text{Mean}(X) \tag{3.12.3}$$

Therefore, the value of p can be determined from the slope of variance-to-mean regression.

This concept was applied to our Arabidopsis Framework Model, assuming that there is a pre-determined number of growth cycles (N) from germination until flowering, with the amount of observed phytomers (determined from leaf count) representing the number of successes X. During each growth cycle, there is a p probability that a new leaf appears and this probability is independent of the previous cycle or the period between successive cycles (phyllochron). Therefore p has a constant value and is similar for both the early and late vegetative stages. Using our leaf number data at the early vegetative stage, we plotted the variance of phytomer number, i.e. leaf number minus one (equation 3.2.1), against the mean (Fig. S6), and p was found to be 0.97 for both Ler and Fei.

In the deterministic model, the number of growth cycles ($N_{\text{deterministic}}$) is the same as the mean number of phytomers that appear. In the stochastic model, however, the number of growth cycles ($N_{\text{stochastic}}$) would be larger and can be determined once p is known (equation 3.12.1). The relationships between variables in the deterministic and stochastic versions are therefore as follows:

$$N_{\text{deterministic}} = N_{\text{stochastic}}p \tag{3.12.4}$$

$$\text{Phyllochron}_{\text{deterministic}}{\cdot}N_{\text{deterministic}} = \text{Phyllochron}_{\text{stochastic}}{\cdot}N_{\text{stochastic}} \tag{3.12.5}$$

Combining the two equations above therefore gives:

$$\text{Phyllochron}_{\text{stochastic}} = \text{Phyllochron}_{\text{deterministic}}{\cdot}p \tag{3.12.6}$$

The new phyllochron values were incorporated into the model, and a binomial sampling function was inserted at the beginning of each growth cycle, so that a leaf only appears 97 % of the time.

The method described above is of a discrete nature. In Arabidopsis, leaf is initiated at the meristem at a regular interval but a slight variability in this interval has been observed (55). Therefore, the variation in leaf number is probably due to the variability in the interval rather than a discrete skipping of growth cycles, as assumed in the binomial distribution. Nevertheless, our stochastic model managed to explain the variance in leaf number from the experiments (Table S9), including the variance in the late vegetative stage when the phyllochron is shorter. This suggested that our discrete method can be used to describe the continuous process, and supported our assumption of *p* being independent of the phyllochron.

4 Model analysis: parameter sensitivity

The behaviour of the Framework Model was tested by standard parameter sensitivity analysis, to determine which modelled processes most affected the model's high-level outputs, namely total biomass and flowering time.

4.1 Analytical procedure

Each parameter value was increased or decreased by 5%, while holding other parameters constant. The input growth conditions for simulation were identical to the reference simulation in Figs. 2I-K. For each perturbation, we compared the flowering time and the fresh biomass at two timepoints to the values for the reference, unperturbed parameter set (Table S11; Figs. S8 - S9). Biomass at flowering showed the effect of each parameter on maximal growth; biomass at 876h (36.5 DAS) showed the effects on vegetative growth rate independent of flowering time.

4.2 Analysis results

In general, perturbations that resulted in an increase (or decrease) in flowering time always displayed an increase (decrease) in biomass at flowering, but not *vice versa*. This was expected, as flowering time determines the duration of vegetative biomass accumulation in the rosette. A 5% parameter change can potentially alter biomass and flowering very substantially, because both outputs integrate the small change in a particular process over many days, greatly amplifying any small effect on the instantaneous growth or developmental rate.

Water content (w) strongly affected fresh biomass (Fig. S9) but our 5% perturbations are outside the limited range of water content reported for many Arabidopsis genotypes (see Discussion; (55)). The next two most sensitive parameters in the model (p51 and p52; see Table S11 for parameter names) doubled the biomass at flowering (Fig. S8), and are associated with the temperature effect on the potential rate of photosynthetic electron transport. Other parameters that affected biomass included leaf carbon content (p30), maximum carboxylation rate (p38), a parameter that relates CO_2 compensation point to temperature (p46), intercellular CO_2 fraction (p49), J_{max}:V_{cmax} ratio (p53), curvature of the electron transport response to irradiance (p55), maximum daily relative growth rate (p63), and parameters associated with specific leaf area (p71 and p72). These are parameters that were either measured or

calibrated in previous independent studies (see Section 3.1, above). RuBisCo-related parameters (p40 – p45) did not display any sensitivity, because the reference condition was light-limiting.

The majority of parameters that displayed strong effects on simulated flowering time came from the PPM (Fig. S9), except for baseline FLC repression (p15), MPTU threshold (p17) and the base temperature used for thermal-time calculation (p1). These parameters were determined in earlier studies through optimisation (1, 6), with the exception of base temperature that was estimated through extrapolation of experimental data (56). Among the sensitive parameters in the PPM were those associated with *TOC1*, *CO* and *FT* (equations 3.4.5-3.4.7, 3.4.14, 3.4.15). This is not surprising, as *TOC1* is the clock component directly upstream of *CO* in the PPM, while *CO* regulates *FT* expression that is linked to the PTM. Accordingly, biomass was increased/decreased with the simulated flowering time. No sensitivity was detected for vernalisation-related parameters (p2 – p7) in the PTM under our non-vernalising input conditions.

In terms of linearity, only the MPTU threshold (p17) caused identical changes in flowering time under both positive and negative perturbations, though this only holds under constant input conditions. However, the resulting effects on biomass were not linear; the absolute increase in biomass was larger than the decrease. This can be explained by the exponential growth behaviour in plants, so a longer vegetative period can result in much higher biomass. We also analysed the biomass simulated at 876h, a time point coinciding with the earliest flowering time among all the perturbations (Table S11; Figs. S8 - S9), to distinguish between the effects of flowering time, growth-related events or both on the final biomass. As expected, no difference in biomass at the early time point was seen when parameters in the PPM were perturbed. Only one parameter (base temperature, p1) affected both vegetative period and growth. In the FM, this parameter is used to calculate thermal-time. Thermal time determines the developmental rate to emergence, to leaf appearance that affects source-sink relations, and to flowering.

5 References

1.   Rasse DP & Tocquin P (2006) Leaf carbohydrate controls over Arabidopsis growth and response to elevated CO2: an experimentally based model. *New Phytol* 172(3):500-513.
2.   Christophe A, *et al.* (2008) A model-based analysis of the dynamics of carbon balance at the whole-plant level in Arabidopsis thaliana. *Funct Plant Biol* 35(11):1147-1162.
3.   Yan HP, Barczi JF, De Reffye P, & Hu BG (2002) Fast algorithms of plant computation based on substructure instances'. *Wscg'2002 Short Communication Papers, Conference Proceedings*:145-152.
4.   Chew YH, *et al.* (2012) An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytol* 194(3):654-665.
5.   Wilczek AM, *et al.* (2009) Effects of Genetic Perturbation on Seasonal Life History Plasticity. *Science* 323(5916):930-934.
6.   Salazar JD, *et al.* (2009) Prediction of Photoperiodic Regulators from Quantitative Gene Circuit Models. *Cell* 139(6):1170-1179.
7.   Farquhar GD, Caemmerer SV, & Berry JA (1980) A Biochemical-Model of Photosynthetic Co2 Assimilation in Leaves of C-3 Species. *Planta* 149(1):78-90.
8.   Farquhar GD, von Caemmerer S, & Berry JA (2001) Models of photosynthesis. *Plant Physiol* 125(1):42-45.
9.   Sun JD, Okita TW, & Edwards GE (1999) Modification of carbon partitioning, photosynthetic capacity, and O-2 sensitivity in arabidopsis plants with low ADP-glucose pyrophosphorylase activity. *Plant Physiol* 119(1):267-276.
10.   Dewar RC, Medlyn BE, & McMurtrie RE (1998) A mechanistic analysis of light and carbon use efficiencies. *Plant Cell Environ* 21(6):573-588.
11.   Ogren E (2000) Maintenance respiration correlates with sugar but not nitrogen concentration in dormant plants. *Physiol Plant* 108(3):295-299.
12.   Stitt M (1996) in *Photosynthesis and the Environment*, ed Baker NR (Kluwer Academic, Dordrecht, The Netherlands), pp 635 - 640.
13.   Eichelmann H & Laisk A (1994) CO_2 Uptake and Electron-Transport Rates in Wild-Type and a Starchless Mutant of Nicotiana-Sylvestris - the Role and Regulation of Starch Synthesis at Saturating Co2 Concentrations. *Plant Physiol* 106(2):679-687.
14.   Gibon Y, *et al.* (2004) Adjustment of diurnal starch turnover to short days: depletion of sugar during the night leads to a temporary inhibition of carbohydrate utilization, accumulation of sugars and post-translational activation of ADP-glucose pyrophosphorylase in the following light period. *Plant J* 39(6):847-862.
15.   Gibon Y, *et al.* (2009) Adjustment of growth, starch turnover, protein content and central metabolism to a decrease of the carbon supply when Arabidopsis is grown in very short photoperiods. *Plant Cell Environ* 32(7):859-874.
16.   Graf A, Schlereth A, Stitt M, & Smith AM (2010) Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proc Natl Acad Sci U S A* 107(20):9458-9463.

17.   Lu Y, Gehan JP, & Sharkey TD (2005) Daylength and circadian effects on starch degradation and maltose metabolism. *Plant Physiol* 138(4):2280-2291.

18.   Thimm O*, et al.* (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J* 37(6):914-939.

19.   Caspar T, Huber SC, & Somerville C (1985) Alterations in Growth, Photosynthesis, and Respiration in a Starchless Mutant of Arabidopsis-Thaliana (L) Deficient in Chloroplast Phosphoglucomutase Activity. *Plant Physiol* 79(1):11-17.

20.   Boyes DC*, et al.* (2001) Growth stage-based phenotypic analysis of arabidopsis: A model for high throughput functional genomics in plants. *Plant Cell* 13(7):1499-1510.

21.   Medford JI, Behringer FJ, Callos JD, & Feldmann KA (1992) Normal and Abnormal-Development in the Arabidopsis Vegetative Shoot Apex. *Plant Cell* 4(6):631-643.

22.   Chenu K*, et al.* (2005) Integrated responses of rosette organogenesis, morphogenesis and architecture to reduced incident light in Arabidopsis thaliana results in higher efficiency of light interception. *Funct Plant Biol* 32(12):1123-1134.

23.   Mundermann L, Erasmus Y, Lane B, Coen E, & Prusinkiewicz P (2005) Quantitative modeling of Arabidopsis development. *Plant Physiol* 139(2):960-968.

24.   Yin XY, Goudriaan J, Lantinga EA, Vos J, & Spiertz HJ (2003) A flexible sigmoid function of determinate growth. *Annals of Botany* 91(3):361-371.

25.   Pouteau S & Albertini C (2009) The significance of bolting and floral transitions as indicators of reproductive phase change in Arabidopsis. *J Exp Bot* 60(12):3367-3377.

26.   Pouteau S*, et al.* (2008) Diversification of Photoperiodic Response Patterns in a Collection of Early-Flowering Mutants of Arabidopsis. *Plant Physiol* 148(3):1465-1473.

27.   Pouteau S, Ferret V, & Lefebvre D (2006) Comparison of environmental and mutational variation in flowering time in Arabidopsis. *J Exp Bot* 57(15):4099-4109.

28.   Granier C*, et al.* (2002) Individual leaf development in Arabidopsis thaliana: a stable thermal-time-based programme. *Ann Bot* 89(5):595-604.

29.   Napp-Zinn K (1957) Untersuchungen iiber das Vernalisationsverhalten einer winterannuellen Rasse von Arabidopsis thaliana. *Planta* 50:177 - 210.

30.   Lee I & Amasino R (1995) Effect of Vernalization, Photoperiod, and Light Quality on the Flowering Phenotype of Arabidopsis Plants Containing the FRIGIDA Gene. *Plant Physiol* 108(1):157 - 162.

31.   Jack T (2004) Molecular and genetic mechanisms of floral control. *Plant Cell* 16:S1-S17.

32.   Sablowski R (2007) Flowering and determinacy in Arabidopsis. *J Exp Bot* 58(5):899-907.

33.   Corbesier L*, et al.* (2007) FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* 316(5827):1030 - 1033.

34.   Imaizumi T, Schultz TF, Harmon FG, Ho LA, & Kay SA (2005) FKF1F-BOX protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science* 309(5732):293-297.

35. Fornara F*, et al.* (2009) Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response. *Dev Cell* 17(1):75-86.

36. Yanovsky MJ & Kay SA (2002) Molecular basis of seasonal time measurement in Arabidopsis. *Nature* 419(6904):308-312.

37. Song YH, Smith RW, To BJ, Millar AJ, & Imaizumi T (2012) FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science* 336:1045 - 1049.

38. Valverde F*, et al.* (2004) Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303(5660):1003-1006.

39. Corbesier L*, et al.* (2007) FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* 316(5827):1030-1033.

40. Gorsuch PA, Pandey S, & Atkin OK (2010) Temporal heterogeneity of cold acclimation phenotypes in Arabidopsis leaves. *Plant Cell Environ* 33(2):244-258.

41. Gorsuch PA, Pandey S, & Atkin OK (2010) Thermal de-acclimation: how permanent are leaf phenotypes when cold-acclimated plants experience warming? *Plant Cell Environ* 33(7):1124-1137.

42. Kumar S, Udawatta RP, & Anderson SH (2010) Root length density and carbon content of agroforestry and grass buffers under grazed pasture systems in a Hapludalf. *Agrofor Syst* 80(1):85-96.

43. Prendergast-Miller M & Sohi SP (2010) Investigating biochar impacts on plant roots and root carbon. in *Soil Organic Matter Conference* (Cote d'Azur, France).

44. Welch SM*, et al.* (2005) Merging genomic control networks and soil-plant-atmosphere-continuum models. *Agric Sys* 86(3):243-274.

45. Welch SM, Dong ZS, Roe JL, & Das S (2005) Flowering time control: gene network modelling and the link to quantitative genetics. *Aust J Agric Res* 56(9):919-936.

46. Tafforeau M, Verdus MC, Norris V, Ripoll C, & Thellier M (2006) Memory Processes in the Response of Plants to Environmental Signals. *Plant Signaling and Behavior* 1(1):9 - 14.

47. Verdus MC, Ripoll C, Norris V, & Thellier M (2012) The Role of Calcium in the Recall of Stored Morphogenetic Information by Plants. *Acta Biotheor* 60(1-2):83-97.

48. Corbesier L, Gadisseur I, Silvestre G, Jacqmard A, & Bernier G (1996) Design in Arabidopsis thaliana of a synchronous system of floral induction by one long day. *Plant J* 9(6):947-952.

49. Penfield S & Hall A (2009) A Role for Multiple Circadian Clock Genes in the Response to Signals That Break Seed Dormancy in Arabidopsis. *Plant Cell* 21(6):1722-1732.

50. Flexas J*, et al.* (2007) Mesophyll conductance to CO2 in Arabidopsis thaliana. *New Phytol* 175(3):501-511.

51. Granier C*, et al.* (2006) PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytol* 169(3):623-635.

52. Sulpice R*, et al.* (2014) Arabidopsis coordinates the diurnal regulation of carbon allocation and growth across a wide range of photoperiods. *Mol Plant* 7(1):137-155.

53.   de Reffye P, Cognée M, Jaeger M, & Traoré B (1998) Stochastic modelling for cotton growth and architecture I. Main stems and primary fruiting branches. *Coton Fibres Trop* 43:282-290.
54.   Wang F*, et al.* (2011) A stochastic model of tree architecture and biomass partitioning: application to Mongolian Scots pines. *Ann Bot* 107(5):781-792.
55.   Guenot B*, et al.* (2012) PIN1-Independent Leaf Initiation in Arabidopsis. *Plant Physiol* 159(4):1501-1510.
56.   Schneider CA, Rasband WS, & Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Meth* 9(7):671-675.

Supplementary figure legends

Figure S1: Model simulation using the original parameter values, with the $J_{max}:V_{cmax}$ ratio (or R_{JV}) taking the value of 2.1. Simulations (solid lines) and experimental data of Col wild type (symbols) at 18, 25, 27 and 38 days after sowing (DAS) for total shoot (a) and individual leaf biomass (b) are illustrated. Leaves are ranked according to the order of appearance. Experimental conditions: ~21.3 ^{0}C; 12:12 light/dark cycle; light intensity = 110 μmol m^{-2} s^{-1}; Average daytime CO_2 level = 375 ppm.

Figure S2: Shoot growth of Ler and Fei until the appearance of flower buds. (a) Leaf number plotted against days after sowing. Only leaves exceeding 1 mm^2 in size (estimated with the naked eye) were considered in the total leaf count (including the cotyledons); (b) Phytomer number, i.e. leaf number minus one, plotted as a function of thermal time after plant emergence. In the plot for Fei, the thermal time at plant emergence has been modified to the value observed in our study (Table S6). Also shown are phytomer numbers as modelled for Col data in the GreenLab functional-structural model, and the modified rate for Ler (Table S6). Error bars showing the standard error of 24 plants are very small and therefore hidden by the data markers.

Figure S3: Simulation for Col (a,b) and Fei (c,d) using a standardised water content of 91 %.

Figure S4: Experimental setup for gas exchange measurement

Figure S5: Experimental data (black and white) and model simulations (light and dark green) for sucrose level (a) and rosette fresh weight (b) at the end of day (ED; white and light green) and end of night (EN; black and dark green), for plants grown under the photoperiod conditions detailed in Fig. 4. Symbols for ED and EN simulation results overlap in panel (a). The simulation for 18h photoperiod in panel (b) used a $J_{max}:V_{cmax}$ value of 1.2, which was extrapolated from the 14h value reported in the literature, and a starch breakdown level reduced to 60% based on the data. Parameter modifications for other photoperiods were similar to those used in Fig. 4. Error bars show the standard deviation of 5 plants.

Figure S6: Plot of variance against mean phytomer number using leaf number data for both Ler and Fei at the early vegetative stage. The slopes of both regressions were close to 0.03, giving a binomial probability of 0.97.

Figure S7: (a) Schematic diagram showing how the timing of phase transition was varied for each simulation, which produced different final leaf numbers at flowering. Dashed lines of similar colours represent the controls that produce the same final number of leaves but at a single rate. Model simulations (solid lines) and experimental data (symbols) of rosette leaf number (b) and individual leaf biomass at 37 DAS (c) for Col and *Pro35S:MIR156* using a piecewise-linear leaf production model. For *Pro35S:MIR156*, the phyllochrons at the early and late vegetative stages were determined with linear regression to be 18.9 °Cd and 7.3 °Cd, respectively. Note that the simulation lines in b are not exactly linear due to the coarse time-step, i.e. increment in degree-days, relative to the phyllochrons.

Figure S8: Sensitivities of parameters (excluding the PPM and water content) to single perturbations of 5 % increase (+) or 5 % decrease (-) in each parameter value. Sensitivities were scored relative to the reference values, (top panel) for flowering time (right axis) and fresh biomass at flowering (left axis), and (bottom panel) for fresh biomass at hour 876. The symbol for each parameter corresponds to the list in Table S11.

Figure S9: Sensitivities of parameters in the PPM and water content to perturbations of 5 % increase (+) and 5 % decrease (-). Sensitivities were scored relative to the reference for flowering time and fresh biomass at flowering (top panel), and for fresh biomass at hour 876 (bottom panel). The symbol for each parameter corresponds to the list in Table S11.

Figure S10: Schematic diagram of the Carbon Dynamic Model (CDM). Carbon is transported (solid arrows) in and/or out of four different pools (grey boxes) through various processes. Dashed arrows indicate information or feedback input.

Figure S11: Schematic diagram of the Arabidopsis Functional-Structural Plant Model (FSPM). Biomass Q is distributed (solid arrows) to different organs (grey boxes) based on their physiological age. The above-ground structure, which determines the area for light interception (S_{ros}), is also considered in the model. Dashed arrows indicate information or feedback input.

Figure S12: The three components of the Photothermal Model (PTM). (a) Photoperiod component, which is divided into three sections by two critical day lengths; (b) Thermal time component, which only considers hourly temperatures above 3 ^{0}C. A filter function is used to account for the different sensitivity to day and night temperatures; (c) Vernalisation component, which has two sub-components. The modifier represents the extent of vernalisation and depends on the accumulated vernalisation effectiveness at different temperatures; (d) The Modified Photothermal Unit (MPTU) is a product of the three components. The MPTU is accumulated every hour until a threshold value is reached to indicate the reproductive switch.

Figure S13: Schematic diagram of the Photoperiodism Model (PPM).

Figure S14: Schematic diagram illustrating the combined CDM and FSPM, which form the vegetative growth component. It is a combination of Figures S10 and S11. Texts in red highlight the new features that were introduced in the current work. The red double-sided arrows indicate the connection points where biomass information between the two models is sent back and forth.

Figure S15: Calibrating the PPM to the photoperiod component of the PTM. The time-series of *FT* mRNA simulated by the PPM for: (a) 8h photoperiod and; (b) 16h photoperiod; (c) A plot of the area under the curve of *FT* mRNA time series at different photoperiods. Letters in brackets link to the associated time-series. These areas were translated into the photoperiod component using the sigmoid function in equation 3.7.1; (d) Comparison between the broken-stick component in the PTM (dashed line, equation 3.3.1) and the calibrated photoperiod component (full line).

Figure S16: Schematic diagram of the multi-scale Framework Model. The four sub-models are represented by the grey blocks. Line arrows: circuit flow; Dashed arrows: information flow; Bold arrows: values being supplied by one component to another.

Figure S17: An example of CO_2 measurement. Shown here are the data for a tray of 5-week-old Ler, where the average daytime CO_2 concentration in the growth chamber was 405 ppm.

Table S1: Parameter values for the Carbon Dynamic Model (CDM)

Symbol	Parameter value	Unit	Parameter description
R	8.314	J K^{-1} mol^{-1}	Gas constant
$K_c(25^0C)$	40.4	Pa	Michaelis constant of carboxylation at 25^0 C
H_c	59400	J mol^{-1}	Activation energy for carboxylation constant
$K_o(25^0C)$	24800	Pa	Michaelis constant of oxygenation at 25^0 C
H_o	36000	J mol^{-1}	Activation energy for oxygenation constant
$V_{c_{max}}(25^0C)$	29.6875	μmol m^{-2} s^{-1}	Maximum carboxylation rate at 25^0 C
H_V	64800	J mol^{-1}	Activation energy for maximum carboxylation rate
$[O_2]$	20500	Pa	O_2 partial pressure
H_J	37000	J mol^{-1}	Activation energy for maximum electron transport
R_{JV}	2.1 1.7		Original value (for 8-h photoperiod) This study (for 12-h photoperiod)
f_{spec}	0.15		Spectral correction factor due to absorbance of irradiance by tissues other than the chloroplast lamella
ST_{br}	0.125		Baseline starch production coefficient
ST_c	0.84		Proportion of night-time starch breakdown
H_r	66400	J mol^{-1}	Activation energy for leaf respiration
SSU_{min}	0.05	g sugar-C m^{-2}	Minimum sugar content in the leaves
GR_{max}	0.408	g g^{-1} d^{-1}	Maximum relative growth rate
α	0.195		Growth respiration coefficient
p_1	0.7		Ratio of intercellular to ambient CO_2 level
p_2	3.69	Pa	Quadratic constant of the CO_2 compensation point
p_3	0.188	Pa K^{-1}	Quadratic constant of the CO_2 compensation point
p_4	0.0036	Pa K^{-2}	Quadratic constant of the CO_2 compensation point
p_5	710	J K^{-1} mol^{-1}	Temperature response parameter for electron transport
p_6	220000	J mol^{-1}	Curvature parameter of J_{max}
p_7	0.7		Curvature of electron transport in response to irradiance
p_8	0.085		Leaf respiration per unit sucrose
p_9	0.016	g C m^{-2}	Basal leaf respiration

Table S2: Parameter values for the Functional-Structural Plant Model (FSPM) optimised for Col

Parameter description	Symbol	Value and unit
Threshold for switching to late vegetative stage		355 ^{0}Cd after plant emergence
Phyllochron at the early vegetative stage		30.3 ^{0}Cd
Phyllochron at the late vegetative stage*		11.9 ^{0}Cd
Seed input	$Q(1)$	1.6 x 10^{-5} g
Duration of leaf expansion	T_l	300 ^{0}Cd (cotyledons)
		400 ^{0}Cd (true leaves)
Duration of root expansion**	T_r	963.495 ^{0}Cd
Leaf lifespan after full expansion	T_s	160 ^{0}Cd
Sink strengths:		
Leaf	P_l	1 (set *a priori*)
Root system	P_r	2.64
Sink variation parameters:		
Leaf	a_l	3.07
	b_l	5.59
Root	a_r	13.03
	b_r	9.58

*For Ler, this was changed to 16.7 ^{0}Cd based on observation in our study

**The values for Ler and Fei are 913.9 ^{0}Cd and 741 ^{0}Cd respectively, which were calculated as 1.3 fold the accumulated thermal time at flowering

Table S3: Parameter values for the Arabidopsis Photothermal Model (PTM)

Parameter	Col	Ler
Threshold MPTU*	3212	2907
Non-vernalising development		
T_b	3	
CSDL	10	
CLDL	14	
D_{LD}	1 (set *a priori*)	
D_{SD}	0.6015	0.6365
P_{day}	1 (set *a priori*)	
P_{night}	0.1782	0.2118
Vernalisation submodel parameters		
T_{Vmin}	-3.5 ^{0}C	
T_{Vmax}	6 ^{0}C	
V_{sat}	960	
F_b	0.4743	0.4259
κ	-5.17	
ω	2.23	
ξ	1.00	

*The threshold values in the table above are the optimised values used in the original model (1, 2). In our study, they were modified to extend the simulation to the flowering time observed in our experiment. The values that we used are 3955 (Col), 3560 (Ler) and 2886 (Fei). As the photothermal model was never optimised for Fei in the original study, the threshold for Fei was determined by assuming that all other parameters adopted the Ler values.

Table S4: Parameter values for the Photoperiodism Model (PPM)

Symbol	Parameter value	Unit	Parameter description
v_1	2.4514	h^{-1}	Coupling constant of light activation of *LHY* transcription
v_2	5.1694	$nM\ h^{-1}$	Maximum rate of light-independent *LHY* transcription
v_3	0.6473	nM	Constant of activation by protein X
v_4	1.5283	$nM\ h^{-1}$	Maximum rate of *LHY* mRNA degradation
v_5	1.8170	nM	Michaelis constant of *LHY* mRNA degradation
v_6	0.8295	h^{-1}	Rate constant of *LHY* mRNA translation
v_7	16.8363	h^{-1}	Rate constant of LHY transport into the nucleus
v_8	0.1687	h^{-1}	Rate constant of LHY transport out of the nucleus
v_9	20.4400	$nM\ h^{-1}$	Maximum rate of cytoplasmic LHY degradation
v_{10}	1.5644	nM	Michaelis constant of cytoplasmic LHY degradation
v_{11}	3.6888	$nM\ h^{-1}$	Maximum rate of nuclear LHY degradation
v_{12}	1.2765	nM	Michaelis constant of nuclear LHY degradation
v_{13}	1.3956	$nM\ h^{-1}$	Maximum activation by protein Y
v_{14}	0.0338	nM	Constant of activation by protein Y
v_{15}	0.5539	nM^{-1}	Maximum rate of *TOC*1 transcription
v_{16}	0.2569	nM	Constant of repression by LHY
v_{17}	3.8231	$nM\ h^{-1}$	Maximum rate of *TOC*1 mRNA degradation
v_{18}	2.5734	nM	Michaelis constant of *TOC*1 mRNA degradation
v_{19}	4.3240	h^{-1}	Rate constant of *TOC*1 mRNA translation
v_{20}	0.3166	h^{-1}	Rate constant of TOC1 movement into the nucleus
v_{21}	2.1509	h^{-1}	Rate constant of TOC1 movement out of the nucleus
v_{22}	0.0013	$nM\ h^{-1}$	Maximum rate of light-dependent cytoplasmic TOC1 degradation
v_{23}	3.1741	$nM\ h^{-1}$	Maximum rate of light-independent cytoplasmic TOC1 degradation
v_{24}	2.7454	nM	Michaelis constant of cytoplasmic TOC1 degradation
v_{25}	0.0492	$nM\ h^{-1}$	Maximum rate of light-dependent nuclear TOC1 degradation
v_{26}	4.0424	$nM\ h^{-1}$	Maximum rate of light-independent nuclear TOC1 degradation
v_{27}	0.4033	nM	Michaelis constant of nuclear TOC1 degradation
v_{28}	0.2431	$nM\ h^{-1}$	Maximum transcription rate of protein X
v_{29}	0.4099	nM	Constant of activation by TOC1 protein
v_{30}	10.1132	$nM\ h^{-1}$	Maximum rate of degradation of *X* mRNA
v_{31}	6.5585	nM	Michaelis constant of *X* mRNA degradation
v_{32}	2.1470	h^{-1}	Rate constant of *X* mRNA translation
v_{33}	1.0352	h^{-1}	Rate constant of protein X movement into the nucleus
v_{34}	3.3017	h^{-1}	Rate constant of protein X movement out of the nucleus
v_{35}	0.2179	$nM\ h^{-1}$	Maximum rate of degradation of cytoplasmic protein X

v_{36}	0.6632	nM	Michaelis constant of cytoplasmic protein X degradation
v_{37}	3.3442	nM h^{-1}	Maximum rate of degradation of nuclear protein X
v_{38}	17.111	nM	Michaelis constant of nuclear protein X degradation
v_{39}	1.8272	nM h^{-1}	Light-dependent component of Y transcription
v_{40}	3.5159	nM h^{-1}	Light-independent component of Y transcription
v_{41}	1.8056	nM	Constant of repression by TOC1
v_{42}	4.2970	nM h^{-1}	Maximum rate of degradation of Y mRNA
v_{43}	1.7303	nM	Michaelis constant of Y mRNA degradation
v_{44}	0.2485	h^{-1}	Rate constant of Y mRNA translation
v_{45}	2.2123	h^{-1}	Rate constant of protein Y movement into the nucleus
v_{46}	0.2002	h^{-1}	Rate constant of protein Y movement out of the nucleus
v_{47}	0.1347	nM h^{-1}	Maximum rate of degradation of cytoplasmic protein Y
v_{48}	1.8258	nM	Michaelis constant of cytoplasmic protein Y degradation
v_{49}	0.6114	nM h^{-1}	Maximum rate of degradation of nuclear protein Y
v_{50}	1.8066	nM	Michaelis constant of nuclear protein Y degradation
v_{51}	0.0051	nM h^{-1}	Maximum rate of Y transcription
v_{52}	0.0604	nM	Constant of repression by LHY
v_{53}	28.3562	h^{-1}	Coupling constant of light activation of Y mRNA transcription
μ_1	3.3064		Hill coefficient of activation by protein X
μ_2	1.0258		Hill coefficient of activation by protein Y
μ_3	1.4422		Hill coefficient of activation by TOC1
μ_4	3.6064		Hill coefficient of repression by TOC1
μ_5	1.0237		Hill coefficient of repression by LHY
v_{COm}	1.1452	h^{-1}	Rate constant of CO (or $TOC1$) mRNA translation
v_{COp}	9.2242	nM h^{-1}	Maximum rate of light-dependent CO protein degradation
k_{COp}	2.0976	nM	Michaelis constant of CO protein degradation
V_{CO}	0.5800	nM h^{-1}	Maximum rate of FT activation by CO
K_{CO}	7.3533	nM	Michaelis constant of FT activation by CO
v_{FT}	1.8974	nM h^{-1}	Maximum rate of FT mRNA degradation
k_{FT}	5.3925	nM	Michaelis constant of FT mRNA degradation
B_{CO}	0	nM h^{-1}	Basal rate of FT transcription

Table S5: Parameter values that link the PPM to the PTM

Parameter	Ler	Col
A	1	
B	-0.3635	-0.3985
C	5.733241	5.733241
n	7	

These values were optimised using Excel Solver by constraining the maximum and minimum values of the sigmoid function to equal D_{LD} and D_{SD} (equation 3.3.1), respectively. The value of D_{SD} (Table S3) was accession-specific in the original PTM (1, 2). n was also constrained to take an integer value.

Table S6: Parameter values for the integration and initialisation of the Framework Model

Parameter description	Symbol	Value and unit
Thermal time at plant emergence	TT_0	$110\ ^0Cd$*
Initial area of cotyledons	S_1(emergence)	$1\ mm^2$
	S_2(emergence)	$1\ mm^2$
Initial sugar content per unit area		0.1 g sugar-C m^{-2}
Initial starch:sugar ratio		2
Leaf carbon content		0.3398 g C per g leaf dry mass (3, 4)
Root carbon content		0.35 g C per g root dry mass (5)

*This value was determined from observation in (6). For Fei, this was changed to 57 ^0Cd based on observation in our study

Table S7: The ratio of $J_{max}:V_{cmax}$ measured at different photoperiods from the literatures

Photoperiod (Light:Dark)	$J_{max}:V_{cmax}$	Source
8:16	2.04 – 2.26	(7)
12:12	1.7	(8)
14:10	1.36	(9)

Note: We assumed that the ratio of $J_{max}:V_{cmax}$ has upper and lower limits of 2.1 and 1.4, respectively. Therefore the ratio was set to 2.1 in the simulations for 4:20, 6:18 and 8:16 photoperiods, and 1.4 in the simulation for 18:6 photoperiod (Fig. 4). In the alternative test for 18:6 photoperiod, we reduced the ratio to 1.2 (Fig. S5b).

Table S8: Model performance in predicting fresh or dry biomass and organ areas, for three accessions, as indicated by normalised root-mean-square error (nRMSE) and linear regressions between modelled and observed values. Predictions are compared to data from 12:12 conditions (as in Figure 2I - 2K). Data are reported for the total shoot across all times measured, and for individual leaves at the specified times in days after sowing (DAS).

	Fresh weight in gram (FW)													
	Col					**Ler**					**Fei**				
	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept
Total shoot	All	6.64	0.9779	1.0365	0.0015	All	1.76	0.9995	0.9784	-0.0074	All	1.61	0.9999	1.0048	0.006
Individual leaves	18	71.89	0.5664	1.4156	-0.0002	18	12.48	0.9249	0.8184	0.0002	18	17.15	0.8261	0.9324	0.0006
	25	219.36	0.5437	2.8244	0.0031	23	21.9	0.7859	1.0417	0.0013	25	9.88	0.9268	0.9819	0.0005
	27	24.74	0.9191	1.4221	-0.0015	29	7.82	0.9585	0.9307	-0.0001	30	0.9589	1.0359	0.0008	0.0875
	38	9.82	0.9181	0.88933	0.0072	37	17.39	0.8352	1.0445	0.0055					
	Dry weight in gram (DW)													
	Col					**Ler**					**Fei**				
	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept
Total shoot	All	4.67	0.9940	1.0517	0.0009	All	5.58	0.9940	0.9603	-0.004	All	4.02	0.9984	1.0145	0.0015
Individual leaves	18	52.08	0.0834	0.2346	0.0001	18	49.29	0.4815	0.7047	-0.0001	18	39.26	0.0162	0.1363	0.0003
	25	107.86	0.1403	1.1155	0.0006	23	22.44	0.9637	0.9798	-0.0003	25	15.12	0.9363	1.2885	-0.0002
	27	23.47	0.8186	1.3068	-0.0005	29	21.11	0.9403	0.9228	-0.0006	30	17.37	0.965	1.3297	-0.0011
	38	10.41	0.9153	1.0182	0.0001	37	16	0.8451	1.0988	-0.0002					
	Area in mm^2													
	Col					**Ler**					**Fei**				
	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept	Time (DAS)	nRMSE (%)	R^2	Slope	y-intercept
Total shoot	All	8.08	0.9834	0.8523	333.1259	All	34.82	0.9805	1.6932	-288.508	All	19.54	0.9992	1.2118	95.6626
Individual leaves	18	83.54	0.6246	1.3379	3.0188	18	10.94	0.9193	0.8855	2.741	18	36.79	0.9237	1.3336	6.3475
	25	181.96	0.6583	2.5901	14.0687	23	13.14	0.9425	1.1201	2.8309	25	25.29	0.9341	1.4131	1.612
	27	40.96	0.9073	1.805	-13.5679	29	15.16	0.9383	1.2317	-0.2434	30	27.75	0.9622	1.4167	2.5531
	38	9.07	0.9367	0.978	21.4443	37	38.05	0.7886	1.4678	19.5537					
Median		**24.74**	**0.9073**	**1.1155**			**16**	**0.9383**	**0.9798**			**17.26**	**0.9476**	**1.1239**	

Table S9: Comparison of simulated and measured variation of leaf number

	Time point (Days after sowing)	Experimental data		Simulation		*P*-value
		Variance	*n*	Variance	*n*	
Ler	4	0	24	0	2400	-
	14	0.1957	24	0.1554	2400	0.3661
	17	0.1721	24	0.2208	2400	0.4786
	21	0.2826	24	0.2833	2400	0.9290
	24	0.4348	24	0.3411	2400	0.3433
	28	0.3895	24	0.4855	2400	0.5360
	36	0.6069	24	0.7454	2400	0.5671
Fei	4	0	24	0	2400	-
	14	0.2011	24	0.1675	2400	0.4641
	17	0.2156	24	0.2271	2400	0.9384
	21	0.375	24	0.3125	2400	0.4654
	24	0.4275	24	0.4430	2400	0.9823
	28	0.5217	24	0.6565	2400	0.5167

Note: *n* represents the number of samples. The alpha value for the two-tailed *F*-test was 0.05. Validation data, which were not used to determine the probability of leaf appearance in Fig. S6, are highlighted in grey. The results are based on the simulations using the MATLAB version.

Table S10: Model performance in predicting fresh and dry biomass and organ areas, for Col and *Pro35S:MIR156*, as indicated by normalised root-mean-square error (nRMSE) and linear regressions between modelled and observed values. Predictions are compared to data from 12:12 conditions (as in Figure 5F). Data are reported for individual leaves at 37 days after sowing. Leaf production rate in the model was fitted to the measured rate using either a spline function or a two-phase piece-wise linear function.

	Spline function							
	Col				*Pro35S:MIR156*			
	nRMSE (%)	R^2	Slope	y-intercept	nRMSE (%)	R^2	Slope	y-intercept
Fresh weight (g)	13.42	0.9479	1.0369	0.0065	12.45	0.9237	1.0752	0.0005
Dry weight (g)	7.38	0.961	0.9368	0	13.33	0.9204	1.0401	-0.0004
Area (mm^2)	9.18	0.9605	0.8111	22.3462	12.94	0.8983	0.885	-1.3881
	Piece-wise linear function							
	Col				*Pro35S:MIR156*			
	nRMSE (%)	R^2	Slope	y-intercept	nRMSE (%)	R^2	Slope	y-intercept
Fresh weight (g)	10.5	0.9813	1.1695	-0.0005	16.86	0.9291	1.2537	-0.0026
Dry weight (g)	8.78	0.9693	1.0428	-0.0005	15.36	0.9214	1.2099	-0.0007
Area (mm^2)	9.09	0.9583	0.9047	-4.1651	11.87	0.9193	1.037	-15.1871

Table S11: Sensitivity of simulated biomass and flowering time to parameter perturbations

Symbol*	Description	Sensitivity to a 5 % increase			Sensitivity to a 5 % decrease		
		Biomass at flowering	Flowering time	Biomass at hour 867	Biomass at flowering	Flowering time	Biomass at hour 867
p1	Base temperature (T_b)	1.096	1.0041	1.046986	0.9087	0.9846	0.944636
p2	Minimum vernalising temperature (T_{Vmin})	1	1	1	1	1	1
p3	Maximum vernalising temperature (T_{Vmax})	1	1	1	1	1	1
p4	Effective hours until saturation of vernalization (V_{sat})	1	1	1	1	1	1
p5	Parameter in the beta function for vernalisation (κ)	1	1	1	1	1	1
p6	Parameter in the beta function for vernalisation (ω)	1	1	1	1	1	1
p7	Parameter in the beta function for vernalisation (ξ)	1	1	1	1	1	1
p15	Baseline FLC repression (for Col) (F_b)	0.7831	0.9516	1	1.2672	1.0504	1
p17	MPTU threshold (for Col)	1.2561	1.0494	1	0.7743	0.9506	1
p18	Night thermal time factor (for Col) (P_{night})	0.9699	0.9856	1	1.0407	1.0041	1
p19	Degree days at plant emergence (TT_0)	1.0635	1	1.051131	1.0793	1	1.091726
p21	Parameter in the beta function for leaf sink variation (a_l)	1.0068	1	1.001641	0.9871	1	0.992745
p22	Parameter in the beta function for leaf sink variation (b_l)	0.9841	1	0.986785	1.0127	1	1.010105
p23	Root sink strength (P_r)	0.9522	1	0.956556	1.0493	1	1.044567
p24	Parameter in the beta function for root sink variation (a_r)	1.086	1	1.090603	0.9159	1	0.911988
p25	Parameter in the beta function for root sink variation (b_r)	0.9462	1	0.939627	1.0608	1	1.06901
p26	Leaf lifespan (T_s)	1	1	1	1	1	1
p27	Duration of cotyledon expansion (T_l)	0.9969	1	0.997063	0.9988	1	0.998791
p28	Duration of true leaf expansion (T_l)	1.0236	1	1.022111	0.9751	1	0.977285
p29	Magnitude of root expansion period relative to flowering time	1.1017	1	1.123683	0.9149	1	0.896096
p30	Leaf carbon content	0.8142	1	0.83365	1.2257	1	1.19468
p31	Root carbon content	0.9522	1	0.956556	1.0493	1	1.044567
p32	Seed biomass	1	1	1	1	1	1
p33	Cotyledon area at emergence	1.05	1	1.050009	0.95	1	0.949991
p34	Initial sucrose content per unit area	1.0046	1	1.004578	0.9954	1	0.995422
p35	Initial starch content per unit sucrose content	1.0032	1	1.003109	0.9968	1	0.996804

p38	Maximum carboxylation rate at 20 ^{0}C (V_{cmax})	1.1906	1	1.166264	0.8217	1	0.840041
p40	Activation energy for carboxylation constant (H_c)	1	1	1	1	1	1
p41	Michaelis-Menten constant of carboxylation at 25 ^{0}C (K_c)	1	1	1	1	1	1
p42	Activation energy for oxygenation constant (H_o)	1	1	1	1	1	1
p43	Michaelis-Menten constant of oxygenation at 25 ^{0}C (K_o)	1	1	1	1	1	1
p44	Activation energy for maximum carboxylation rate (H_v)	1	1	1	1	1	1
p45	O_2 partial pressure	1	1	1	1	1	1
p46	Quadratic constant of the CO_2 compensation point (p_2)	0.8748	1	0.888409	1.139	1	1.12161
p47	Quadratic constant of the CO_2 compensation point (p_3)	1.0254	1	1.022543	0.9754	1	0.978148
p48	Quadratic constant of the CO_2 compensation point (p_4)	0.9982	1	0.998445	1.0017	1	1.001468
p49	Fraction of ambient CO_2 level in the intercellular space (p_1)	1.1076	1	1.094403	0.8901	1	0.902142
p50	Activation energy for maximum electron transport (H_J)	0.9657	1	0.969511	1.0349	1	1.030921
p51	Curvature parameter of J_{max} (p_6)	0.9217	1	0.930299	2.0635	1	1.876576
p52	Temperature response parameter for electron transport (p_5)	2.0635	1	1.876576	0.9218	1	0.930472
p53	J_{max}:V_{cmax} ratio (R_{JV})	1.1906	1	1.166264	0.8217	1	0.840041
p54	Spectral correction factor due to absorbance of irradiance by tissue other than the chloroplast lamella (f_{spec})	0.9725	1	0.975643	1.0278	1	1.024702
p55	Curvature of electron transport in response to irradiance (p_7)	1.1546	1	1.134911	0.8721	1	0.885991
p56	Activation energy for leaf respiration (H_r)	0.989	1	0.990326	1.0106	1	1.009414
p57	Leaf respiration per unit sucrose (p_8)	0.9804	1	0.982553	1.019	1	1.016929
p58	Basal leaf respiration (p_9)	0.9295	1	0.937295	1.0736	1	1.064778
p60	Baseline starch production coefficient (ST_{br})	1.0004	1	1.000173	0.9994	1	0.999827
p61	Proportion of night-time starch breakdown (ST_c)	1.0254	1	1.024961	0.9734	1	0.973916
p62	Minimum sugar content in the leaves (SSU_{min})	0.9789	1	0.981085	1.0209	1	1.018656
p63	Maximum relative growth rate (GR_{max})	1.1904	1	1.190534	0.821	1	0.82104
p64	Growth respiration coefficient (α)	0.939	1	0.945759	1.0625	1	1.055105
p65	Minimum zenithal angle (α_i for $1 \le i \le i_{max}$)	0.9857	1	0.987131	1.0132	1	1.011833

p66	Maximum increase in zenithal angle	0.98	1	0.981344	1.0204	1	1.019088
p67	Thermal time since plant emergence for the juvenile-to-adult transition	0.9972	1	0.98143	0.9965	1	1.006564
p71	Specific leaf area for cotyledons	1.2145	1	1.179133	0.8087	1	0.832873
p72	Rate of decrease in specific leaf area	0.8446	1	0.870962	1.1897	1	1.15089
p78	Sigmoidal link to the clock (B)	1.0717	1.0072	1	0.9502	0.9825	1
p79	Sigmoidal link to the clock (C)	1.2225	1.0412	1	0.8096	0.9547	1
p80	Sigmoidal link to the clock (n)	1	1	1	1	1	1
p82	Phyllochron at the juvenile phase	0.978	1	0.968647	1.0072	1	1.013042
p83	Phyllochron at the adult phase	1.0069	1	0.996891	0.9902	1	1.0019
v1	Coupling constant of light activation of *LHY* transcription	0.9698	0.9887	1	1.0305	1.0031	1
v2	Maximum rate of light-independent *LHY* transcription	1	1	1	0.9899	0.999	1
v3	Constant of activation by protein X	0.9899	0.999	1	1	1	1
v4	Maximum rate of *LHY* mRNA degradation	1.0613	1.0062	1	0.9019	0.9774	1
v5	Michaelis constant of *LHY* mRNA degradation	0.96	0.9835	1	1.0407	1.0041	1
v6	Rate constant of *LHY* mRNA translation	1	1	1	0.9798	0.9979	1
v7	Rate constant of LHY transport into the nucleus	1	1	1	0.9899	0.999	1
v8	Rate constant of LHY transport out of the nucleus	1	1	1	1	1	1
v9	Maximum rate of cytoplasmic LHY degradation	0.9899	0.999	1	1	1	1
v10	Michaelis constant of cytoplasmic LHY degradation	1	1	1	0.9899	0.999	1
v11	Maximum rate of nuclear LHY degradation	1	1	1	0.9798	0.9979	1
v12	Michaelis constant of nuclear LHY degradation	0.9899	0.999	1	1	1	1
v13	Maximum activation by protein Y	0.855	0.963	1	1.0305	1.0031	1
v14	Constant of activation by protein Y	1.0928	1.0134	1	0.883	0.9753	1
v15	Maximum rate of *TOC*1 transcription	0.855	0.963	1	1.0305	1.0031	1
v16	Constant of repression by LHY	0.9899	0.999	1	0.8549	0.965	1
v17	Maximum rate of *TOC*1 mRNA degradation	1.093	1.0216	1	0.749	0.9352	1
v18	Michaelis constant of *TOC*1 mRNA degradation	0.8551	0.9609	1	1.1249	1.0247	1
v19	Rate constant of *TOC*1 mRNA translation	0.9502	0.9825	1	1.0305	1.0031	1
v20	Rate constant of TOC1 movement into the nucleus	1.2221	1.034	1	0.7655	0.9496	1
v21	Rate constant of TOC1 movement out of the	0.864	0.9733	1	1.1355	1.0257	1

	nucleus						
v22	Maximum rate of light-dependent cytoplasmic TOC1 degradation	1	1	1	1	1	1
v23	Maximum rate of light-independent cytoplasmic TOC1 degradation	1.1569	1.0278	1	0.8547	0.9702	1
v24	Michaelis constant of cytoplasmic TOC1 degradation	0.9696	0.9969	1	1.0203	1.0021	1
v25	Maximum rate of light-dependent nuclear TOC1 degradation	1	1	1	1	1	1
v26	Maximum rate of light-independent nuclear TOC1 degradation	1.0927	1.0093	1	0.9115	0.9784	1
v27	Michaelis constant of nuclear TOC1 degradation	0.96	0.9835	1	1.0613	1.0062	1
v28	Maximum transcription rate of protein X	0.883	0.9753	1	0.8552	0.9599	1
v29	Constant of activation by TOC1 protein	0.9696	0.9969	1	1.0101	1.001	1
v30	Maximum rate of degradation of X mRNA	0.8547	0.9702	1	0.8459	0.9588	1
v31	Michaelis constant of X mRNA degradation	0.8547	0.9691	1	0.855	0.964	1
v32	Rate constant of X mRNA translation	0.883	0.9753	1	0.8551	0.9609	1
v33	Rate constant of protein X movement into the nucleus	1.0101	1.001	1	0.9698	0.9866	1
v34	Rate constant of protein X movement out of the nucleus	0.9697	0.9928	1	1	1	1
v35	Maximum rate of degradation of cytoplasmic protein X	0.96	0.9835	1	0.9697	0.9907	1
v36	Michaelis constant of cytoplasmic protein X degradation	1.0101	1.001	1	0.9798	0.9979	1
v37	Maximum rate of degradation of nuclear protein X	0.9899	0.999	1	0.9798	0.9979	1
v38	Michaelis constant of nuclear protein X degradation	0.9798	0.9979	1	0.9899	0.999	1
v39	Light-dependent component of Y transcription	1	1	1	0.9899	0.999	1
v40	Light-independent component of Y transcription	0.864	0.9733	1	1.1143	1.0237	1
v41	Constant of repression by TOC1	1.0305	1.0031	1	0.9502	0.9825	1
v42	Maximum rate of degradation of Y mRNA	1.1143	1.0237	1	0.855	0.963	1
v43	Michaelis constant of Y mRNA degradation	0.8546	0.9722	1	1.1143	1.0237	1
v44	Rate constant of Y mRNA translation	0.883	0.9753	1	1.0929	1.0206	1
v45	Rate constant of protein Y movement into the nucleus	1.0203	1.0021	1	0.9697	0.9938	1
v46	Rate constant of protein Y movement out of the nucleus	0.9899	0.999	1	1.0101	1.001	1
v47	Maximum rate of degradation of	1.0101	1.001	1	0.9899	0.999	1

	cytoplasmic protein Y						
v48	Michaelis constant of cytoplasmic protein Y degradation	0.9899	0.999	1	1.0101	1.001	1
v49	Maximum rate of degradation of nuclear protein Y	1.2223	1.0381	1	0.7405	0.9342	1
v50	Michaelis constant of nuclear protein Y degradation	0.7486	0.9403	1	1.2225	1.0422	1
v51	Maximum rate of Y transcription	0.883	0.9753	1	1.093	1.0216	1
v52	Constant of repression by LHY	1.1036	1.0226	1	0.8549	0.965	1
v53	Coupling constant of light activation of Y mRNA transcription	1.0101	1.001	1	0.9798	0.9979	1
v54	Hill coefficient of activation by protein Y (μ_2)	1.3006	1.0535	1	0.563	0.892	1
v55	Hill coefficient of activation by protein X (μ_1)	0.9404	0.9815	1	1.0203	1.0021	1
v56	Hill coefficient of activation by TOC1 (μ_3)	0.9698	0.9877	1	1.0305	1.0031	1
v57	Hill coefficient of repression by TOC1 (μ_4)	0.9696	0.9949	1	1.0203	1.0021	1
v58	Hill coefficient of repression by LHY (μ_5)	1.0927	1.0103	1	0.9115	0.9784	1
v59	Rate constant of CO (or $TOC1$) mRNA translation (v_{COm})	0.8276	0.9568	1	1.2223	1.0391	1
v60	Maximum rate of light-dependent CO protein degradation (v_{COp})	1.0927	1.0093	1	0.921	0.9794	1
v61	Michaelis constant of CO protein degradation (k_{COp})	0.9502	0.9825	1	1.0717	1.0072	1
v62	Maximum rate of FT activation by CO (V_{CO})	0.8096	0.9547	1	1.245	1.0484	1
v63	Michaelis constant of FT activation by CO (K_{CO})	1.2001	1.0319	1	0.8459	0.9588	1
v64	Maximum rate of FT mRNA degradation (v_{FT})	1.2339	1.0473	1	0.8007	0.9537	1
v65	Michaelis constant of FT mRNA degradation (k_{FT})	0.8185	0.9558	1	1.2339	1.0473	1
v66	Basal rate of FT transcription (B_{CO})	1	1	1	1	1	1
w	Water content	2.3528	1	2.352911	0.6349	1	0.634911

*Note that the symbol/numbering used for each parameter in this table only corresponds to the MATLAB code used for running the sensitivity analysis, and differs from the symbols in the model description section.

1.	Chew YH, *et al.* (2012) An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytol* 194(3):654-665.
2.	Wilczek AM, *et al.* (2009) Effects of Genetic Perturbation on Seasonal Life History Plasticity. *Science* 323(5916):930-934.
3.	Gorsuch PA, Pandey S, & Atkin OK (2010) Temporal heterogeneity of cold acclimation phenotypes in Arabidopsis leaves. *Plant Cell Environ* 33(2):244-258.
4.	Gorsuch PA, Pandey S, & Atkin OK (2010) Thermal de-acclimation: how permanent are leaf phenotypes when cold-acclimated plants experience warming? *Plant Cell Environ* 33(7):1124-1137.
5.	Kumar S, Udawatta RP, & Anderson SH (2010) Root length density and carbon content of agroforestry and grass buffers under grazed pasture systems in a Hapludalf. *Agrofor Syst* 80(1):85-96.
6.	Boyes DC, *et al.* (2001) Growth stage-based phenotypic analysis of arabidopsis: A model for high throughput functional genomics in plants. *Plant Cell* 13(7):1499-1510.
7.	Pons TL (2012) Interaction of temperature and irradiance effects on photosynthetic acclimation in two accessions of Arabidopsis thaliana. *Photosynth Res* 113(1-3):207-219.
8.	Flexas J, *et al.* (2007) Mesophyll conductance to CO2 in Arabidopsis thaliana. *New Phytol* 175(3):501-511.
9.	Bunce JA (2008) Acclimation of photosynthesis to temperature in Arabidopsis thaliana and Brassica oleracea. *Photosynthetica* 46(4):517-524.

Figure S1: Model simulation using the original parameter values, with the $J_{max}:V_{cmax}$ ratio (or R_{JV}) taking the value of 2.1. Simulations (solid lines) and experimental data of Col wild type (symbols) at 18, 25, 27 and 38 days after sowing (DAS) for total shoot (a) and individual leaf biomass (b) are illustrated. Leaves are ranked according to the order of appearance. Experimental conditions: ~21.3 ^{0}C; 12:12 light/dark cycle; light intensity = 110 μmol m^{-2} s^{-1}; Average daytime CO_2 level = 375 ppm.

Figure S2: Shoot growth of Ler and Fei until the appearance of flower buds. (a) Leaf number plotted against days after sowing. Only leaves exceeding 1 mm^2 in size (estimated with the naked eye) were considered in the total leaf count (including the cotyledons); (b) Phytomer number, i.e. leaf number minus one, plotted as a function of thermal time after plant emergence. In the plot for Fei, the thermal time at plant emergence has been modified to the value observed in our study (Table S6). Also shown are phytomer numbers as modelled for Col data in the GreenLab functional-structural model, and the modified rate for Ler (Table S6). Error bars showing the standard error of 24 plants are very small and therefore hidden by the data markers.

Figure S3: Simulation for Col (a,b) and Fei (c,d) using a standardised water content of 91 %.

Figure S4: Experimental setup for gas exchange measurement

Figure S5: Experimental data (black and white) and model simulations (light and dark green) for sucrose level (a) and rosette fresh weight (b) at the end of day (ED; white and light green) and end of night (EN; black and dark green), for plants grown under the photoperiod conditions detailed in Fig. 4. Symbols for ED and EN simulation results overlap in panel (a). The simulation for 18h photoperiod in panel (b) used a J_{max}:V_{cmax} value of 1.2, which was extrapolated from the 14h value reported in the literature, and a starch breakdown level reduced to 60% based on the data. Parameter modifications for other photoperiods were similar to those used in Fig. 4. Error bars show the standard deviation of 5 plants.

Figure S6: Plot of variance against mean phytomer number using leaf number data for both Ler and Fei at the early vegetative stage. The slopes of both regressions were close to 0.03, giving a binomial probability of 0.97.

Figure S7: (a) Schematic diagram showing how the timing of phase transition was varied for each simulation, which produced different final leaf numbers at flowering. Dashed lines of similar colours represent the controls that produce the same final number of leaves but at a single rate. Model simulations (solid lines) and experimental data (symbols) of rosette leaf number (b) and individual leaf biomass at 37 DAS (c) for Col and *Pro35S:MIR156* using a piecewise-linear leaf production model. For *Pro35S:MIR156*, the phyllochrons at the early and late vegetative stages were determined with linear regression to be 18.9 °Cd and 7.3 °Cd, respectively. Note that the simulation lines in b are not exactly linear due to the coarse time-step, i.e. increment in degree-days, relative to the phyllochrons.

Figure S8: Sensitivities of parameters (excluding the PPM and water content) to single perturbations of 5 % increase (+) or 5 % decrease (-) in each parameter value. Sensitivities were scored relative to the reference values, (top panel) for flowering time (right axis) and fresh biomass at flowering (left axis), and (bottom panel) for fresh biomass at hour 876. The symbol for each parameter corresponds to the list in Table S11.

Figure S9: Sensitivities of parameters in the PPM and water content to perturbations of 5 % increase (+) and 5 % decrease (-). Sensitivities were scored relative to the reference for flowering time and fresh biomass at flowering (top panel), and for fresh biomass at hour 876 (bottom panel). The symbol for each parameter corresponds to the list in Table S11.

Figure S10: Schematic diagram of the Carbon Dynamic Model (CDM). Carbon is transported (solid arrows) in and/or out of four different pools (grey boxes) through various processes. Dashed arrows indicate information or feedback input.

Figure S11: Schematic diagram of the Arabidopsis Functional-Structural Plant Model (FSPM). Biomass *Q* is distributed (solid arrows) to different organs (grey boxes) based on their physiological age. The above-ground structure, which determines the area for light interception (S_{ros}), is also considered in the model. Dashed arrows indicate information or feedback input.

Figure S12: The three components of the Photothermal Model (PTM). (a) Photoperiod component, which is divided into three sections by two critical day lengths; (b) Thermal time component, which only considers hourly temperatures above 3 ^{0}C. A filter function is used to account for the different sensitivity to day and night temperatures; (c) Vernalisation component, which has two sub-components. The modifier represents the extent of vernalisation and depends on the accumulated vernalisation effectiveness at different temperatures; (d) The Modified Photothermal Unit (MPTU) is a product of the three components. The MPTU is accumulated every hour until a threshold value is reached to indicate the reproductive switch.

Figure S13: Schematic diagram of the Photoperiodism Model (PPM).
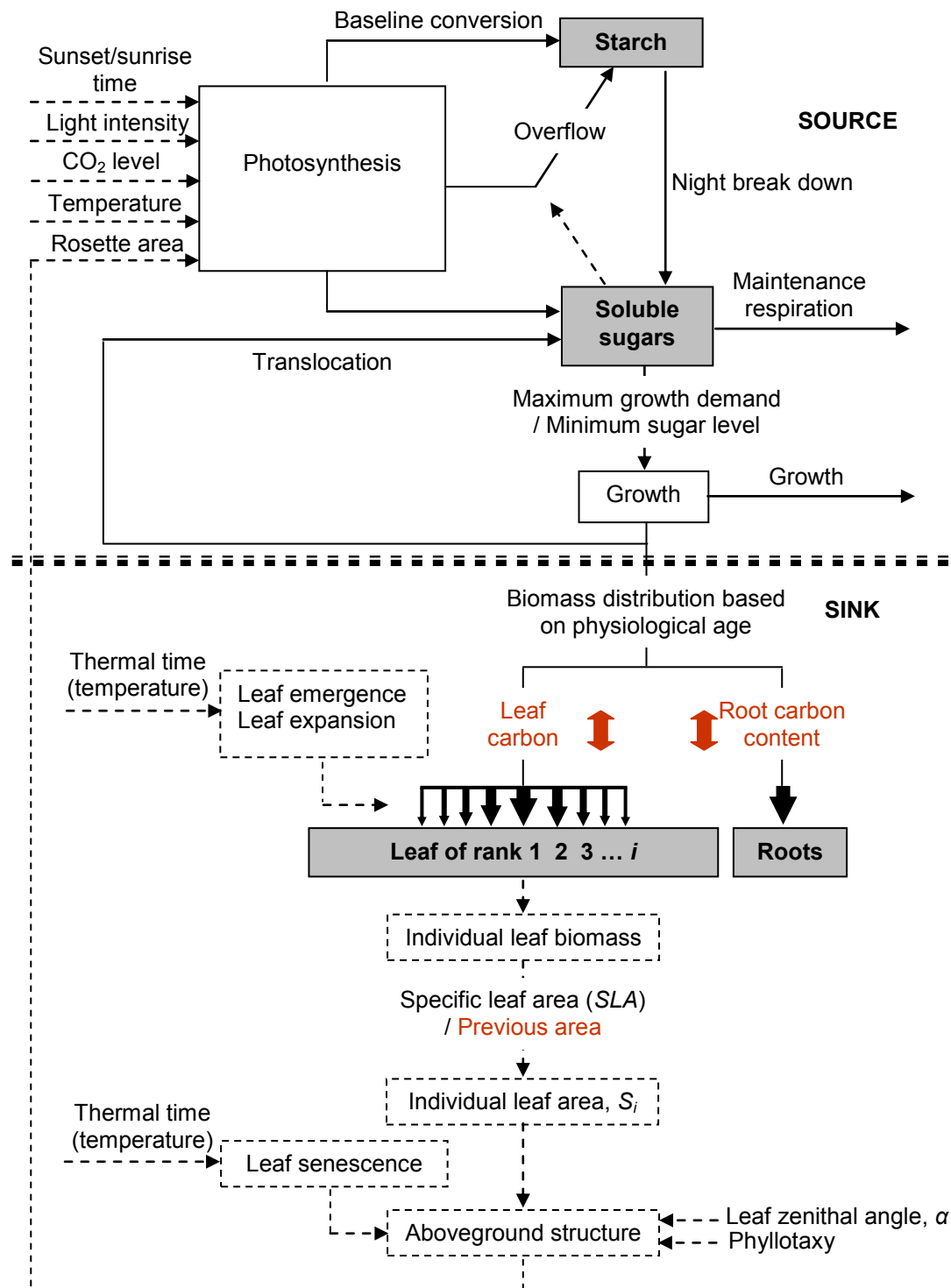
Figure S14: Schematic diagram illustrating the combined CDM and FSPM, which form the vegetative growth component. It is a combination of Figures S10 and S11. Texts in red highlight the new features that were introduced in the current work. The red double-sided arrows indicate the connection points where biomass information between the two models is sent back and forth.
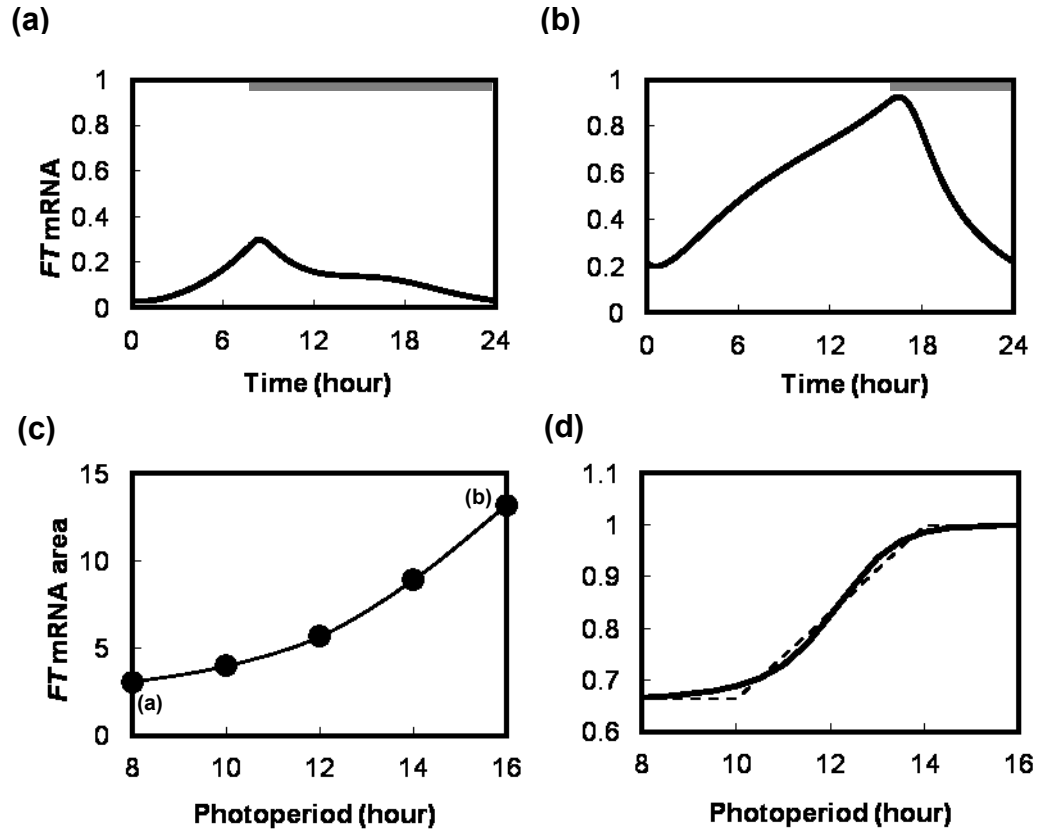
Figure S15: Calibrating the PPM to the photoperiod component of the PTM. The time-series of *FT* mRNA simulated by the PPM for: (a) 8h photoperiod and; (b) 16h photoperiod; (c) A plot of the area under the curve of *FT* mRNA time series at different photoperiods. Letters in brackets link to the associated time-series. These areas were translated into the photoperiod component using the sigmoid function in equation 3.7.1; (d) Comparison between the broken-stick component in the PTM (dashed line, equation 3.3.1) and the calibrated photoperiod component (full line).
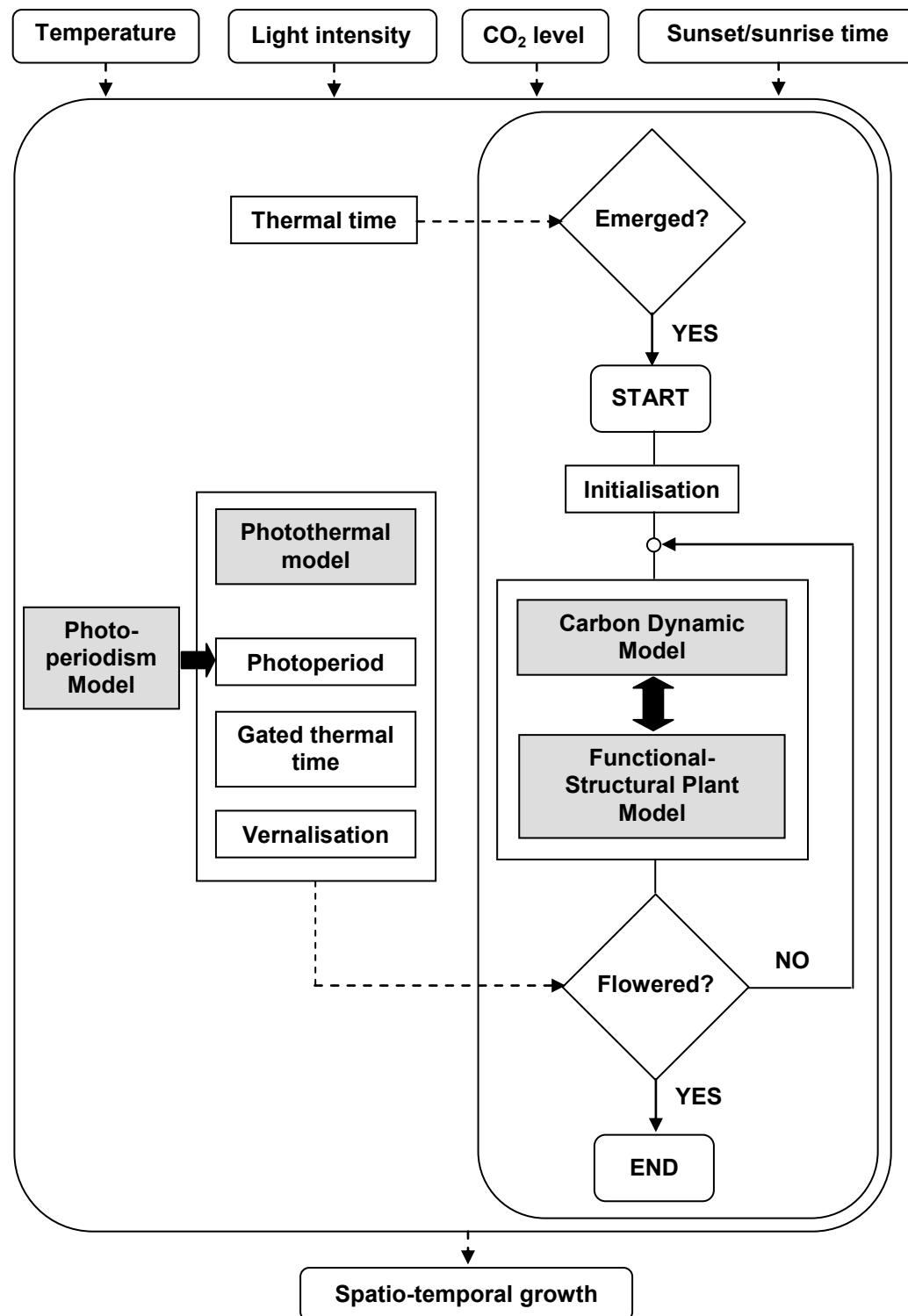
Figure S16: Schematic diagram of the multi-scale Framework Model. The four sub-models are represented by the grey blocks. Line arrows: circuit flow; Dashed arrows: information flow; Bold arrows: values being supplied by one component to another.
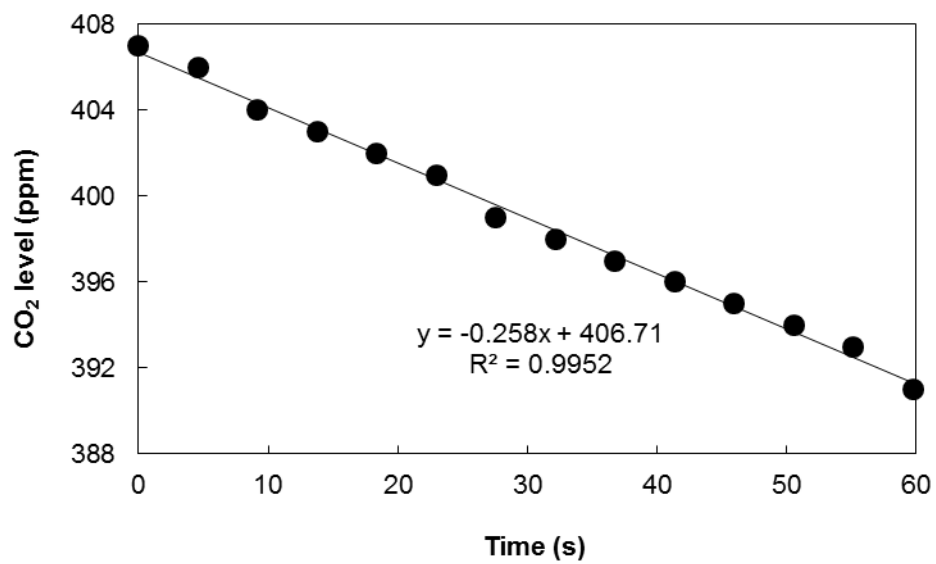
Figure S17: An example of CO_2 measurement. Shown here are the data for a tray of 5-week-old Ler, where the average daytime CO_2 concentration in the growth chamber was 405 ppm.