Multiscale digital Arabidopsis predicts individual organ and whole-organism growth

Chew, Yin Hoon; Wenden, Benedicte; Flis, Anna; Mengin, Virginie; Taylor, Jasper; Davey, Christopher L.; Tindal, Christopher; Thomas, Howard; Ougham, Helen J.; De Reffye, Philippe; Stitt, Mark; Williams, Mathew; Muetzelfeldt, Robert; Halliday, Karen J.; Millar, Andrew J.

Published in:
Proceedings of the National Academy of Sciences of the United States of America

DOI:
10.1073/pnas.1410238111

Publication date:
2014

Citation for published version (APA):
Chew, Y. H., Wenden, B., Flis, A., Mengin, V., Taylor, J., Davey, C. L., Tindal, C., Thomas, H., Ougham, H. J., De Reffye, P., Stitt, M., Williams, M., Muetzelfeldt, R., Halliday, K. J., & Millar, A. J. (2014). Multiscale digital Arabidopsis predicts individual organ and whole-organism growth. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(39), E4127-E4136. https://doi.org/10.1073/pnas.1410238111

A Multi-Scale Digital Arabidopsis Predicts Individual Organ and Whole-Organism Growth

Yin Hoon Chew[1], Bénédicte Wenden[2], Anna Flis[3], Virginie Mengin[3], Jasper Taylor[4], Christopher L. Davey[5], Christopher Tindal[1], Howard Thomas[5], Helen J. Ougham[5], Philippe de Reffye[6], Mark Stitt[3], Mathew Williams[7], Robert Muetzelfeldt[4], Karen J. Halliday[1] & Andrew J. Millar[1,a]

[1] SynthSys, C. H. Waddington Building, University of Edinburgh, King's Buildings, Edinburgh EH9 3JD, UK [2] INRA and Université Bordeaux, UMR 1332 de Biologie du Fruit et Pathologie, F-33140 Villenave d'Ornon, France [3] Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam-Golm, Germany [4] Simulistics Ltd., 2B Pentland Park, Loanhead, EH20 9PA, UK [5] Edward Llwyd Building, 2.14, IBERS Penglais Campus, Aberystwyth University, SY23 2FG, UK [6] Cirad-Amis, UMR AMAP, TA 40/01 Avenue Agropolis, F-34398 Montpellier, Cedex 5, France [7] School of GeoSciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, UK

Submitted to Proceedings of the National Academy of Sciences of the United States of America

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 modelled 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 multi-scale 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 new 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, multi-scale models, assembling knowledge from systems biology to ecophysiology, will help to understand and to engineer plant behaviour from the genome to the field.

plant growth model | multi-scale | digital organism | crop science | ecology

Introduction

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 favourable season, and thus the associated environment, increases reproductive success (3). This synchronisation 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, in order to optimise 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-recognised 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, multi-scale models might be developed in a

Significance

Plants respond to environmental change by triggering biochemical and developmental networks, across multiple scales. Multi-scale 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 modelling.

Reserved for Publication Footnotes

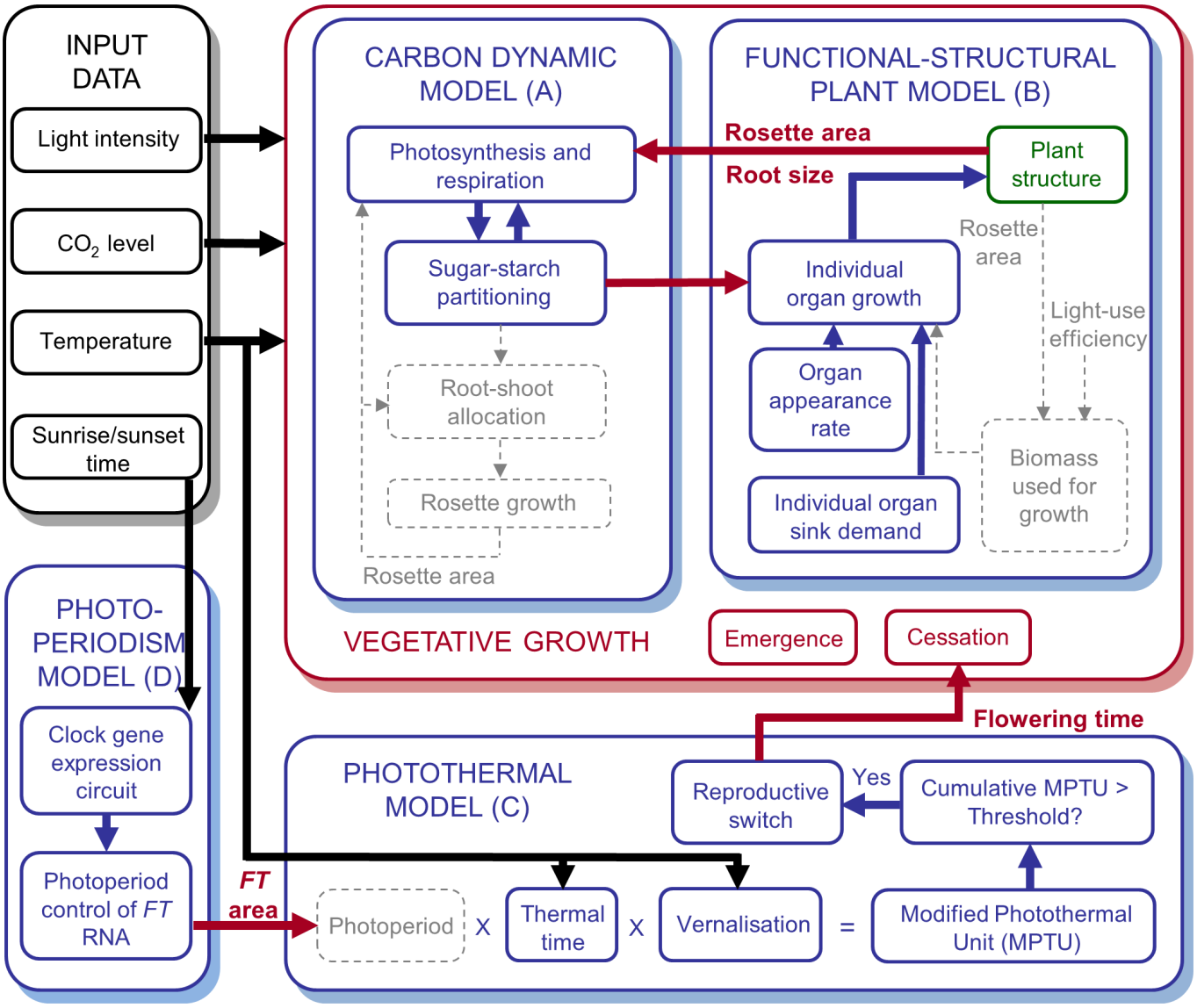

Fig. 1. Overview of the Framework Model. The Framework Model takes environmental data as input (black) to four existing Arabidopsis models (blue shadowed boxes), which are: (A) a Carbon Dynamic Model (CDM) that describes carbon assimilation and resource partitioning (31); (B) a Functional-Structural 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). Upon integration, several original components were discarded (grey) while new connections were created (red).

distributed, community approach. Experts in each domain could model a particular aspect of biology in detail and re-use 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 Modelling community. In that case, sub-models 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 modelling 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 co-ordinated data acquisition by researchers from multiple disciplines. The quantitative compatibility of the constituent models is therefore a key question.

In order to develop a quantitative model that links multiple, interacting processes from metabolism to development, we implemented the modular modelling 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 characterise different aspects of plant biology, 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,

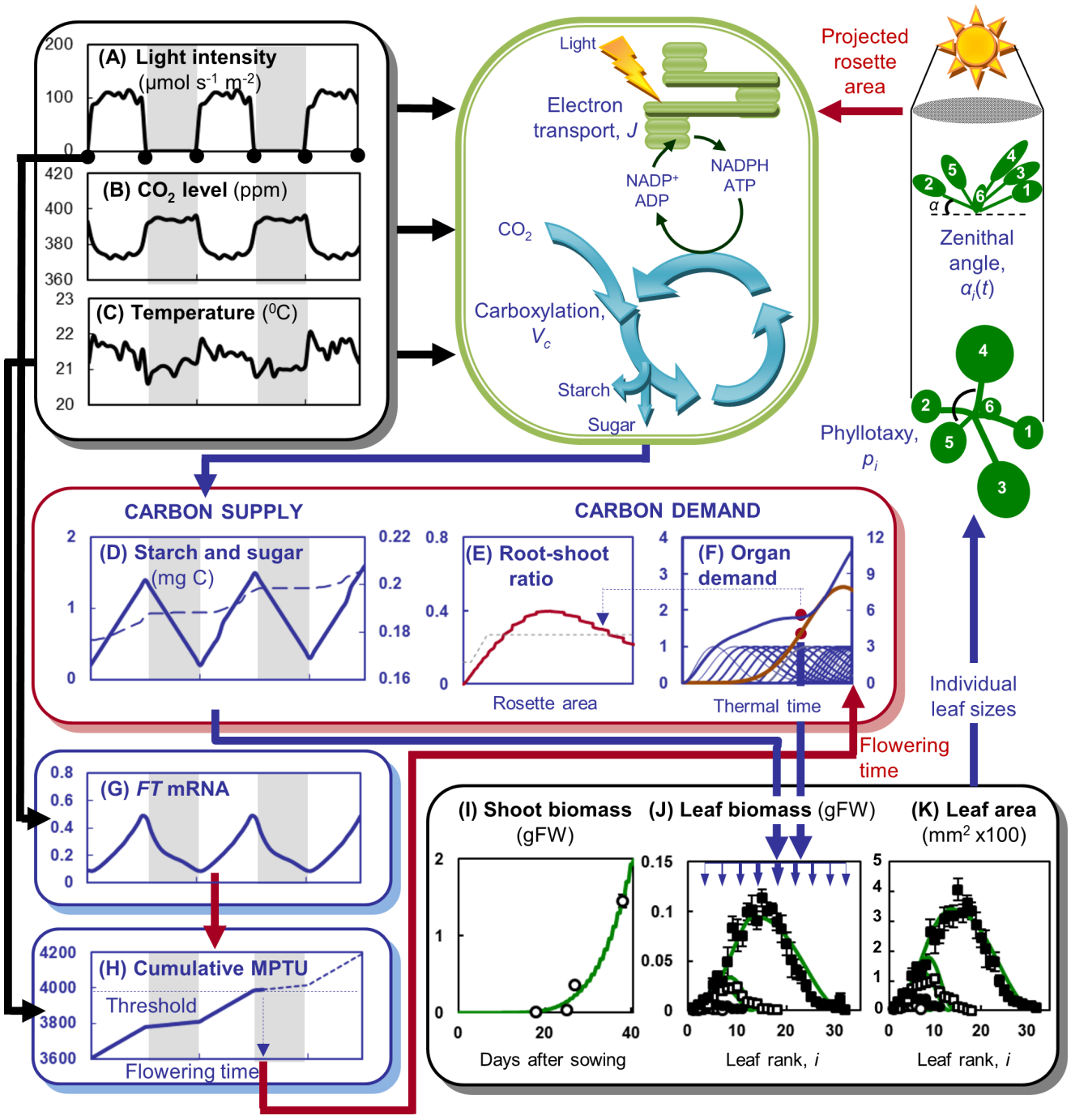

Fig. 2. The Framework Model's workflow predicts whole-plant and individual organ growth data. Input data required are hourly light intensity (A), CO_2 level (B) and temperature (C), illustrated for simulated three 12h L (open):12h D (shaded area) cycles. Carbon supply (D) is used as sugar (dashed line) or stored as starch (solid line). Carbon is allocated at each hourly time step according to the organ demand (E, 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) (grey 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 wild type (green lines) closely match experimental data, for (I) total shoot biomass, (J) for leaf biomass and (K) for leaf area, at 18 (open circles), 25 (filled circles), 27 (open squares) and 38 (filled squares) days 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 light/dark cycle; light intensity, 110 μmol m^{-2} s^{-1}; mean daytime CO_2 level, 375 ppm. The error bars show the standard errors of 5 plants. The colour code links to the model components in Fig. 1.

quantitative agreement. The FM was validated against metabolic, physiological and biomass data for multiple genetic backgrounds and tested for several growth conditions. Numerical simulations using FM enabled us to understand the physiological relevance

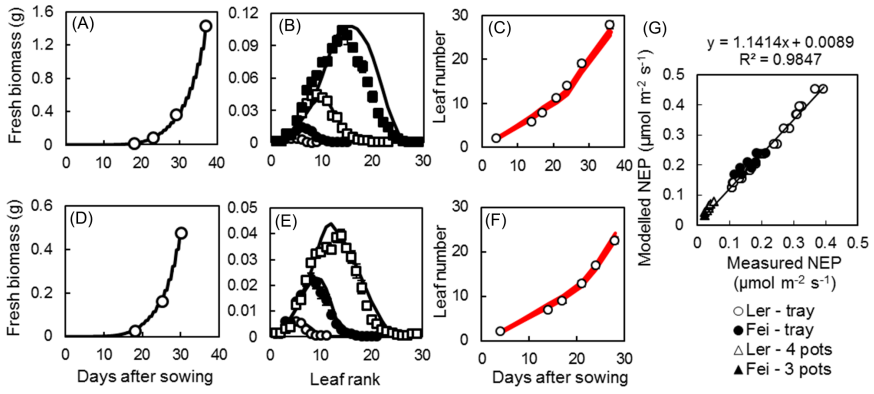

Fig. 3. The Framework Model predicts plant growth and gas exchange data for different accessions. Model simulations (solid lines) and experimental data (symbols) of total shoot biomass, individual leaf biomass and leaf number for Ler (A,B,C) and Fei (D,E,F) are shown. Time points of measurement in B are 18 (open circles), 23 (filled circles), 29 (open squares) and 37 (filled squares) days after sowing (DAS). Time points of measurement in E are 18 (open circles), 25 (filled circles) and 30 (open squares) DAS. The thickness of the red lines in C and F represents a region with one standard deviation above and below the mean values from the stochastic simulations of leaf number for 2400 model runs. The plot of modelled and measured Net Ecosystem Production (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 light/dark cycle; light intensity = 130 µmol m^{-2} s^{-1}; Average daytime CO_2 concentration = 375 ppm. Error bars in A, B, D and E show the standard errors of: n = 10 plants for total shoot biomass; n = 5 plants for individual leaf biomass. Error bars in C and F (smaller than the symbols) represent the standard deviation of 24 plants.

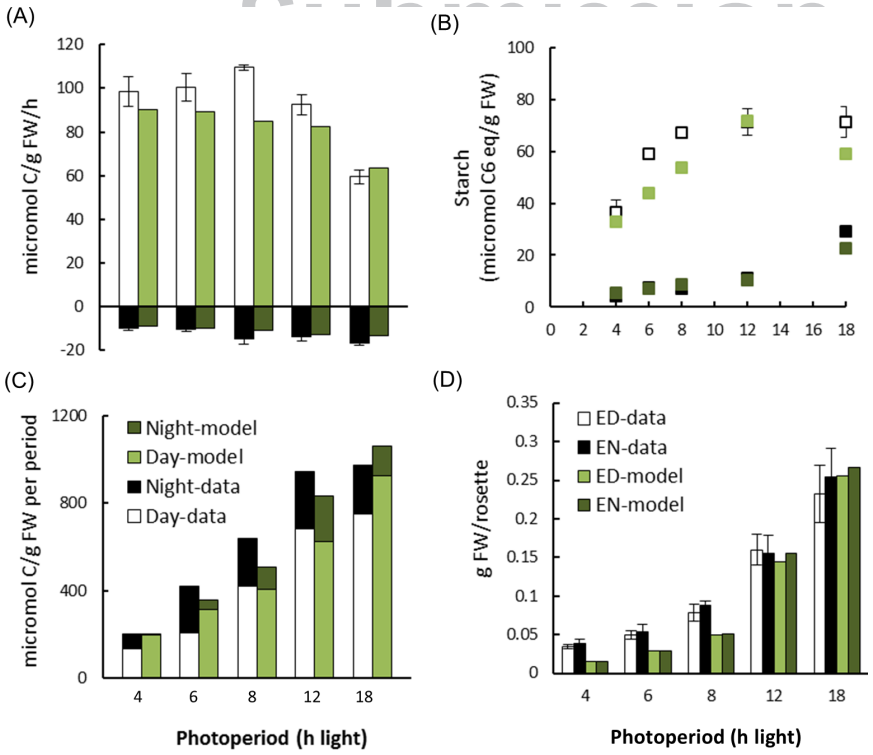

Fig. 4. Testing the Framework Model under different photoperiods. Experimental data (black and white; (51) are compared with model simulations (light and dark green) in the photoperiods indicated, for (A) carbon assimilation and respiration rates; (B) starch levels; (C) amount of growth per day or night 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 standard deviation of 5 plants.

of an observed developmental process. Finally, we compared new data with model predictions to quantitatively explain a developmental phenotype, supporting one of two proposed mechanisms.

Model Description

Components of the Framework Model

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 multi-scale 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:

A. 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

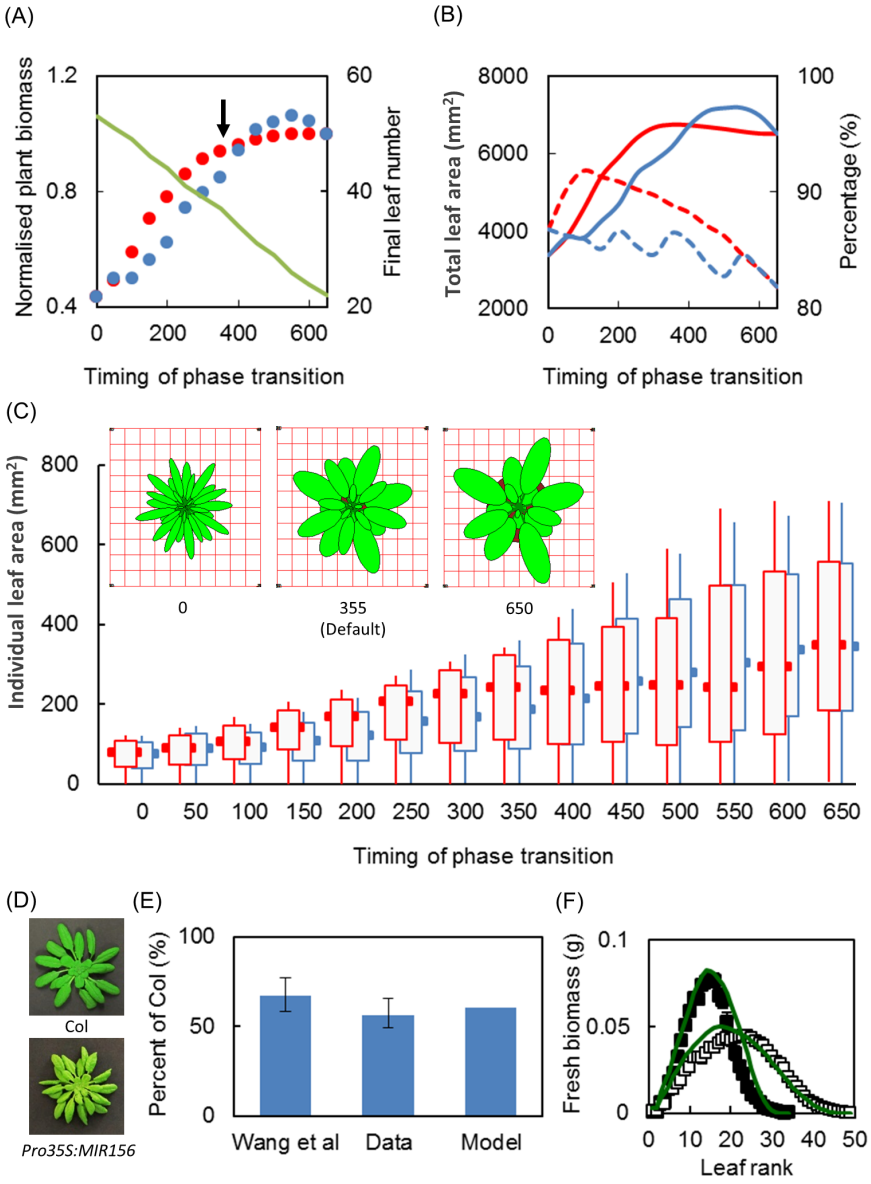

Fig. 5. Leaf production rate balances biomass and leaf area for photosynthesis. Simulation results with time-varying leaf production rates (red) and the associated controls with constant rates (blue) are shown for: (A) plant biomass (symbols, left axis) and final leaf number at flowering (green line, right axis); biomass is normalised 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 phase transition point in our model. The timing of the phase transition (*y*-axes) are expressed in thermal time after plant emergence. (D) Rosette images of 37-day-old Col wild type (upper) and the greater numberof smaller leaves in *Pro35S:MIR156* (lower). (E) Area of the largest leaf in *Pro35S:MIR156*, relative to Col wild type (100%), in data of Wang et al (54), our experimental data and model simulation. Error bars show the standard deviation of 5 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 (filled squares) and *Pro35S:MIR156* (open squares) at 37 DAS. Experimental conditions: ~20.7 °C; 12:12 light/dark cycle; light intensity = 100 μmol m^{-2} s^{-1}; Average daytime CO_2 concentration = 405 ppm. Error bars show the standard errors of 5 plants.

was a discrete-time model with a 6s time step, constructed using data of Columbia (Col) wild-type plants grown under 8h Light (L):16h Dark (D) conditions.

B. 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 (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 parameterised using data of Col wild-type plants grown under 12hL:12hD conditions. Only a subset of the large original model (34) was applicable to our study. The relevant sub-set of parameter values and developmental structures was rewritten into a conventional, dynamic form that was compatible with the other sub-models, as a discrete-time model with an hourly time step.

C. 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 vernalisation pathway that promotes flowering in the spring after a long chilling period over winter (36), and by warm ambient temperature (37). Each of these factors is represented by effective mathematical functions in the model, without mechanistic detail, and it was parameterised 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.

D. A Photoperiodism Model (PPM), which is a gene dynamic model of the circadian clock (39) and the photoperiod pathway (6). This was a conventional ordinary-differential-equation (ODE) model, usually solved with an adaptive time step of minutes or less. The model was parameterised using data from Col and Ler wild-type plants grown under 16hL:8hD and 8hL:16hD 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, while redundant model components were replaced (Fig. 1 and Supplementary Information). Unit conversions were required for compatibility,

and two parameter values (8 and 9 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:

1. The model's time step was standardised to one hour 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. 2A-C), and thereby resolves diel behaviour.

2. 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. 2E-F).

3. To facilitate the replacement step 2 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.

4. 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.

5. 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 replaced the empirical light-use efficiency (LUE) component, which was previously estimated from experimental data through model inversion (34).

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

7. 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).

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

9. 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 the modelling work and analysis were conducted in MATLAB (Mathworks, Cambridge, UK) (see Supplementary Information). The Plant Systems Modelling (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 will be publicly accessible from PlaSMo upon publication, in MATLAB and Simile formats. Simile provides a visual modelling environment with a Graphical User Interface, plotting tools and an animated display of simulated plant growth (see Supplementary Video) (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. Wild-type Col plants were grown in 12hL:12hD 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 8h 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 sub-model, the FM overestimated growth (Fig. S1). However, the literature shows that Arabidopsis grown in an 8h photoperiod have altered photosynthetic physiology compared to our reference 12h photoperiod. Specifically, the ratio of maximum electron transport to the maximum rate of carboxylation (J_{max}:V_{cmax}) decreases as photoperiod increases (45-47) (Table S7). The CDM's original value for J_{max}:V_{cmax} has only been tested in an 8h photoperiod (31). Substituting the value measured in a 12h photoperiod was sufficient for the FM to fit the Col biomass data (Fig. 2I-K). The R^2 between measured and modelled values of fresh biomass, dry biomass and area of the rosette were 0.98, 0.99, and 0.98, respectively, with normalised Root Mean Square Error (nRMSE) less than 10% (Table S8). The median values of R^2 and nRMSE for all the data, including individual leaf predictions, were 0.91 and 24.7% respectively. The dynamic operation of the model in Simile is illustrated in the Supplementary Video.

The FM was next tested by comparison to 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 to Ler at the same time points (Fig. S2a). However, leaf appearance rate in Fei matched the Col rate when plotted against thermal time after seedling emergence (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 (Table S6). With only these changes, the model's match to data of Ler and Fei plants was as good as for Col (Fig. 3A,B,D,E), with median R^2 (and nRMSE) of 0.94 (16%) and 0.95 (17.3%), respectively (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 standardised water content of 91%. This caused slight overestimation of fresh biomass for Fei, but less significant effects for Col (Fig. S3).

We additionally tested the applicability of our multi-scale FM to ecosystem studies, by comparing model simulations to measured trace gas exchange data (see Supplementary Information). 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) (Fig. S4). The model accurately predicted measured gas exchange from 26 days after sowing until flowering time (R^2 = 0.98) (Fig. 3G). 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 (Table S11; Figs. S8 - S9), because of the longer (shorter) duration of biomass accumulation in the rosette. Flowering time was controlled by parameters of the photoperiodism model (PPM), by the overall flowering threshold and by the baseline *FLC* repression, as expected in our non-vernalising conditions (see Supplementary Information). Vegetative growth was also assessed at a fixed time, 36.5 days after sowing. Of the twelve parameters that most affected fresh biomass at this timepoint, two parameters directly controlled the water and carbon content of the modelled biomass. Each of these parameters represents a complex physiological process. Eight parameters represented photosynthetic processes and two 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 4h, 6h, 8h, 12h and 18h 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 (Table S7). The simulations also replicated the relevant environmental conditions (see Supplementary Information, Section 3.11; (51)).

Carbon assimilation and respiration rates were slightly underestimated (10.7% and 6% lower in the 12h photoperiod, for example) on the simulated day corresponding to the day of measurement (Fig. 4A). The resulting net carbon fixation allowed the model to reproduce the full amount of starch accumulation by the end of the 12h photoperiod (Fig. 4B), but starch levels were underestimated (by 10-26%) in shorter photoperiods (Fig. 4B). The model closely matched the starch levels remaining at the end of the night (Fig. 4B). However, in short photoperiods, the lower amount of starch accumulation in the light meant that the amount of starch mobilised per night was underestimated in the model. Additionally, part of the mobilised starch was used to maintain a higher sucrose level than observed in the data (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. 4C). The model more closely matched the observed growth increment in the 12h photoperiod (Fig. 4C), where the simulated starch and sucrose levels matched observations (Fig. 4B; 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. 4D). This indicates that further parameters in addition to J_{max}:V_{cmax} are important for modelling growth, especially in the extreme 4h photoperiod. In contrast, the FM accurately predicted the biomass in the 12h photoperiod protocol, to within the experimental error (Fig. 4D). 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 4h and 12h photoperiod, biomass increased strongly and the relative growth rate (RGR, mg FW produced per day per unit existing biomass) increased almost linearly with light fluence (51). This is the response that is expected if the conversion efficiency of carbon into biomass is constant. This linear relation between daily light fluence and growth was lost in long photoperiods. Whereas light fluence increased by 50% between the 12h and 18h photoperiod, RGR increased by only 18% (51). Observed changes in the 18h photoperiod included higher starch levels at dawn and a reduction in specific leaf area (i.e. increased leaf thickness) (51). Both of these are expected to reduce growth rates; incomplete starch mobilisation will sequester carbon from growth, while 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 nighttime 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 14h photoperiod, reproduced the observed biomass (Fig. 4D). This result was also recapitulated by extrapolating J_{max}:V_{cmax} below the published value for 14h photoperiods and reducing starch breakdown, but without considering the increase in leaf thickness (Fig. S5b). Thus these three factors are sufficient to account quantitatively for the altered growth rate under long photoperiods, though 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 non-symmetrical 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 (Fig. S6 and Supplementary Information, Section 3.12). 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-values > 0.05) (Fig. 3C,F and Table S9), while accounting for 11.3% (Ler) and 12.7% (Fei) of the variance in biomass measured at flowering time. The standard deviations 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 standard deviation in leaf initiation (plastochron) reported in Col (12.72; see 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. 3C, 3F, S2). We explored the significance of this developmental timing, by simulating earlier or later transition points (Fig. S7a). 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. 5A). Most leaves were small: median and third-quartile leaf areas fell to 32.6% and 33.5% of the reference value (Fig. 5C). 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. 5B) and, for transition points at 100-400 degree-days, in a greater proportion of large leaves (third quartile area above control; Fig. 5C) that partly escaped shading, resulting in higher biomass than in the controls (Fig. 5A). Simulations with a later transition point, hence a longer interval of slow leaf production, increased biomass (6% increase from transitions at 500-650 degree-day; Fig. 5A). The associated controls increased biomass up to 10.9%. The plant's observed behaviour, represented by the reference transition point, seemed sub-optimal. However, the later transition points reduced the percentage of functional leaves at flowering from 88.9% to 81.8% (Fig. 5B). 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. 5C). Thus the higher biomass of these simulated plants depended upon 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 maximum size, indicating that the proportion of large leaves was high. Taken together, our analysis suggested that increasing the leaf production rate at mid-vegetative 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. 5C). This relationship has been described as a 'dual effect' in plants overexpressing microRNA156 (*Pro35S:MIR156*), which have a short plastochron relative to

wild-type plants (54). To test if our model could reproduce the behaviour of these developmentally-altered plants, and explain the 'dual effect', we grew *Pro35S:MIR156* plants alongside Col wild type for 37 days. Consistent with the previous study, *Pro35S:MIR156* plants had a higher leaf production rate and smaller leaves compared to the wild type (Fig. 5D and S7b). The size of the largest leaf in *Pro35S:MIR156* was only 57% of that in wild type (Fig. 5E). To test whether the leaf production rate alone was sufficient to explain this phenotype, we simulated the growth of the wild type and *Pro35S:MIR156* for 37 days, fixing the leaf production rate in the model to the measured rates in each genotype. 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 the individual leaves in the mutant, including their smaller size relative to the wild type (Fig. 5F and Table S10). As an additional test, we repeated the simulations with the model's simpler, piece-wise linear leaf production rate, using the default values for Col and refitting the piece-wise function to the data for *Pro35S:MIR156* (Fig. S7b,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 wild-type 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 Framework Model.

Discussion

We present methods, examples and validation for one approach to developing a multi-scale, whole-plant model of *Arabidopsis thaliana*, inspired by crop science, by integrating existing models from different laboratories. The resulting Framework Model (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. This suggests that a distributed, community-wide effort could successfully extend and refine the FM by integrating further, focussed models into the larger framework.

Our approach stems from the recognition of potential synergies among diverse plant modellers (57), 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 re-written, 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 12h 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 modelling is overfitting, which becomes more significant in models of high complexity. This 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 re-optimising 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 optimised 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 over-fitted models.

We conducted a sensitivity analysis to examine the behaviour of the FM. This identified 7 out of 18 photosynthetic parameters that are highly sensitive, though this number was likely underestimated because RuBisCo-related parameters 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 parameterised (58, 59). 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 (60). Together, our work and that of others highlight the need for improved systems understanding and mathematical representation to predict plant behaviour 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, Columbia accession. Nonetheless, significant variability among laboratories was recently reported even in standardised 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 12h photoperiods (Fig. 3 and 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, below). 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 (55)), 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-0), because Fei-0 was selected for its increased leaf number in a previous study (48). Simulation of the FM showed that early emergence was sufficient to explain the higher leaf number without altered leaf appearance rate (phyllochron) compared to Col, consistent with our data (Fig. 3). Phyllochron can also easily be determined through observation or automated imaging systems (61, 62) 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, in order to test further regulation by the PPM (6) and PTM (1).

The Framework Model also reproduced the measured biomass of plants grown in 12h photoperiods under slightly different conditions (Fig. 4D), as part of a large, independent data set 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.

4A). This highlights 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 mobilisation at night, and a fixed minimum sugar level. While 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). Firstly, 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), and 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 to long photoperiods (Fig. S5a). Together, these effects caused the model to underestimate growth at night in short photoperiods (Fig. 4C). 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. Secondly, experimental data show that the assumption of almost complete (84%) starch mobilisation 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. 4D). 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 (63-66).

We illustrate the potential of the FM to understand the effect of developmental programmes upon growth and the final rosette form, in four examples. We compared FM simulations with leaf appearance data to discover early seedling emergence in Fei-0 (noted above), and to introduce stochastic leaf production in the model that reproduced the varying leaf number observed in Arabidopsis rosettes (Fig. 3). However, this developmental variation accounted for rather little ($\sim$12%) of the observed variation in rosette biomass. Our simulations of phyllochron (time to leaf appearance) for Ler and Fei had lower standard deviations compared to the standard deviation 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 to plastochron (67), though the two measures were tightly related.

Secondly, 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. On the other hand, 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 to the biomass-maximising, slow-production strategy, this developmental programme distributes carbon investment (and thus leaf size) more widely, a feature characteristic of bet hedging strategies that could be tested in ecological studies (68, 69).

Thirdly, our model reproduced the smaller leaf size phenotype of the developmentally mis-regulated *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: 1) the existence of a "compensatory mechanism" whereby plastochron length and leaf size affect each other reciprocally, so as to reduce changes to the overall plant biomass; or 2) a "common regulator" that controls each of the two traits (54). The combined operation of the CDM and FSPM in the Framework Model provides a parsimonious explanation for the dual effect. High leaf production rate requires carbon resources to be shared among more leaves (Fig. 2F), leading to a decrease in individual leaf growth. Using this mechanism alone, the Framework Model matched the mutant leaf size distribution as accurately as it did the wild type (Fig. 5F). Partitioning of a given amount of carbon among a a larger number of leaves is a sufficient compensatory mechanism (54), though 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 signalling (70-72) to whole-plant phenotypes, and extend to applications that modulate organ size, for example in pruning (73, 74).

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 utilisation 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 Framework Model, and the approach used to build it, provides 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. This is particularly important to understand the pervasive effects of environmental stresses or pleiotropic biological regulators, such as the circadian clock. Finally, multi-scale digital plant models might contribute to link systems biologists with ecophysiology and crop science, where significant synergies may be gained.

Acknowledgements.

We thank C.I. Gracia for assisting in the experiments, Tomasz Zielinski and Martin Beaton for infrastructure support and Ronan Sulpice for the Fei-0 (Fei) seeds. Y.H.C. was a recipient of the Darwin Trust PhD studentship. This study was supported by the Biotechnology and Biological Sciences Research Council awards BBR (BB/F010605/1 to HJO, AJM and RM) and ROBuST (BB/F005237/1 to KJH, MW, AJM and others) and by the European Commission FP7 collaborative project TiMet (contract no. 245143 to AJM, MS and others). **Author Contributions** K.J.H., H.T., H.J.O., R.M. and A.J.M. conceived the study. Y.H.C., B.W., M.S., M.W., K.J.H. and A.J.M. designed the experiments. Y.H.C., B.W., A.F. and V.M. performed the experiments and analysed the experimental data. Y.H.C. refactored and integrated the models in MATLAB, performed the simulations and analysed the simulation data. C.T. constructed the PlaSMo repository. C.L.D., P.de R., R.M. and J.T. refactored the models in Simile and validated them. Y.H.C., R.M., M.W., and A.J.M. wrote the manuscript with input from all authors.

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. Lopez-Bucio J, Cruz-Ramirez A, & Herrera-Estrella L (2003) The role of nutrient availability in regulating root architecture. *Curr Opin Plant Biol* 6(3):280-287.
12. Yin XY & 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 C-3 photosynthesis from the chloroplast to the ecosystem. *Plant Cell Environ* 36(9):1641-1657.
17. Nedbal L, Cerveny J, & Schmidt H (2009) in *Photosynthesis in silico*, eds Laisk A, Nedbal L, Govindjee (Springer, 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 Model* 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 Climate* 19(11):2122-2143.
23. Overstreet CM, Nance RE, & Balci O (2002) Issues in enhancing model reuse. *International Conference on Grand Challenges for Modeling and Simulation*, 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. *Environmental Modelling & Software* 39:39-49.
26. Hunter PJ & Viceconti M (2009) The VPH-Physiome Project: Standards and Tools for Multiscale Modeling in Clinical Applications. *IEEE Reviews in Biomedical Engineering* 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 SB & 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-U127.
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 Sys Biol* 1:13.
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. Mendez-Vigo B, de Andres MT, Ramiro M, Martinez-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 For 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* 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* 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. 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.
56. Granier C, *et al.* (2002) Individual leaf development in Arabidopsis thaliana: a stable thermal-time-based programme. *Ann Bot* 89(5):595-604.
57. Thomas H (2008) Systems biology and the biology of systems: how, if at all, are they related? *New Phytol* 177(1):11-15.
58. 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.
59. 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.
60. Asseng S, *et al.* (2013) Uncertainty in simulating wheat yields under climate change. *Nature Climate Change* 3(9):827-832.
61. Tisne S, *et al.* (2013) Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. *Plant J* 74(3):534-544.
62. 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.
63. Scialdone A, *et al.* (2013) Arabidopsis plants perform arithmetic division to prevent starvation at night. *eLife* 2:e00669.
64. Seaton DD, Ebenhöh O, Millar AJ, & Pokhilko A (2014) Regulatory principles and experimental approaches to the circadian control of starch turnover. *J Royal Soc Interface* 11(91):20130979.
65. Pokhilko A, Flis A, Sulpice R, Stitt M, & Ebenhoh 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.
66. Pal SK, *et al.* (2013) Diurnal changes of polysome loading track sucrose content in the rosette of wildtype Arabidopsis and the starchless *pgm* mutant. *Plant Physiol* 162: 1246-1265.
67. Clerget B, Dingkuhn M, Goze E, Rattunde HFW, & Ney B (2008) Variability of phyllochron, plastochron and rate of increase in height in photoperiod-sensitive Sorghum varieties. *Ann Bot* 101(4):579-594.
68. Simons AM (2009) Fluctuating natural selection accounts for the evolution of diversification bet hedging. *Proceedings of the Royal Society B-Biological Sciences* 276(1664):1987-1992.
69. Simons AM (2011) Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proceedings of the Royal Society B-Biological Sciences* 278(1712):1601-1609.
70. Wahl V, *et al.* (2013) Systemic regulation of flowering by trehalose-6-phosphate signaling in *Arabidopsis thaliana*. *Science* 339: 704-707
71. 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 2:e00260
72. Yu S *et al.* (2013) Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. eLife 2:e00269
73. 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.
74. 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.