



RESEARCH PAPER

A multi-model framework for the Arabidopsis life cycle

Argyris Zardilis^{1,*}, Alastair Hume^{1,2,*} and Andrew J. Millar^{1,*}

¹ SynthSys and School of Biological Sciences, University of Edinburgh, Max Born Crescent, Edinburgh, EH9 3BF, UK

² EPCC, University of Edinburgh, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, UK

* Correspondence: andrew.millar@ed.ac.uk

Received 28 July 2018; Editorial decision 10 October 2018; Accepted 20 November 2018

Editor: Bertrand Muller, INRA, France

Abstract

Linking our understanding of biological processes at different scales is a major conceptual challenge in biology and aggravated by differences in research methods. Modelling can be a useful approach to consolidating our understanding across traditional research domains. The laboratory model species Arabidopsis is very widely used to study plant growth processes and has also been tested more recently in ecophysiology and population genetics. However, approaches from crop modelling that might link these domains are rarely applied to Arabidopsis. Here, we combine plant growth models with phenology models from ecophysiology, using the agent-based modelling language Chromar. We introduce a simpler Framework Model of vegetative growth for Arabidopsis, FM-lite. By extending this model to include inflorescence and fruit growth and seed dormancy, we present a whole-life-cycle, multi-model FM-life, which allows us to simulate at the population level in various genotype × environment scenarios. Environmental effects on plant growth distinguish between the simulated life history strategies that were compatible with previously described Arabidopsis phenology. Our results simulate reproductive success that is founded on the broad range of physiological processes familiar from crop models and suggest an approach to simulating evolution directly in future.

Keywords: Agent-based modelling, Arabidopsis, computational modelling, ecophysiology, growth model, life history, population ecology, systems biology.

Introduction

Understanding the links between biological processes at multiple scales, from molecular regulation to populations and evolution, is a major challenge in understanding life. As systems become more complex we need models to describe our understanding and help our thinking when trying to explain and make predictions across scales. This approach has also been proposed in attempts to engineer crop traits starting from genetics or from genomes (Welch *et al.*, 2005; Yin and Struik, 2008, 2010; Parent and Tardieu, 2014; Wu *et al.*, 2016; Chenu *et al.*, 2018), where simpler models have demonstrated both the potential of crop modelling in general and the significant demands of detailed models for empirical data that

vary in availability (Hammer *et al.*, 2006; Asseng *et al.*, 2013). For micro-organisms, comprehensive models link the metabolic and molecular level with the cellular (Karr *et al.*, 2012) and population scales (Weiße *et al.*, 2015), whereas contemporary work in more complex organisms has necessarily focused more narrowly (Buckley and Mott, 2013; Lynch, 2013; Zhu *et al.*, 2013; Klose *et al.*, 2015; Le Novère, 2015; Hepworth *et al.*, 2018).

The concentration of plant science research on the laboratory model species Arabidopsis offers an opportunity for broad understanding that includes mechanistic models (Chew *et al.*, 2014a; Voss *et al.*, 2014; Millar *et al.*, 2019). The Framework

Model (FMv1; Chew et al., 2014b) represented vegetative growth of Arabidopsis in lab conditions, starting from four independent models that represent photosynthesis and carbon storage (Rasse and Tocquin, 2006), plant structure and carbon partitioning among organs (Christophe et al., 2008), flowering phenology (Chew et al., 2012), and the circadian clock gene circuit and its output to photoperiodic flowering (Salazar et al., 2009). Later updates focused on plant phenotypes controlled by the clock, such as tissue elongation and starch metabolism (FMv2; Chew et al., 2017, Preprint), or temperature and organ-specific inputs to flowering (Kinmonth-Schultz et al., 2018, Preprint). The Framework Models align with community efforts to link understanding of crop plant processes at multiple scales, for benefits in agriculture (Wu et al., 2016; Zhu et al., 2016). Among the limitations of the Framework Models, growth was limited to the vegetative stage, ending upon flower induction. Without reproduction, the models had no seed yield or link to evolutionary fitness. Without seed dormancy, they lacked a major determinant of Arabidopsis life history in the field. Their representation of the circadian clock was also unnecessarily detailed for many studies outside chronobiology.

Other models have considered reproductive success through growth, including for Arabidopsis. One simplified approach relates growth and fitness only to the duration of the developmental period and not to its timing in the year, ignoring environmental influences (Prusinkiewicz et al., 2007). On the other hand, ecological phenology models of the Arabidopsis life cycle consider natural environmental conditions but ignore physical growth and development (Chuine and Beaubien, 2001; Chuine, 2010; Burghardt et al., 2015). We apply a declarative agent-based modelling approach (Honorato-Zimmer et al., 2018) that facilitates model composition, to develop FM-life, an extension of the Framework Model to the whole Arabidopsis life cycle. FM-life includes a simpler model of vegetative growth, FM-lite, without the clock circuit, and a new model of inflorescence growth including reproduction. We introduce a clustering approximation, in order to simulate FM-life tractably at the population scale over decades. Testing the FM-life model with contrasting environmental and genetic inputs shows that ecological questions can increasingly be informed by mechanistic understanding of growth processes (Millar, 2016; Doebeli et al., 2017).

Materials and methods

Chromar

FMv1 and FMv2 are both available as MATLAB programs. Here we use a declarative, agent-based language, Chromar, which represents the models very concisely and supports simulations (Honorato-Zimmer et al., 2017). Chromar models are designed to be human-readable. As no agent-based approach is broadly familiar in biological modelling, we introduce simple examples below. Briefly, agents in Chromar conform to a small number of types. Each type defines the attributes of the agent. A type defined as Leaf (mass:real) represents leaves with a mass attribute that is a real number, for example Leaf (mass=5). Agents operate stochastically, according to rules that specify:

- (i) Agent level dynamics, where agents are created or destroyed. For example, organogenesis of a leaf with mass m_0 at rate k corresponds to the following rule: $\emptyset \xrightarrow{k} \text{Leaf}(\text{mass} = m_0)$.

- (ii) Attribute level dynamics, which alter the attributes of an agent. For example, growing a leaf by mass g at rate k_g corresponds to: $\text{Leaf}(\text{mass} = m) \xrightarrow{k_g} \text{Leaf}(\text{mass} = m + g)$.

Two further features of Chromar are particularly useful here, *fluents* for describing time-dependent, deterministically changing values such as environmental inputs, and *observables* used for capturing system-wide state easily. Observables are used to manage multiple descriptions of the same process at different levels of abstraction, such as the individual leaves of a plant and the total number of leaves in a rosette. Fluents and observables can be used directly in expressions and be combined with the normal set of mathematical expressions, so we could have a function, $f(n, \text{temp})$, that takes as arguments and applies some mathematical operation to an observable for the number of leaves, n , and a fluent, temp , for the temperature.

Here we made use of a particular implementation of Chromar as an embedded domain-specific language inside the general purpose programming language Haskell (Honorato-Zimmer et al., 2018). This means that we can leverage the power of a general-purpose programming language inside the rules, for example for describing attribute dynamics or rule rates.

Phenology models in Chromar

In many phenology models, the simulated plant accumulates a conceptual development indicator in every time unit as a function of the contributing environmental factors, until a threshold is reached for transition to the next developmental stage. For example, in a seed type Seed(dev:real), the dev attribute measures development towards germination. A phenology rule for germination affected by temperature and moisture, starting from dev value d , could be:

$$\text{Seed}(\text{dev} = d) \xrightarrow{1.0} \text{Seed}(\text{dev} = d + f(\text{temp}, \text{moist}))$$

On average once every time unit the dev attribute of a particular seed will be increased from the present value, d , by a function of the contributing factors temp and moist. Further parameters might represent how sensitive the seed is to the environmental factors. At the threshold D_g , the seed germinates to a plant and resets the development measure to 0:

$$\text{Seed}(\text{dev} = d) \xrightarrow{1.0} \text{Plant}(\text{dev} = 0) [d > D_g]$$

where the expression inside the square brackets is used to indicate conditional activity of the rule. The rule is active only when the expression evaluates to true.

The component models

The models presented here represent the full life cycle in three stages: seed dormancy (A in Fig. 1A), vegetative growth up to flowering (B in Fig. 1A), and the reproductive stage up to seed dispersal (C in Fig. 1A). Each model (A, B, C) includes a phenology component that represents only timing. The vegetative and reproductive stage models also represent biomass growth at the organ level, based on the carbon budget of the plant. We varied genetic parameters that affect only the timing components of A (seed dormancy, ψ_i) and B (floral repression during vegetative growth, f_i), for comparison to Burghardt et al. (2015). Each parameter value for an individual plant, i , can be fixed or selected probabilistically from a distribution as described (Burghardt et al., 2015). The three models were integrated in a whole life-cycle model of one plant (FM-life), and then extended to a population of such plants.

Seed dormancy model (A)

The seed dormancy model is the Chromar version of the model of (Burghardt et al., 2015), which is based in turn on (Alvarado and Bradford, 2002). It represents the development of a newly dispersed seed

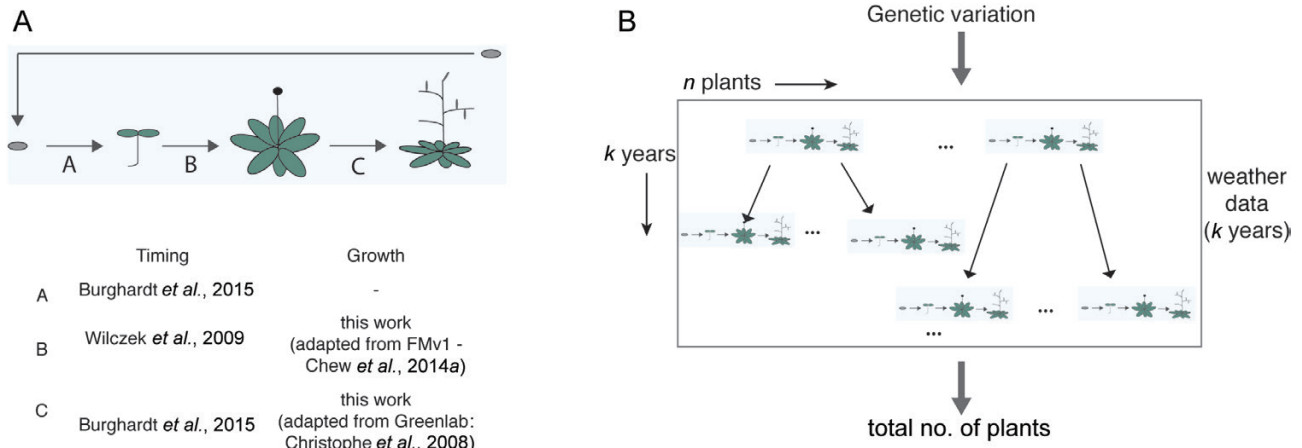


Fig. 1. Overview of the FM-life and population models used in this study. (A) Overview of the models used for the growth and timing components for the three developmental stages: seed dormancy (a), vegetative period (b), and reproductive period (c). (B) Diagram of the population level model. Inputs to the model are the distribution of values of the two genetic parameters (f_i , ψ_i) and weather data from some location for a number of years. The output is some population measure of interest; an example might be the total number of plants after k years.

from $\text{dev}=0$ to a threshold value, D_g , where the seed germinates. Above baseline levels of temperature T_b and of moisture (see below), increasing moisture and temperature speed the progress towards germination. The additional developmental units added (hydrothermal units, htu) at every time unit are described by:

$$\text{htu}(t) = \begin{cases} (\Psi(t) - \Psi_b(t)) \times (T(t) - T_b) & \text{if } T_b < T(t) \leq T_o \text{ and } \Psi_b(t) < \Psi(t) \\ (\Psi(t) - \Psi_b(t)) \cdot (T_o - T_b) & \text{if } T(t) > T_o \text{ and } \Psi_b(t) < \Psi(t) \\ 0 & \text{otherwise} \end{cases}$$

where $\Psi(t)$ and $T(t)$ give the moisture and temperature levels at time t , respectively. The definition distinguishes between operating in sub-optimal and supra-optimal temperatures (below or above the optimum, T_o , respectively). The baseline moisture is used to represent the dormancy level of the seed. If Ψ_b is high, the seed accumulates htu slowly for a given set of environment conditions, whereas if Ψ_b is low, development is faster in the same conditions. From an initial dormancy level, ψ_i , seeds lose dormancy (Ψ_b becomes smaller) over time at a rate r that is also a function of the environmental conditions, moisture and temperature, and represents the observed process of after-ripening. ψ_i is also used to represent the genetic effect on dormancy, where high ψ_i represents stronger dormancy.

In Chromar, the Seed type captures information about the seed development process: Seed(gntp:(real,real),dev:real,r:real). The gntp attribute stores the genotype of the organism, ψ_i (seed dormancy level) and f_i (floral repression level), which is passed on to the agents representing the later stages of development and transmitted unchanged to the next generation. dev stores the cumulative development indicator (sum of htu up to the current time point), and r stores the after-ripening up to the current time point. The development rule is the following:

$$\text{Seed}(gntp = a, \text{dev} = d, r = r) \xrightarrow{1.0} \text{Seed}(\text{dev} = d + \text{htu}(\text{temp}, \text{moist}, f(r, \text{temp}, \text{moist}), a, \psi_i), r = f(r, \text{temp}, \text{moist}))$$

where *temp* and *moist* are fluents describing temperature and moisture. We use the 'dot' (.) operator in the expression above for accessing the two genetic parameters of the gntp attribute. The following rule represents germination, starting the vegetative stage:

$$\text{Seed}(gntp = a, \text{dev} = d, r = r) \xrightarrow{1.0} \text{Plant}(gntp = a), \text{Root}(\dots), \text{Leaf}(\dots), \text{Leaf}(\dots) [d > D_g]$$

The abstract Plant agent represents the plant at the vegetative stage, along with agents for the root and the two cotyledon leaves. The initial configuration of the organs at germination is as introduced by Chew *et al.* (2014b). Note that the genotype attribute is passed from seed to emerged plant unchanged.

We can easily change the time step of the model by changing the rate and development update size. For example, in our population simulations later for efficiency reasons we use a daily time step instead of the hourly one of the original model:

$$\text{Seed}(gntp = a, \text{dev} = d, r = r) \xrightarrow{1.0/24.0} \text{Seed}(\text{dev} = d + 24 \times \text{htu}(\text{temp}, \text{moist}, f(r, \text{temp}, \text{moist}), a, \psi_i), r = f(r, \text{temp}, \text{moist}))$$

This means that we make one bigger update every day instead of smaller ones every hour. The expected update to the developmental sum is the same ($\text{htu}(\dots)$) but the variability is increased.

In order to illustrate the functioning of the model we show simulations of the htu accumulation over a year for real climate data in Valencia, Spain for different values of the initial dormancy level, ψ_i . There are two periods during the year that are favourable for development towards germination, spring and early autumn (Fig. 2B). Lowly dormant ($\psi_i=0$) and medium dormant seeds ($\psi_i=1.25$) sown in the spring manage to accumulate enough development to both germinate in autumn. Highly dormant seeds, on the other hand, progress slower and miss the autumn window of favourable weather and therefore have to wait until spring in the following year. This shows a case where differences in genetics (continuous parameters) leads to qualitatively different timings that affect the current and subsequent lifecycles.

Vegetative growth model (FM-lite) (B)

For the vegetative stage we introduce a simplified version of FMv1 (Chew *et al.*, 2014b) for use in studies that do not focus on circadian timing. FM-lite has three constituent models represented in Chromar with modifications to environmental responses (see below), and without the fourth, circadian clock model of FMv1.

Timing

The timing component is the simpler flowering phenology model of (Wilczek *et al.*, 2009) rather than the augmented version in FMv1 (combination of Chew *et al.* (2012) with Salazar *et al.* (2009) models). Vegetative development extends from $\text{dev}=0$ to a threshold value, D_f , where the plant flowers. The main contributing environmental factors are photoperiod, ambient temperature, and vernalization, giving the modified photothermal units, mptu , at a time t as:

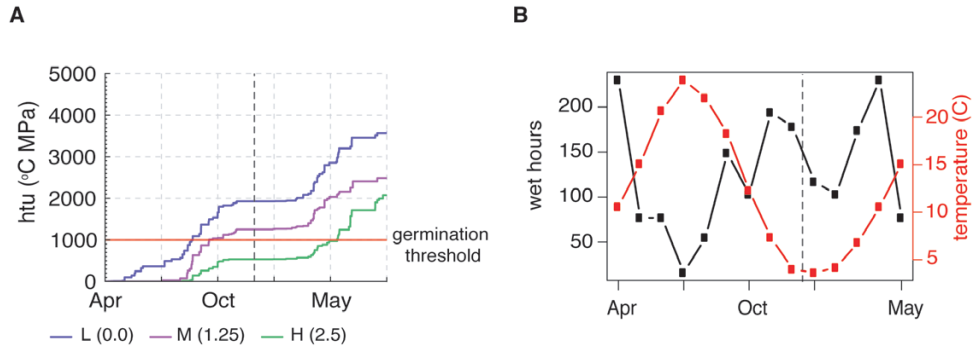


Fig. 2. Differences in germination behaviour of seeds sown on 10 April in Valencia (Spain) (A) Each line represents the progress towards germination (hydrothermal units) of seeds with different genetic backgrounds (different values of initial dormancy, ψ_i) where each line is the average of 10 runs. Small differences due to the initial dormancy value can lead to qualitatively different behaviour. The low (L, blue line) and medium (M, purple line) dormant seeds both germinate in early autumn while the highly (H, green line) dormant seeds miss the favourable autumn season and germinate in the spring in the following year (see also (B)) (B) Temperature and moisture (determinants of germination) during the same period in Valencia. Wet hours is the number of hours per month with at least some moisture in the soil. This is for illustration only and for the simulations in (A) the hourly values of soil moisture were used instead. The black line is the monthly average temperature as used in the simulation of development in (A).

$$mptu(t) = photoperiod(t) \cdot thermal(t) \cdot vernalisation(t)$$

The vernalization term both accounts for the observed requirement for a specific duration of exposure to cold and is also used to represent the genetic effect on the progress towards flowering, modelled as $vernalisation(t) = f(wc, f_i)$, where wc is the exposure to cold accumulated up to t and f_i is the genetic parameter for the initial floral repression, as in Wilczek et al. (2009).

In Chromar, the plant type: `Plant(gntp:(real,real),dev:real,wc:real)` includes the genotype attributes as noted above, the development so far (`dev`), and finally the accumulated winter chilling (`wc`). The development rule is then:

$$\begin{aligned} & \text{Plant}(gntp = a, dev = d, wc = w) \xrightarrow{1.0} \\ & \text{Plant}(dev = d + mptu(temp, dl, af_i, w), wc = f(w)) \end{aligned}$$

where $temp$ and dl are fluents for temperature and day length respectively, and w is the present value of wc . The transition to a flowering plant, `FPlant`, follows:

$$\text{Plant}(gntp = a, dev = d, wc = w) \xrightarrow{1.0} \text{FPlant}(gntp = a) \quad [d > D_f]$$

Similar time step changes are possible by changing the rate and update size as discussed above for the seed dormancy timing model.

Growth

As in FMv1 (Chew et al., 2014b), the growth component includes a carbon budget for the plant from Rasse and Tocquin (2006), which in turn includes photosynthesis rate equations based on the Farquhar et al. (1980) model. Growth at the organ level (rosette leaves and root) is represented based on the Greenlab model (Christophe et al., 2008). We will consider a sucrose carbon pool (c), a starch carbon pool (s), and one pool for the biomass of the root and each of the rosette leaves (Fig. 2A). In Chromar we have the following agents to store the state (amount of carbon, or total biomass) of these pools:

- (i) `Cell(c,s:real)`. An agent that stores the amount of carbon in the sucrose (c attribute) and starch pools (s attribute). The amounts are carbon totals at the whole plant level.
- (ii) `Leaf(m:real,i:int)`. An agent that represents a rosette leaf. It has attributes for its mass (m) and its index of appearance (i).
- (ii) `Root(m:real)`. An agent that represents the root with an attribute for its mass (m).

For each organ we have a growth flow from the sucrose carbon pool to the mass of the organ (growth rule, Fig. 3). The growth amount depends on the demand function of the organ ($d(i,t)$ rule rate function) and its ‘sink strength’ ($g(m)$), which varies among organs. The value of the demand function varies over time between 1 (maximum demand) and 0 (no demand) at the end of the expansion period of the organ and it is used to time growth. The sink strength function $g(m)$, which is fixed throughout the life of an organ, gives the magnitude of growth of each organ type relative to other organ types. The amount of carbon requested by an organ at every time unit is $g(m) \times d(i,t)$. Depending on the metabolic status of the whole plant (level of c pool) and the requests from other organs, an organ will receive either the full expected amount or a portion of it. This view of growth is a simplification that works well for the range of lab conditions that the models were developed in but may not work as well for the full range of natural conditions we consider later. For example, while we consider sink strengths to be constant here, drought, considered in models of crop species, might change the relative sink strengths to allocate more assimilate to the root, which will change the carbon balance and affect whole plant growth (Kage et al., 2004).

A flow in the opposite direction (`mobl` rule, Fig. 3) represents carbon mobilization from the organs if the central sucrose pool (`Cell(c)`) is reduced to a critical level. Thus, each organ can be either a net sink or source of carbon. For each organ, we also have a flow leaving the system from the c pool for the cost of the maintenance respiration and other processes of the organ (`maint` rule, Fig. 3). Photosynthetic carbon fixation is represented by the assimilation process (`assim` rule, Fig. 3). The amount of assimilate at every time unit is the product of the photosynthesis rate, which is a function of environmental conditions at that time step, and the projected area of the rosette. Here we use an observable, a_{ros} , for the effective rosette area, which is a function of the global state of the rosette at the current time (derived from the masses of all the current leaves) and takes into account the effect of shading, as in Chew et al. (2014b). The carbon partitioning function includes a baseline partitioning to starch, then support of a target sucrose level, with excess sucrose supporting growth and a final overflow to additional starch production, as in Chew et al. (2014b). At night, no photosynthesis occurs and carbon from the starch pools flows to the sucrose pool (`sdgr` rule, Fig. 3). Finally, we have the creation of new leaves, which impacts the above processes indirectly by creating more demand for growth and adding maintenance costs (`leaf cr` rule, Fig. 3). Leaves are created by the main apical meristem (`Vaxis` agent) along with an `LAxis` agent that can give rise to lateral branches after flowering (see next section).

It is interesting to note that in FMv1, carbon partitioning between processes and organs is done explicitly whereas in our Chromar representation, partitioning is an emergent, stochastic effect of competition for

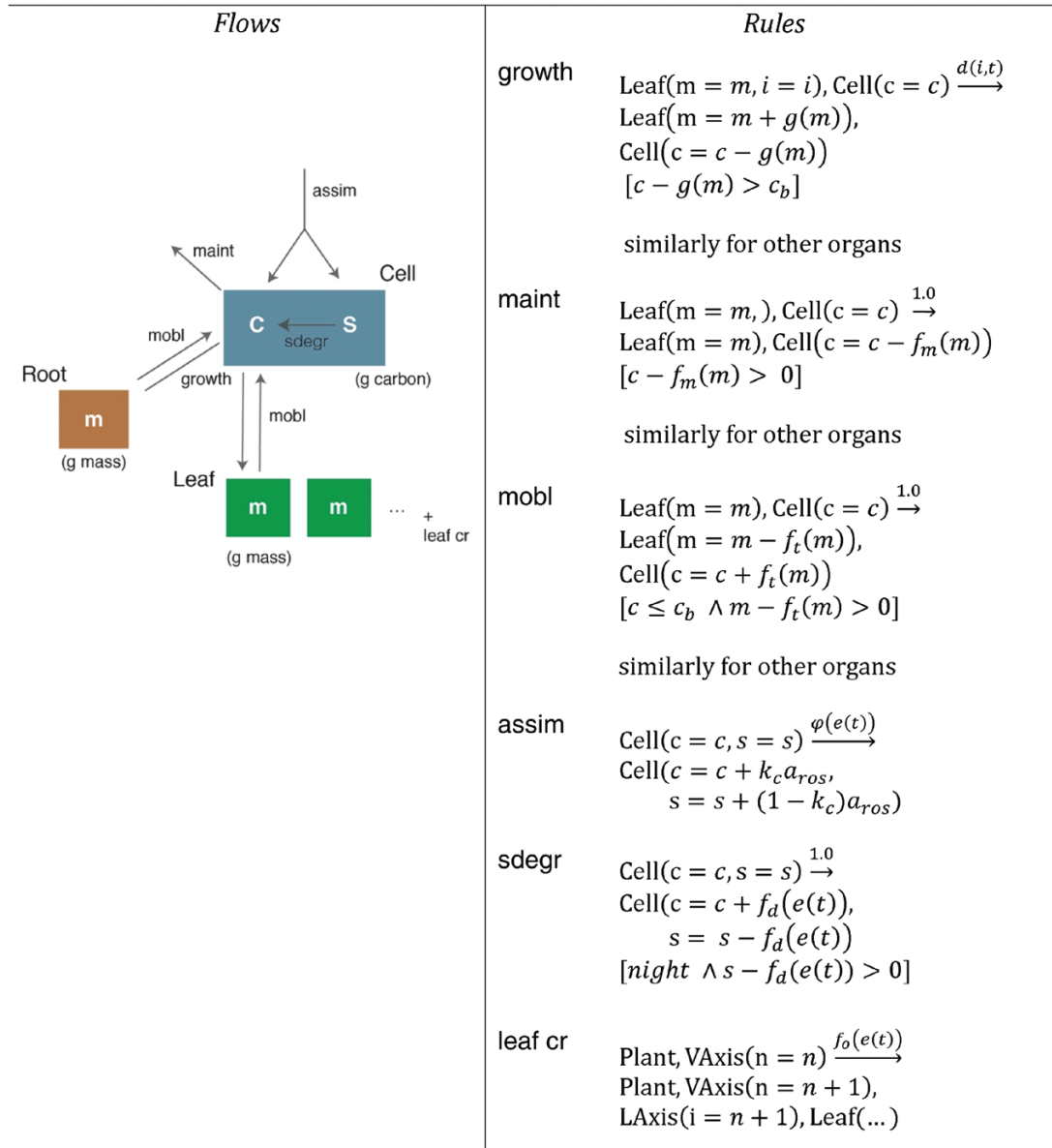


Fig. 3. An overview of the dynamics on FM-lite (growth component of vegetative stage). The dynamics take the form of flows between different reservoirs of carbon, here shown in a graphical way (left) along with the corresponding Chromar rules (right).

the finite amount of sucrose carbon in the main reservoir. For example, partitioning of carbon among organs for growth is done explicitly in FMv1 by dividing the demand of each organ by the sum of the demands of all other organs:

$$g(m) \times \frac{d(i, t)}{\sum d(i, t)}.$$

In the Chromar representation we do not have this explicit division by the global demand, which means that the amount of carbon that an organ gets is higher at each growth event but growth events are rarer because not all growth request are successful (competition). The competition therefore recovers the explicit partitioning.

Modifications for natural conditions

FMv1 was developed for lab conditions. As an initial approach to reflect plant responses to the broader range of relevant conditions in nature, we made the following changes:

- (i) The rate of photosynthesis is set to 0 below 0 °C.
- (ii) The maintenance cost for an organ is also 0 below 0 °C.
- (iii) The rate of photosynthesis is affected by soil moisture through stomatal closure. The photosynthesis rate is affected by a stomata term $f_{\text{stom}}(\text{moist})$, which is a simple phenomenological function that relates soil moisture and stomatal closure (France and Thornley, 1984).

These conservative changes give a lower bound on the effects of natural weather conditions.

Comparison of FM-lite with FMv1

In addition to the weather responses, Wilczek flowering model, and emergent carbon partitioning among organs, our model representation uses the stochastic rule-based Chromar as opposed to the deterministic MATLAB program of FMv1. In order to compare the model representations, we simulated growth in the two models for a fixed number of hours in lab conditions, where the modifications to weather responses have no effect. The two models were simulated in lab conditions (22 °C, 12/12 light/dark cycles) for 800 growth hours and showed

comparable results (Fig. 4). FMv1 was simulated in MATLAB while FM-lite was simulated in the Haskell implementation of Chromar and the results were averaged over five runs. The rosette mass results are the closest since they represent the development of multiple Leaf agents, masking the stochastic effects on each Leaf. The difference between the final rosette mass of the FMv1 and FM-lite (averaged over five runs) simulations is within 10% of the final rosette mass in FMv1. The stochasticity is more apparent for the root where the growth curves are further apart. The difference between final root mass in FMv1 and FM-lite (averaged over five runs) is ~20% of the final root mass in FMv1. Sucrose carbon levels are also more variable in FM-lite, since the growth rule (removing sucrose carbon from the central pool) provides organs with a larger amount but less frequently than the small fixed amount at every time step in FMv1 (see previous section).

Reproductive stage model (C)

Timing

The timing component is a thermal time model from Burghardt et al. (2015), representing the development of the inflorescence and seed from $dev=0$ at flowering, to a threshold value, D_s , where the plant disperses its seeds. Here there is no genetic input and the thermal units that accumulate at time t are simply the value of the temperature at t above a base temperature T_b :

$$tu(t) = \begin{cases} T(t) - T_b & \text{if } T(t) > T_b \\ 0 & \text{otherwise} \end{cases}$$

Writing into Chromar we have an `FPlant(dev:real)` type for a flowered plant and the following rule for its development that follows from the tu definition above:

$$FPlant(dev = d) \rightarrow FPlant(dev = d + tu(temp))$$

Finally, the transition to seed happens when the accumulated development reaches D_s :

$$FPlant(attr = a, dev = d) \rightarrow Seed(attr = a, dev = 0, r = 0) [d > D_s]$$

Note that the genotype attribute of the parent plant is transferred to the seeds unchanged. Like the seed dormancy and flowering timing models we can again change the time step by modifying the rate and update size.

Growth

The growth component of the reproductive stage model is loosely related to the Greenlab model (Christophe et al., 2008). The metabolic processes affecting the carbon budget of the plant are the same as in vegetative growth but with additional organ types to represent the Arabidopsis inflorescences. Organs appear in units (metamers) with a metamer identifier. Each growth unit on the main axis consists of an internode (stem between leaves), a leaf, and a lateral meristem that can give rise to a lateral axis. We consider only the primary axis and secondary, lateral branches, thus metamers on the lateral axis lack a further lateral meristem. All fruits on an axis are represented on its last metamer, replacing the leaf; this metamer also lacks a meristem. Two indices represent metamer position:

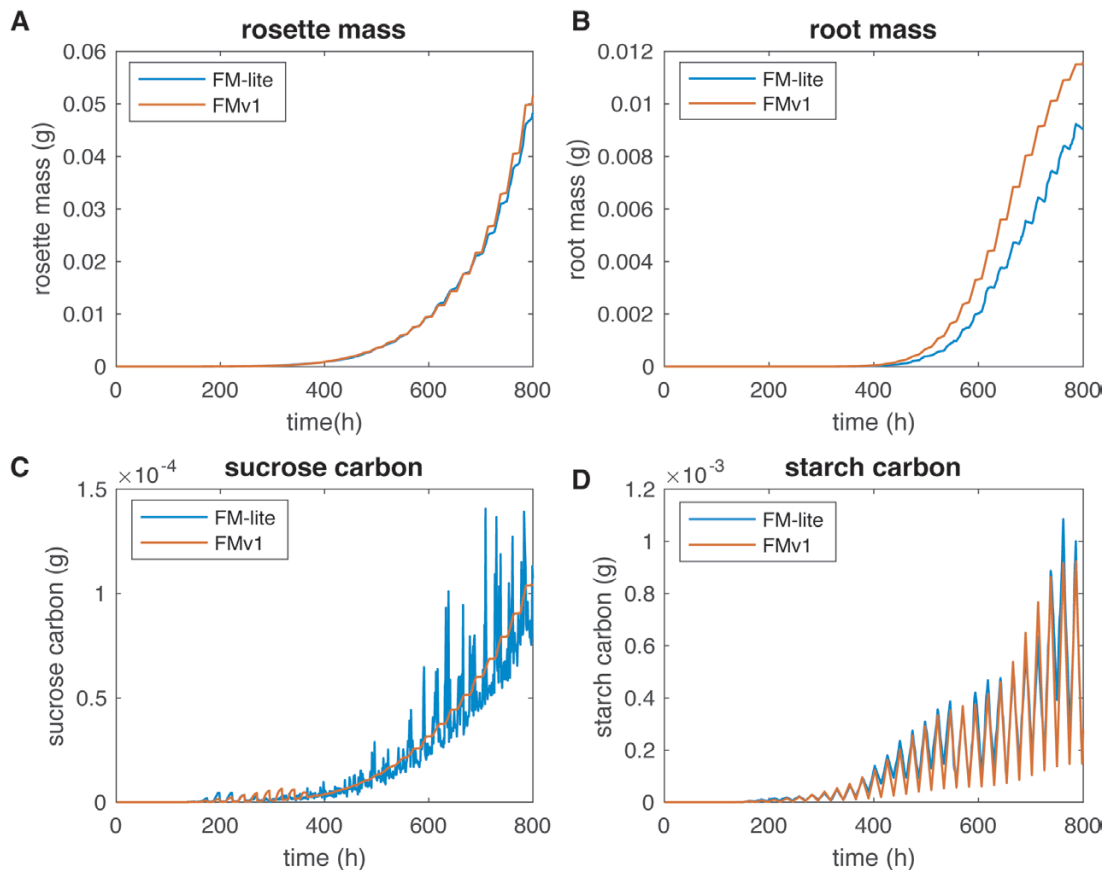


Fig. 4. A comparison of the original FM implementation (FMv1) with the adaptation used in this work (FM-lite) for 800 h of growth. FM-lite simulations were performed in the Haskell implementation of Chromar and results were averaged over five runs. FMv1 simulations were carried out in the MATLAB environment. (A) Comparison of simulated rosette mass trajectories between FMv1 and FM-lite. (B) Comparison of simulated root mass trajectories between FMv1 and FM-lite. (C) Comparison of simulated sucrose carbon between FMv1 and FM-lite. (D) Comparison of simulated starch carbon between FMv1 and FM-lite.

the index of the metamer along its axis and the index of the parent metamer along the primary axis (left part of Fig. 5). We define the following new agent types to represent this structure:

- (i) $\text{INode}(i, pi: \text{int}, m: \text{real})$ to represent the internode (stem between successive leaves). Attribute i is the temporal index of appearance in its axis (primary or lateral) and attribute pi is the parent primary metamer. The cotyledons have indices 1 and 2 on the primary axis, for example.
- (ii) $\text{LLeaf}(i, pi: \text{int}, m: \text{real})$ to represent a leaf on the lateral axes.
- (iii) $\text{Fruit}(i, pi: \text{int}, m: \text{real})$ to represent a fruit on the axis.

The maximum number of inflorescence metamers on the main axis is taken to be 20% of the number of rosette leaves at flowering time (n_f) and given by $v_{\max}(n_f)$ (Pouteau and Albertini, 2009). The maximum number of growth units on each lateral axis is given by $l_{\max}(i)$, a decreasing function of the index of the lateral axis starting from a maximum of 6 at the axis after the cotyledons (index 3) and going to a minimum of 1 at the topmost lateral branch (Mündermann *et al.*, 2005). The topmost lateral axis can only appear with a delay after the apical fruit has appeared on the primary axis. Each successive lateral branch going down can only start developing with a delay after the fruit of the axis above it has appeared. The delay associated with lateral axis growth, given in the rules by t_{del} , is a function of the metabolic state of the plant, as described (Christophe *et al.*, 2008).

The new organ types have associated sink strengths and demand functions. The cauline leaves on the main axis contribute to the photosynthetically active area and can shade the rosette leaves underneath them. The lateral leaves contribute to photosynthesis without shading. Internodes and fruits do not contribute to photosynthesis. Seeds are not directly represented, so a birth function $b(m)$ is required to calculate the number of

seeds for a given fruit mass m at seed dispersal time, as described below. The root demand function has been modified relative to FMv1 so that it has two parts, one in the vegetative stage that coincides with rosette growth and one in the reproductive stage that coincides with inflorescence and fruit growth.

Whole life cycle model, FM-life

The Chromar framework allows us simply to concatenate the rules of timing and growth components of the three models above, to represent the whole life cycle. Then given an initial state with the genetic attributes of the plant (gnpt attribute of agents) and the environmental conditions for a particular location, $e(t)$, we can simulate an entire life cycle from seed to seed. The timing components of the model give us the timing within the year of the growth period (vegetative + reproductive stages) and therefore the environmental conditions that the plant is exposed to during growth. The growth components predict growth at the individual organ level with these environmental conditions and therefore give us the environmentally determined seed number given by the $b(m)$ function. The model files are available at: https://github.com/azardilis/ChromarFM/releases/tag/fm-life_v1.0.

Population level model and plotting conventions

Since FM-life estimates the number of seeds at the end of the life cycle, these can initiate multiple independent copies of the model in the next generation. We then have a classical evolutionary birth process, sometimes called a branching process since it unfolds in a tree-like way. The potential number of individuals in generation i , n_i , is equal to the sum of the number of seeds produced by the individuals in the previous generation

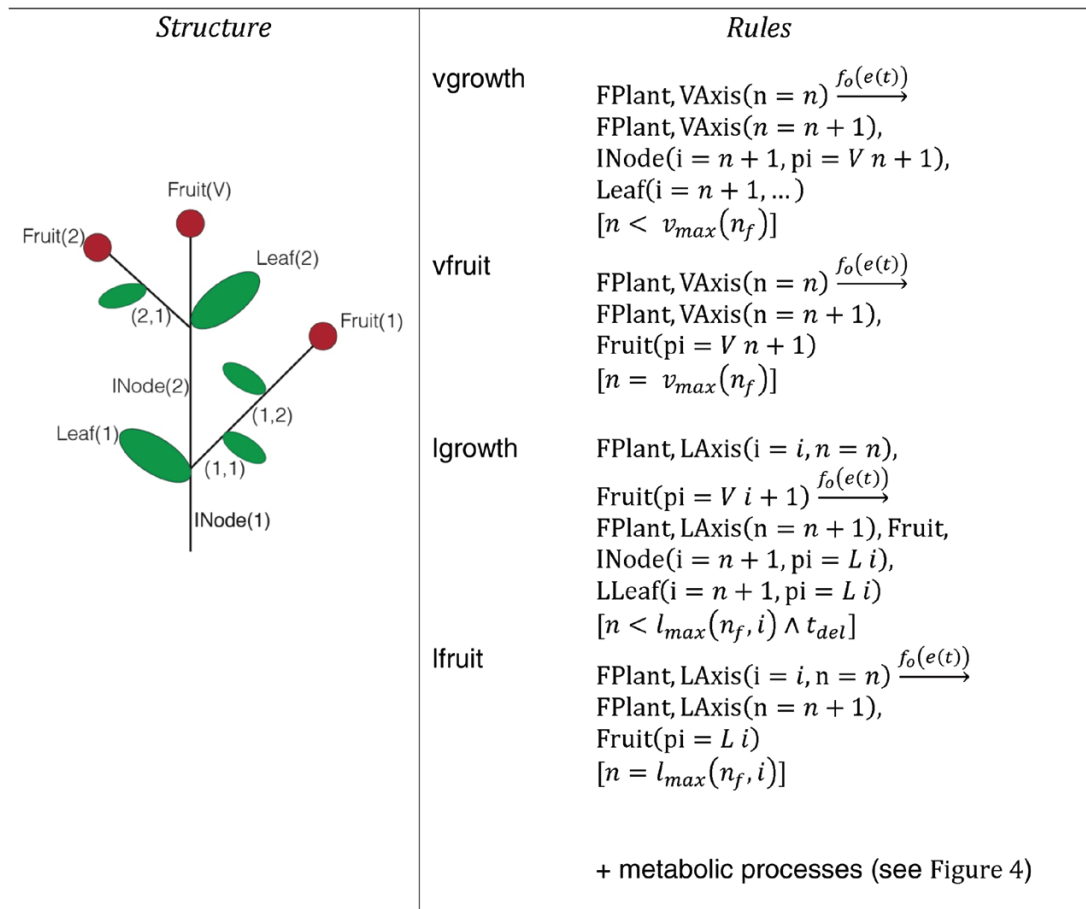


Fig. 5. Overview of the structural part of the growth component of the reproductive stage model. The numbering scheme used to keep track of the positions of the organs in the inflorescence architecture is shown on the left. On the right the Chromar rules used to grow a structure like the one on the left panel (see main text for details).

(see Discussion). Dormant seed never die in the model and may germinate after several years (Burghardt et al., 2015).

Since we are using an individual-based model, n_i becomes computationally prohibitive to simulate over decades of population growth. In order to overcome this limitation, we simulated the timing (phenology) and growth components sequentially and used conservative birth functions $b(m)$. Figure 6 introduces the plotting conventions for these results. The timing components were first simulated with $b(m)=1$, such that each plant makes one seed, as in Burghardt et al. (2015) but with a daily time step (see discussion in model description) as opposed to an hourly one. The phenological simulation results in an unbranched sequence of developmental stage timings for each lineage (Fig. 6C). The simulation results for several decades typically revealed a small number of life cycle growth strategies, from clusters of individual life cycles. The clusters were generated using k -means clustering, where k is chosen by visual inspection of the life cycle plots (Fig. 6A). Alternative clustering approaches might be an area for future work. Figure 6A shows the distribution over a year of all individual life cycles that conformed to two contrasting life cycle strategies under environmental conditions for Valencia (see Results). Cluster membership depends on the dates and durations of multiple developmental stages. This is hard to visualize, because the timing of any single developmental stage partially overlaps among different strategies. Figure 6B therefore summarizes the median dates of all three developmental transitions in each strategy, here illustrated by (i) a summer growth strategy and (ii) a winter growth strategy. In the next stage, the growth models were simulated once per cluster, with the environmental conditions associated with the typical timing of that cluster (median vegetative and reproductive stages). This returns the typical biomass of organs over time, including the fruit mass at seed dispersal (m_1 for cluster 1, m_2 for cluster 2; Fig. 6D). Finally, each life cycle is assigned the fruit mass m associated with its cluster, and thereby a growth-based, reproductive success $b(m)$ that evaluates to 0 in some cases. Thus, the second stage recovers a version of the branching lineage tree, where some lineages die out (Fig. 6E).

Our output population measure is the total population of plants over all lineages over all generations. For example, consider a lineage with three generations starting with a plant with final fruit mass m_{11} . For the next generation we have $b(m_{11})$ and then $b(m_{11}) \times b(m_{21})$. The population measure for that lineage is $1 + b(m_{11}) + b(m_{11}) \times b(m_{21})$. The population measure for multiple lineages starting from multiple plants in the initial population is the sum of the population measures of all the lineages. This requires a birth function, which we use in a very simple form, as follows:

$$b(m) = \begin{cases} 1 & \text{if } m > m_0 \\ 0 & \text{otherwise} \end{cases}$$

A plant produces one seed or none, the latter in life cycles with fruit mass at seed dispersal m less than a threshold m_0 . Below, we make some conservative choices for the value of the reproductive threshold value, m_0 , to explore the effect on the output population measure.

Finally, we distinguish three sources of variability in the population model: (i) weather varies between years, (ii) genetic parameters can vary among the initial population if their values are chosen probabilistically, and (iii) simulation results vary due to stochasticity in the model representation.

Weather data

For the phenology model simulation we used the weather data that accompanied the Burghardt et al. (2015) model, available from a Dryad repository (Burghardt et al. 2014). In this dataset weather inputs over 60 years were generated stochastically for four locations in Europe: Halle, Valencia, Norwich, and Oulu. The weather inputs include values for temperature (in °C), moisture (water potential in MPa), and day length (in hours).

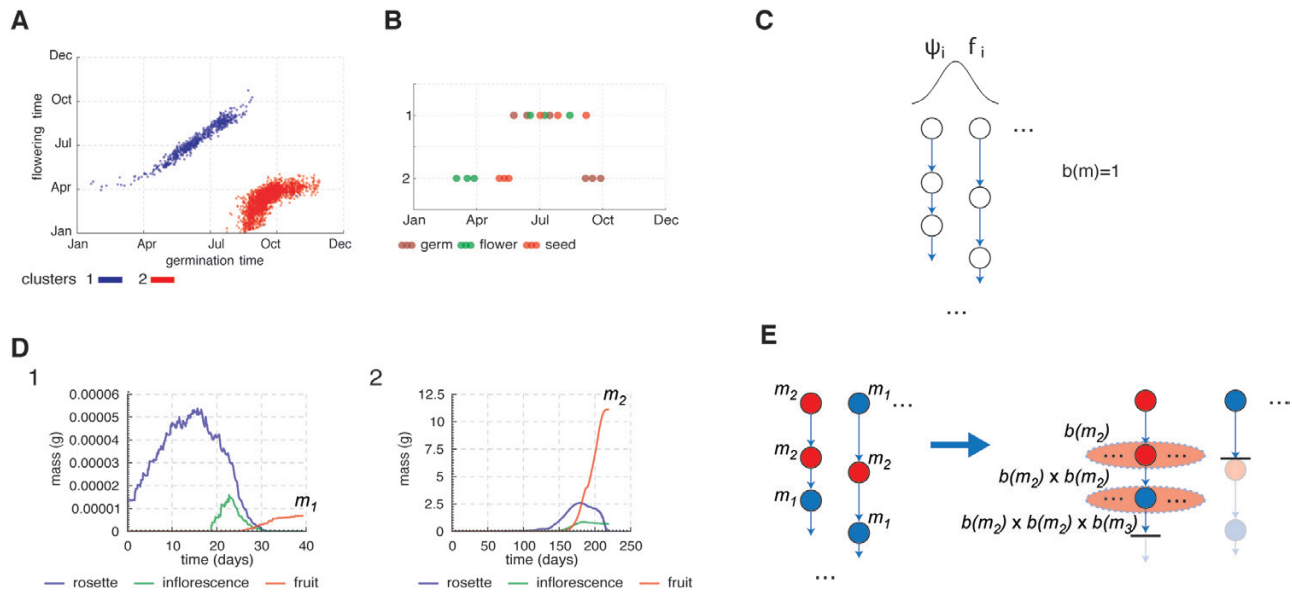


Fig. 6. The two-stage simulation of the population level model. (A) The distributions of developmental events (germination, flowering) from the phenology-only simulation of the population model (C) with the identification of two clusters representing two distinct strategies (blue, red). (B) The 25th, 50th, and 75th percentiles of the distributions of developmental events of the two clusters from (A) (cluster labels on y-axis). Germination time distributions are in brown, flowering time distributions are in green and seed dispersal distributions are in red. (C) Illustration of the phenology only simulation with growth independent birth function $b(m)=1$ (each plant makes a single seed). The choice of, ψ_i (seed dormancy level) and f_i (initial floral repression) is done probabilistically through a normal distribution. (D) Results of simulations of the growth models for the median dates of developmental events (B) for the clusters identified in (A). m_1, m_2 are the final fruit masses for a plant with median developmental event timings for clusters 1 and 2, respectively (as identified and shown in A, B). (E) Illustration of the assignment of fruit masses to the lifecycles of the phenology-only simulation (C) according to their clusters. m_1, m_2 refer to final fruit masses for a median plant in clusters 1 and 2 (from growth simulations in (D) and same two clusters in (A, B)). This recovers the full branching population process where the horizontal bars indicate a lineage dying out ($b(m)=0$ at that generation, no reproduction) and the potential rest of the lineages is shaded. The breadth of the ellipses indicates the reproductive success when there is reproduction ($b(m)>0$).

For the growth simulations we used weather data from the ECMWF ERA-Interim dataset over the years 2010–2011 (Dee *et al.*, 2011). A program provided by Mathew Williams and Luke Smallman (School of GeoSciences, University of Edinburgh) that uses methods from Williams *et al.* (2001) was used to generate hourly inputs given daily averages from the dataset for temperature and radiation. For the soil moisture input used in the photosynthesis rate calculation, we used a daily average of soil moisture values from the dataset and assumed that to be constant throughout the day (swvl parameters in the ERA dataset). The soil moisture parameter here is a number in arbitrary units from 0 to 1 that represents the ‘wetness’ of the soil while the soil moisture used above measures water potential and is given in MPa.

Results

The population of FM-life models (see Methods) allows us to test how growth processes that alter reproductive success affect the life history strategies of Arabidopsis growing in different environmental conditions (location) and with different genetic parameters in the initial population. We can therefore explore the genotype \times environment interaction, using a population measure. To illustrate this potential, we compare simulation results for two previously studied locations, Valencia (Spain) and Oulu (Finland), and two opposing combinations of genetic parameters, high seed dormancy/high floral repression (HH) and low seed dormancy/low floral repression (LL). Within an initial population of 100 seeds, the seed dormancy levels, ψ_i , were assigned probabilistically, sampling from a normal distribution with mean 0.0 and standard deviation 1, or the low dormancy level (L) and mean 2.5 with the same standard deviation for the high dormancy case (H). Floral repression was fixed at either 0.598 for the low level (L) and 0.737 for the high level (H), values that were chosen to reflect the behaviour of natural populations of Arabidopsis in Wilczek *et al.* (2009). Both parameter choices follow Burghardt *et al.* (2015). The simulation time period was 60 years and, as in Burghardt *et al.* (2015), we discarded the first 15 years of the simulation to focus on stable life history strategies. A key difference from the earlier work is that even our conservative choice of birth function (see Methods) allows some lineages to die out.

Valencia

Figure 7 shows the results of the two-stage simulation for a population of the LL genotype in Valencia (Fig. 7A). We identify four possible life history strategies based on the timing (phenology) components of the FM-life model:

- (i) Strategy 1, *summer-only strategy* where the entire growth is in the summer. The growth period is quite short and the conditions unfavourably hot and dry. In the growth simulation, the rosette leaves senesce before the reproductive stage (purple curve). The drought effect on photosynthesis severely limits the carbon available for fruit mass (orange curve).
- (ii) Strategy 2, *spring strategy* where the entire growth period is in the spring. The growth period is only slightly longer than the *summer-only* strategy but it falls in more favourable weather conditions. The rosette lifetime extends beyond

flowering to support fruit growth, which combined with favourable weather gives high fruit mass.

- (iii) Strategy 3, *winter-repr strategy* spans the winter/early spring period. A short vegetative period in the end of summer/early autumn ends with flowering and a long reproductive stage over the winter/early spring. The rosette is senescing when favourable conditions return in early spring, seriously limiting fruit development.
- (iv) Strategy 4, *winter-veg strategy* again spans the winter/early spring period. The life cycle duration is similar to strategy (iii) but slightly later germination delays flowering until spring. The rosette grows all winter, overlapping with a short reproductive stage and supporting high fruit mass.

Plants with life cycle strategies 2 and 4 predicted orders of magnitude more fruit mass than plants with life cycle strategy 3 and or the least successful strategy 1 (Fig. 7C). This result clearly ranked the strategies available to plants of the LL genotype, although the absolute values of the predicted biomass are less certain (see Discussion). The 100 plants amassed 4905 potential lifecycles over 45 years of phenological simulation (Fig. 7A). Without a minimum mass threshold (m_0) for reproduction, 66% of potential life cycles followed the more successful *spring* and *winter-veg* strategies (2 and 4; Fig. 7C). Figure 7D shows the sequential transitions between strategies. For example, 60% of potential plants following the successful *spring* strategy (2) disperse their seeds early enough for the next generation to adopt the *winter-veg* strategy (4), achieving two generations per year. These transitions underlie the bimodal distribution of life cycle times reported by Burghardt *et al.* (2015) for this simulation.

Simulation of the HH genotype (Fig. 7E,J) identified similar strategies. Since the seed have longer dormancy, the population amassed fewer potential life cycles (2954 as opposed to 4905 in the LL case; Fig. 7F). The growth and final fruit masses are different because of slight variation in timing of the growth period but strategies 2 and 4 are again more successful than strategies 1 and 3 (Fig. 7H). A higher fraction of potential life cycles followed the successful strategies (78% as opposed to 66% in the LL case; Fig. 7G). Higher seed dormancy reduced the germination in the summer and early autumn that led to the less successful strategies 1 and 3, so any strategy was likely to be followed by either strategy 2 or 4 in the next generation (Fig. 7).

In order to calculate the population success, we make two choices for the reproduction mass threshold, m_0 , which eliminate one or both of the least successful strategies. Choosing a value $m_0=2\times 10^{-5}$ g (the mass of a single seed) eliminated the *summer-only* strategy from both genotypes, which gives a population of 1210 plants over 45 years in the LL case (Fig. 7C). The HH genotype allows a larger percentage of viable life cycles but we predict fewer plants (1020) since the number of potential life cycles was lower (Fig. 7H). Choosing a value $m_0=6\times 10^{-3}$ g left only two viable strategies, 2 and 4, for both genotypes. Reciprocal transitions between the strategies were still possible but *winter-veg* was strongly favoured (Fig. 7E, J). The LL genotype predicted 240 plants in total

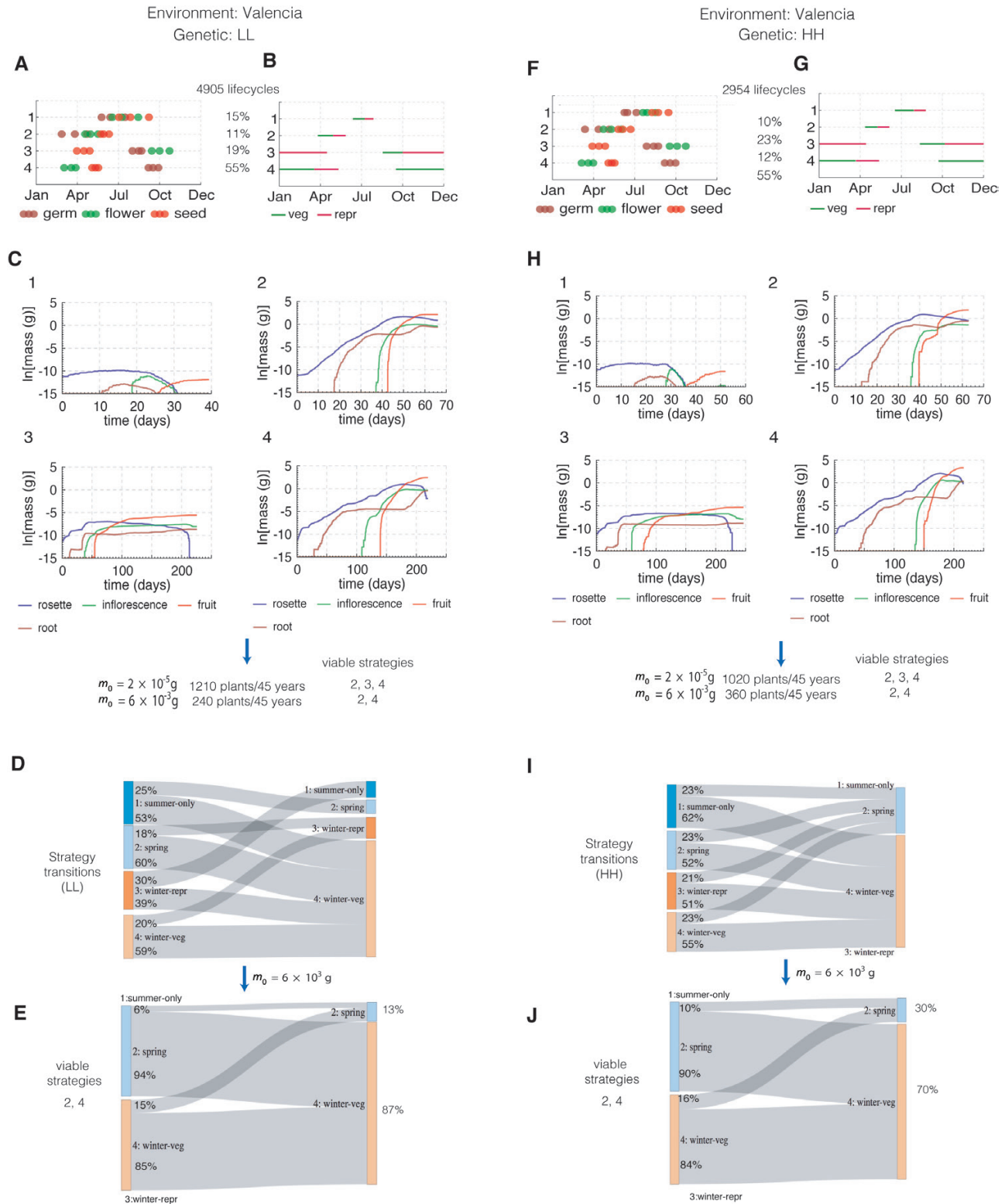


Fig. 7. Population experiments in Valencia in two different genetic backgrounds (LL, low dormancy, low floral repression; and HH, high dormancy, high floral repression). (A) Results of the phenology-only simulation results for the LL genotype. The 25th, 50th, and 75th percentiles of the distribution of developmental events are shown for each identified cluster. (B) Growth stages over a year for each cluster from (A) according to the median time of the distribution of developmental events for each cluster. (C) Growth simulations over the growth period shown in (B) for each cluster. (D) Probabilities of successive strategies. (E) Probabilities of successive strategies after eliminating life cycles with strategies 1 and 3 using a reproductive threshold. (F–J) As for (A–E) but for the HH genotype in the same location.

over 45 years, compared with 360 plants for the HH genotype: G×E interaction favoured the HH genotype despite its smaller number of potential life cycles. Thus, modelling the growth processes not only distinguished among the potential life cycle strategies within a genotype but also between the genotypes.

Oulu

The equivalent simulations were performed for conditions in Oulu, Finland in the same LL and HH genetic backgrounds (Fig. 8). The results indicated three potential life cycle strategies (Fig. 8A, B, E, F):

- (i) Strategy 1, *summer-only strategy* where the entire life cycle occurs in the summer. The vegetative period is short, the rosette is very small and supports negligible fruit growth (Fig. 7C).
- (ii) Strategy 2, *winter-repr strategy* where a life cycle of almost a year has a very short vegetative stage, followed by a long reproductive stage over the winter. Again, the very small rosette supports little fruit growth in the following spring.
- (iii) Strategy 3, *winter-veg strategy* where the plant overwinters in the vegetative stage. Unlike in Valencia, the rosette grows little over the winter. Rapid rosette growth in the following spring supports a substantial inflorescence and fruit development, though the predicted fruit mass is smaller than in Valencia.

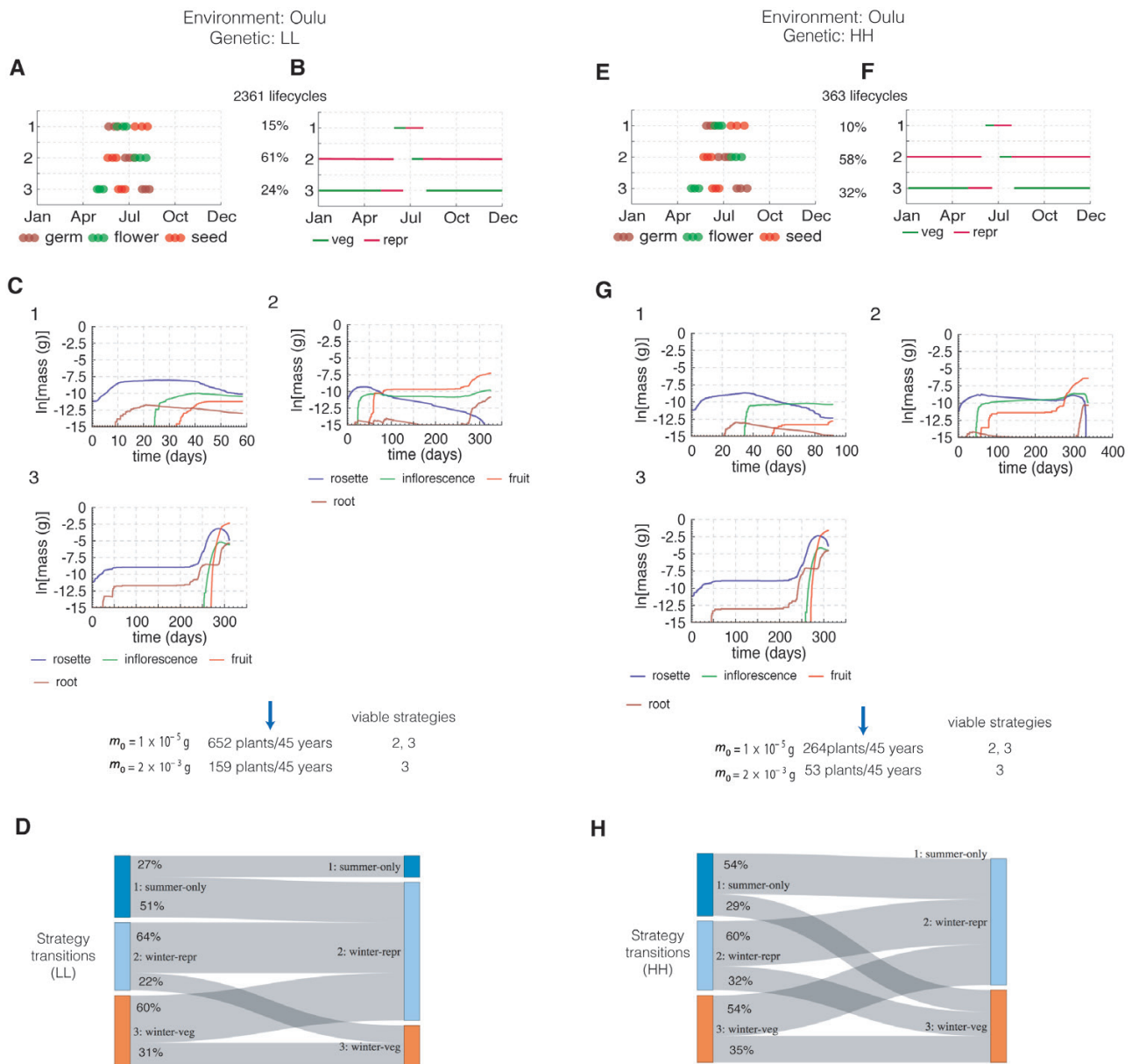


Fig. 8. Population experiments in Oulu in two different genetic backgrounds (LL, low dormancy, low floral repression; and HH, high dormancy, high floral repression). (A) Results of the phenology-only simulation results for the LL genotype. The 25th, 50th, and 75th percentiles of the distribution of developmental events are shown for each identified cluster. (B) Illustration of growth stages over a year for each cluster according to the median time of the distribution of developmental events for each cluster. (C) Growth simulations over the growth period shown in (B) for each cluster. (D) Probabilities of successive strategies. (E–J) As for (A–E) but for HH genotype in the same location.

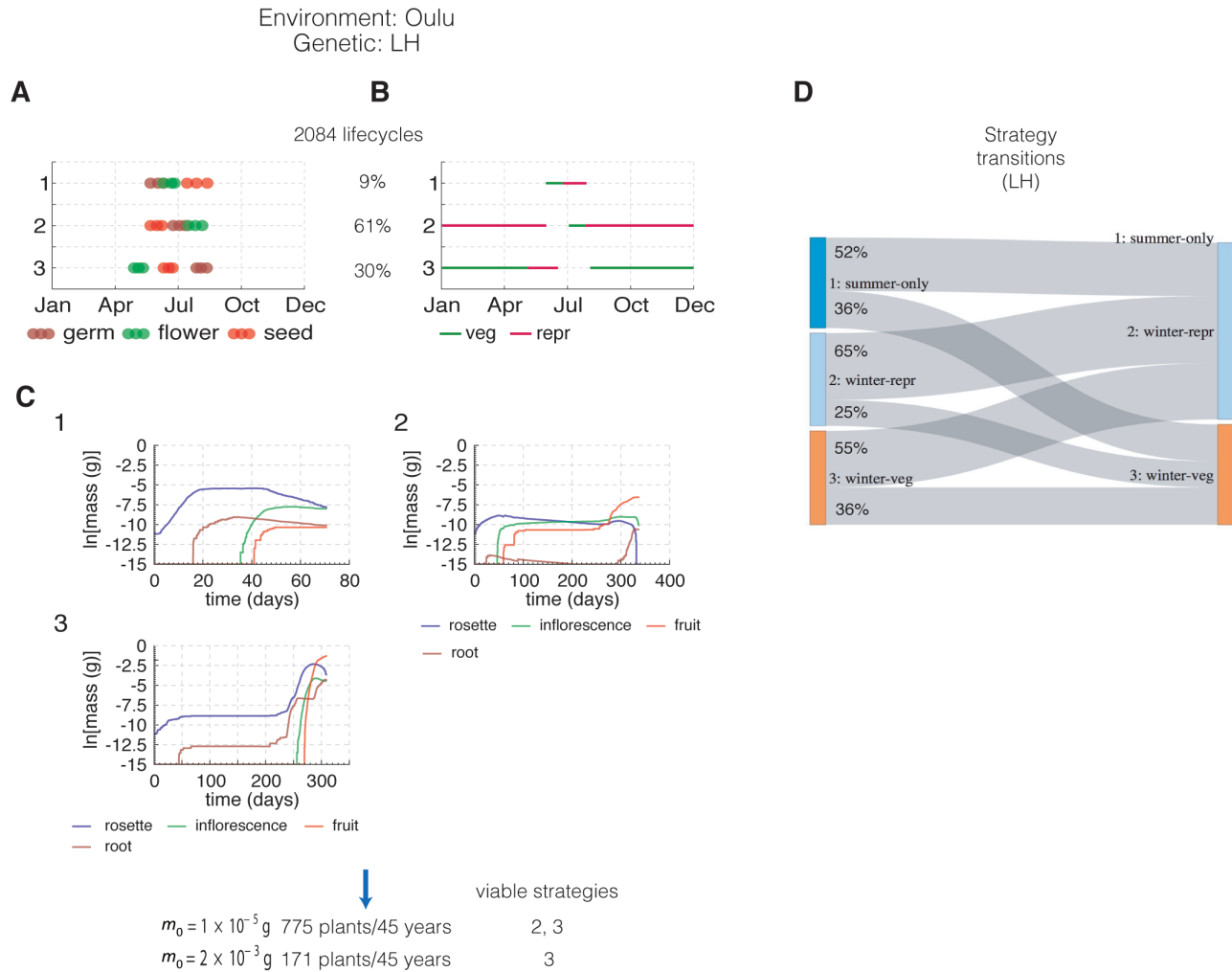


Fig. 9. Simulation results for a combined variant in Oulu (LH, low dormancy, high floral repression). (A) Results of the phenology-only simulation results. The 25th, 50th, and 75th percentile of the distribution of developmental events are shown for each identified cluster. (B) Illustration of growth stages over a year for each cluster according to the median time of the distribution of developmental events for each cluster. (C) Growth simulations over the growth period shown in (B) for each cluster. (D) Probabilities of successive strategies.

The severe winter conditions limited the number of potential life cycles to 2361 for the LL genotype or 363 for HH. A higher proportion of HH life cycles followed the successful *winter-veg* strategy (3; 32% against 24% in LL; Fig. 8B, G). Surprisingly, a majority of life cycles for both genotypes followed the *winter-repr* strategy 2. Applying the reproductive threshold mass, m_0 , eliminated one or both of strategies 1 and 2 (Fig. 8C, G), suggesting a strong selective pressure for greater floral repression to reduce the number of *winter-repr* life cycles. With $m_0=2 \times 10^{-3}$ g, the LL genotype yielded 159 plants over 45 years compared with 53 plants for HH. All G×E combinations had actively growing plants at the end of the simulation. Interestingly, plants of the HH genotype had higher average reproductive success per plant in Oulu yet the LL plants were more successful by our population measure. The faster development of LL plants allowed more, short lifecycles within the simulated interval (consistent with the phenology model alone).

Discussion

We present a whole-life-cycle multi-model for growth and reproduction of *Arabidopsis*, FM-life, combining phenology

models that time the developmental stages and growth models to predict organ biomass. The simple, FM-lite model of vegetative growth and its extension to the reproductive stage in FM-life simulate broader, mechanistically founded components of fitness at the individual plant level compared with the phenology models alone. Most insights from the component models naturally remain (Rasse and Tocquin, 2006; Christophe et al., 2008; Wilczek et al., 2009; Burghardt et al., 2015). Multi-models are helpful in emphasizing interactions. The cauline leaves in our inflorescence model, for example, extend the duration of photosynthetic competence. As cauline leaves can be produced 6 months later than early rosette leaves in the *winter-veg* strategy (Fig. 8), they remain active photosynthetic sources (Earley et al., 2009; Leonardos et al., 2014) when the rosette leaves are senescing. The growth models provided the fruit mass that we used as an indicator of reproductive success, such that metabolic and developmental processes of growth informed a more mechanistic understanding of ecological, population dynamics over multiple generations.

The growth model allowed us to discriminate among alternative life cycle strategies in each G×E combination, by selecting

against strategies that were compatible with the phenology models alone but had qualitatively worse growth. In previous work, strategies with high seed dormancy in southern Valencia and low dormancy in northern Oulu were noted to align with the behaviour of the cognate wild populations (Atwell *et al.*, 2010; Chiang *et al.*, 2011; Méndez-Vigo *et al.*, 2011; Burghardt *et al.*, 2015). In each G×E combination, individual plants in our simulations might adopt alternative life cycle strategies. The less-successful strategies were lethal in our model, eliminating >95% of potential lifecycles (simulated by the phenology model alone) for the LL genotype in Valencia, for example (Fig. 7A, C). Thus, our results supported the observed genotypic distinction between Valencia and Oulu, because the requirement for a minimum fruit mass eliminated more lineages of the less-successful genotype in each case (Figs 7C, H, 8C, G).

Our approach might appear conservative, as the binary birth function (one seed/no seed) ignored variation in seed mass among life cycle strategies, which might otherwise reinforce the advantage of successful strategies. The successful genotype LL in Oulu, however, had lower reproductive success per plant than HH, suggesting a more subtle balance of advantage. Genotypes with low dormancy and high floral repression (LH) are observed in far northern locations (Atwell *et al.*, 2010). We therefore simulated the LH genotype (Fig. 9). LH plants delayed flowering time enough to reduce the frequency of potential *summer-only* life cycles to 9% compared with 15% in LL (Fig. 9B) and increased the fruit mass of the successful *winter-veg* life cycle close to the HH genotype (Fig. 9C). The LH model predicted slightly higher

reproductive success overall, returning 171 life cycles (Fig. 9C) compared with 159 for the LL variant (Fig. 8C), consistent with the observation of LH genotypes at this location.

A limitation of our work arises from the fact that the phenology component models of FM-life have been validated against field data (Wilczek *et al.*, 2009; Burghardt *et al.*, 2015) whereas the growth component models have not (Rasse and Tocquin, 2006; Christophe *et al.*, 2008). This will clearly affect our results, illustrated by the fact that selecting a single year's weather data for Valencia and fixing genetic parameter values eliminated the summer strategy (Fig. 10A, B), despite the remaining variability due to the model's stochastic representation. Biomass simulations are inevitably sensitive to the timing of the growth period, because a longer interval of exponential growth in good conditions rapidly changes absolute biomass, as illustrated in Fig. 10C, D. We therefore modified the growth models to account for severe winter conditions and to limit photosynthesis in dry conditions. These changes are conservative as severe weather conditions also affect growth in other ways. We already mentioned the effect of drought on partitioning, for example, but other metabolic processes like maintenance respiration are also known to be affected. These effects have been studied in crop species (McDowell, 2011), but little is known on the responses of Arabidopsis. Despite the changes, the FM-life model predicted unreasonably high fruit mass in some cases. The binary birth function ensured that this had no effect on our population measure.

On the other hand, in the least successful strategies it is possible that the model underestimates the predicted fruit mass, for

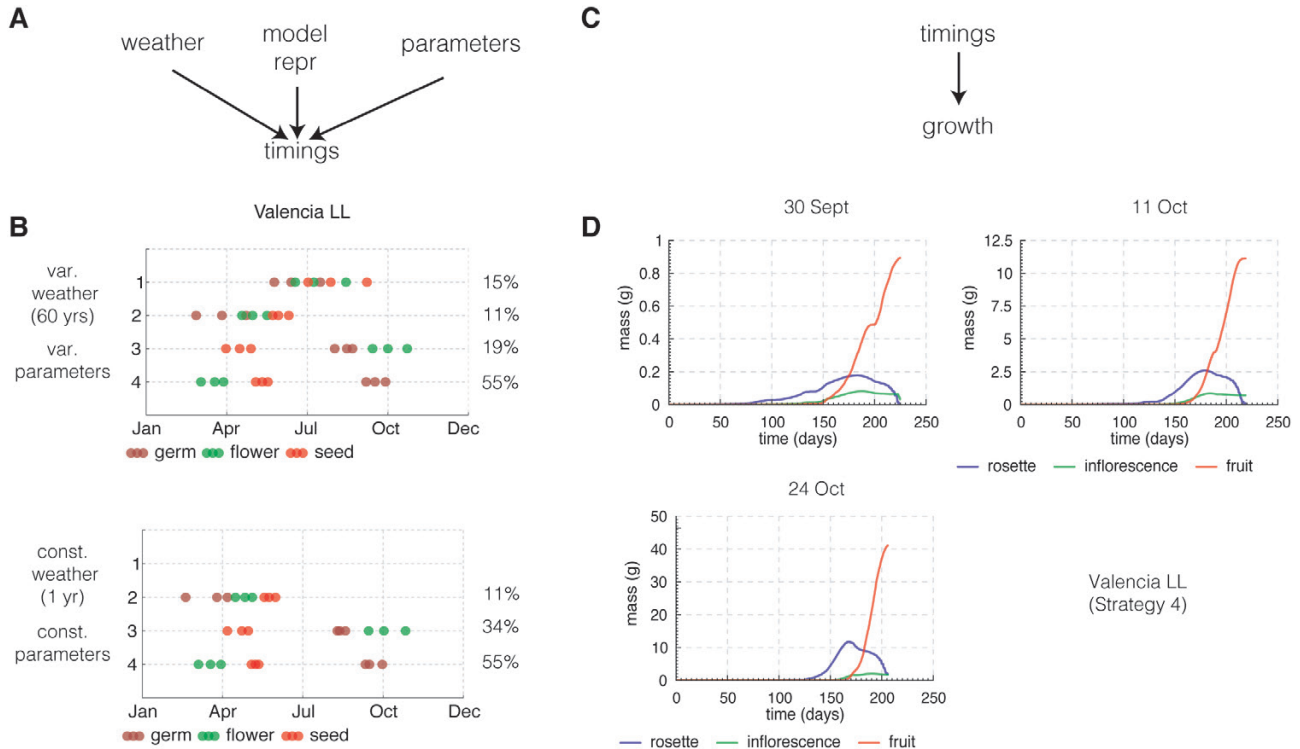


Fig. 10. Robustness of timing and growth simulation results. (A) Sources of variability for the timing results. (B) Timing results for variable weather (60 years) and variable genetic parameters chosen from a distribution and timing results for simulations with constant weather (same 1-year weather over 60 years) and constant parameters. The timings of the strategies are very similar but strategy 1 life cycles have moved to strategy 3. This could be because of the particular year of weather data we used for the simulations. (C) Sources of variability for growth results. (D) Growth simulation results starting from three different germination dates (corresponding to 25th, 50th, and 75th percentiles of the distribution of germination times) of strategy 4 in Valencia LL (Fig. 8).

Table 1. Table of relevant parameter values. Developmental thresholds for the timing components of the three stages are from the Burghardt et al. (2015) model

Name	Units	Value	Description
D_g	°C MPa	1000	Developmental threshold for germination
D_f	°C h	2604	Developmental threshold for flowering
D_s	°C	8448	Developmental threshold for seed dispersal
m_0	g	various	Reproductive threshold
ψ_i	MPa	Low(L)=0.0 High(H)=2.5	Initial seed dormancy (distribution means)
f_i	None	Low(L)=0.598 High(H)=0.737	Initial floral repression

example because we ignore the contribution of stem photosynthesis in the assimilation rule. We chose the threshold fruit mass m_0 conservatively, so the lowest threshold is the mass of a single seed. A wide range of reasonable thresholds would eliminate the same one or two strategies in each G×E scenario, as the strategies give such large predicted differences in biomass. As the ranking of strategies is so clear, minor alterations of the biomass simulations will not affect our qualitative comparison across genotypes and locations.

Among possible gaps in understanding of the environmental effects on growth in natural settings or in our representation, we repeat our previous caution (Chew et al., 2014b, 2017, Preprint) that models of nutrient balance for Arabidopsis will be helpful. Rosette biomass in the Framework Model is understandably sensitive to photosynthetic parameters (Chew et al., 2014b), yet these have not been validated in Arabidopsis across the wide range of photoperiods and temperatures simulated here (Walker et al., 2013). FM-life predicts a discretized fruit mass per cluster and hence reproductive success for a typical representative of each life cycle strategy, approximating an underlying, continuous distribution of fruit mass among plants. The accuracy of this approximation will depend on the variation of plants within clusters. The benefit lies in computational tractability, allowing us to simulate differential reproductive success that is informed by understanding of growth processes.

Much like FMv1, our version here (FM-life) also takes the same breadth-first strategy in modelling where the main goal is to show that links across scales are possible. The modular nature of the models means that domain experts can add detailed alternatives for particular processes in future. The aim of the Framework Model is to provide the context for such studies, in this case reaching to ecological outcomes, or in the FMv2, testing more detailed effects of circadian regulation upon metabolism and hence biomass (Chew et al., 2017, Preprint).

Our approach builds upon previous models that predict fitness and population processes in Arabidopsis, which have focused on developmental components of fitness or on phenology (Prusinkiewicz et al., 2007; Satake et al., 2013; Springthorpe and Penfield, 2015). Linking these components sharpens ecological insight, by understanding the performance of genetic variants in the environment that underlies differences in fitness (see discussions in Donohue et al., 2015; Doebeli et al., 2017) and can thus inform evolutionary hypotheses. Adding genetic variation between generations will in future model Arabidopsis evolution explicitly, perhaps after competing genetic variants

in silico using adaptive dynamics approaches (Brännström et al., 2013; Weiße et al., 2015). Our approach giving a mechanistically founded reproductive success can also inform crop models where the main focus is yield. With a more detailed nutrient balance model yield (e.g. total seed mass) can be assessed in different G×E scenarios or under different management interventions. Thus, the FM-life model offers a further tool to bridge among disciplines in plant biology, ecology, and evolution.

Acknowledgements

We are grateful to L. Burghardt and K. Donohue (Duke University) for sharing data and model definitions during a post-publication embargo period, and to L. Smallman and M. Williams (University of Edinburgh) for expert assistance with meteorological inputs and ecophysiological simulation and to L. McNally (University of Edinburgh) for discussion of evolutionary aspects. AZ was supported by an iCASE studentship from the Biotechnology and Biological Sciences Research Council (award BB/K011294/1) with Simulistics Ltd, and AH by a FLIP Fellowship (BB/M017605/1).

References

- Alvarado V, Bradford K. 2002. A hydrothermal time model explains the cardinal temperatures for seed germination. *Plant, Cell & Environment* **25**, 1061–1069.
- Asseng S, Ewert F, Rosenzweig C, et al. 2013. Uncertainty in simulating wheat yields under climate change. *Nature Climate Change* **3**, 827.
- Atwell S, Huang YS, Vilhjálmsson BJ, et al. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627–631.
- Brännström Å, Johansson J, von Festenberg N. 2013. The Hitchhiker's Guide to Adaptive Dynamics. *Games* **4**, 304–328.
- Buckley TN, Mott KA. 2013. Modelling stomatal conductance in response to environmental factors. *Plant, Cell & Environment* **36**, 1691–1699.
- Burghardt L, Metcalf C, Wilczek A, Schmitt J, Donohue K. 2014. Data from: Modeling the influence of genetic and environmental variation on the expression of plant life cycles across landscapes. Dryad Digital Repository. <https://doi.org/10.5061/dryad.nv0p1>.
- Burghardt LT, Metcalf CJ, Wilczek AM, Schmitt J, Donohue K. 2015. Modeling the influence of genetic and environmental variation on the expression of plant life cycles across landscapes. *The American Naturalist* **185**, 212–227.
- Chenu K, Van Oosterom EJ, McLean G, et al. 2018. Integrating modelling and phenotyping approaches to identify and screen complex traits: transpiration efficiency in cereals. *Journal of Experimental Botany* **69**, 3181–3194.
- Chew YH, Seaton DD, Mengin V, Flis A, Mugford ST, Smith AM, Stitt M, Millar AJ. 2017. Linking circadian time to growth rate quantitatively via carbon metabolism. *BioRxiv* 105437. [Preprint].

- Chew YH, Smith RW, Jones HJ, Seaton DD, Grima R, Halliday KJ.** 2014a. Mathematical models light up plant signaling. *The Plant Cell* **26**, 5–20.
- Chew YH, Wenden B, Flis A, et al.** 2014b. Multiscale digital Arabidopsis predicts individual organ and whole-organism growth. *Proceedings of the National Academy of Sciences, USA* **111**, E4127–E4136.
- Chew YH, Wilczek AM, Williams M, Welch SM, Schmitt J, Halliday KJ.** 2012. An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytologist* **194**, 654–665.
- Chiang GC, Bartsch M, Barua D, Nakabayashi K, Debieu M, Kronholm I, Koornneef M, Soppe WJ, Donohue K, De Meaux J.** 2011. *DOG1* expression is predicted by the seed-maturation environment and contributes to geographical variation in germination in *Arabidopsis thaliana*. *Molecular Ecology* **20**, 3336–3349.
- Christophe A, Letort V, Hummel I, Cournède P-H, de Reffye P, Lecœur J.** 2008. A model-based analysis of the dynamics of carbon balance at the whole-plant level in *Arabidopsis thaliana*. *Functional Plant Biology* **35**, 1147–1162.
- Chuine I.** 2010. Why does phenology drive species distribution? *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences* **365**, 3149–3160.
- Chuine I, Beaubien EG.** 2001. Phenology is a major determinant of tree species range. *Ecology Letters* **4**, 500–510.
- Dee DP, Uppala S, Simmons A, et al.** 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137**, 553–597.
- Doebeli M, Ispolatov Y, Simon B.** 2017. Towards a mechanistic foundation of evolutionary theory. *eLife* **6**, e23804.
- Donohue K, Burghardt LT, Runcie D, Bradford KJ, Schmitt J.** 2015. Applying developmental threshold models to evolutionary ecology. *Trends in Ecology & Evolution* **30**, 66–77.
- Earley EJ, Inghand B, Winkler J, Tonsor SJ.** 2009. Inflorescences contribute more than rosettes to lifetime carbon gain in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany* **96**, 786–792.
- Farquhar GD, von Caemmerer S, Berry JA.** 1980. A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta* **149**, 78–90.
- France J, Thornley JHM.** 1984. *Mathematical models in agriculture*. London: Butterworths.
- Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, van Eeuwijk F, Chapman S, Podlich D.** 2006. Models for navigating biological complexity in breeding improved crop plants. *Trends in Plant Science* **11**, 587–593.
- Hepworth J, Antoniou-Kourounioti RL, Bloomer RH, et al.** 2018. Absence of warmth permits epigenetic memory of winter in *Arabidopsis*. *Nature Communications* **9**, 639.
- Honorato-Zimmer R, Millar AJ, Plotkin GD, Zardilis A.** 2017. Chromar, a language of parameterised agents. *Theoretical Computer Science*, doi: 10.1016/j.tcs.2017.07.034.
- Honorato-Zimmer R, Millar AJ, Plotkin GD, Zardilis A.** 2018. Chromar, a rule-based language of parameterised objects. *Electronic Notes in Theoretical Computer Science* **335**, 49–66.
- Kage H, Kochler M, Stützel H.** 2004. Root growth and dry matter partitioning of cauliflower under drought stress conditions: measurement and simulation. *European Journal of Agronomy* **20**, 379–394.
- Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B Jr, Assad-Garcia N, Glass JI, Covert MW.** 2012. A whole-cell computational model predicts phenotype from genotype. *Cell* **150**, 389–401.
- Kinmonth-Schultz HA, MacEwen MJ, Seaton DD, Millar AJ, Imaizumi T, Kim S-H.** 2018. Mechanistic model of temperature influence on flowering through whole-plant accumulation of FT. *BioRxiv* 267104. [Preprint].
- Klose C, Venezia F, Hussong A, Kircher S, Schäfer E, Fleck C.** 2015. Systematic analysis of how phytochrome B dimerization determines its specificity. *Nature Plants* **1**, 15090.
- Le Novère N.** 2015. Quantitative and logic modelling of molecular and gene networks. *Nature Reviews. Genetics* **16**, 146–158.
- Leonardos ED, Rauf SA, Weraduwege SM, Marillia E-F, Taylor DC, Micallef BJ, Grodzinski B.** 2014. Photosynthetic capacity of the inflorescence is a major contributor to daily-C-gain and the responsiveness of growth to elevated CO₂ in *Arabidopsis thaliana* with repressed expression of mitochondrial-pyruvate-dehydrogenase-kinase. *Environmental and Experimental Botany* **107**, 84–97.
- Lynch JP.** 2013. Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. *Annals of Botany* **112**, 347–357.
- McDowell NG.** 2011. Mechanisms linking drought, hydraulics, carbon metabolism, and vegetation mortality. *Plant Physiology* **155**, 1051–1059.
- Méndez-Vigo B, Picó FX, Ramiro M, Martínez-Zapater JM, Alonso-Blanco C.** 2011. Altitudinal and climatic adaptation is mediated by flowering traits and *FRI*, *FLC*, and *PHYC* genes in *Arabidopsis*. *Plant Physiology* **157**, 1942–1955.
- Millar AJ.** 2016. The intracellular dynamics of circadian clocks reach for the light of ecology and evolution. *Annual Review of Plant Biology* **67**, 595–618.
- Millar AJ, Urquiza-García U, Freeman PL, Hume A, Plotkin GD, Sorokina O, Zardilis A, Zielinski T.** 2019. Practical steps to digital organism models, from laboratory model species to ‘Crops in silico’. *Journal of Experimental Botany* doi: 10.1093/jxb/ery435.
- Mündermann L, Erasmus Y, Lane B, Coen E, Prusinkiewicz P.** 2005. Quantitative modeling of Arabidopsis development. *Plant Physiology* **139**, 960–968.
- Parent B, Tardieu F.** 2014. Can current crop models be used in the phenotyping era for predicting the genetic variability of yield of plants subjected to drought or high temperature? *Journal of Experimental Botany* **65**, 6179–6189.
- Pouteau S, Albertini C.** 2009. The significance of bolting and floral transitions as indicators of reproductive phase change in *Arabidopsis*. *Journal of Experimental Botany* **60**, 3367–3377.
- Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E.** 2007. Evolution and development of inflorescence architectures. *Science* **316**, 1452–1456.
- Rasse DP, Tocquin P.** 2006. Leaf carbohydrate controls over Arabidopsis growth and response to elevated CO₂: an experimentally based model. *New Phytologist* **172**, 500–513.
- Salazar JD, Saithong T, Brown PE, Foreman J, Locke JC, Halliday KJ, Carré IA, Rand DA, Millar AJ.** 2009. Prediction of photoperiodic regulators from quantitative gene circuit models. *Cell* **139**, 1170–1179.
- Satake A, Kawagoe T, Saburi Y, Chiba Y, Sakurai G, Kudoh H.** 2013. Forecasting flowering phenology under climate warming by modelling the regulatory dynamics of flowering-time genes. *Nature Communications* **4**, 2303.
- Springthorpe V, Penfield S.** 2015. Flowering time and seed dormancy control use external coincidence to generate life history strategy. *eLife* **4**, 05557.
- Voss U, Bishopp A, Farcot E, Bennett MJ.** 2014. Modelling hormonal response and development. *Trends in Plant Science* **19**, 311–319.
- Walker B, Ariza LS, Kaines S, Badger MR, Cousins AB.** 2013. Temperature response of in vivo Rubisco kinetics and mesophyll conductance in *Arabidopsis thaliana*: comparisons to *Nicotiana tabacum*. *Plant, Cell & Environment* **36**, 2108–2119.
- Weiß AY, Oyarzún DA, Danos V, Swain PS.** 2015. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences, USA* **112**, E1038–E1047.
- Welch S, Roe J, Das S, Dong Z, He R, Kirkham M.** 2005. Merging genomic control networks and soil-plant-atmosphere-continuum models. *Agricultural Systems* **86**, 243–274.
- Wilczek AM, Roe JL, Knapp MC, et al.** 2009. Effects of genetic perturbation on seasonal life history plasticity. *Science* **323**, 930–934.
- Williams M, Rastetter EB, Shaver GR, Hobbie JE, Carpino E, Kwiatkowski BL.** 2001. Primary production of an arctic watershed: an uncertainty analysis. *Ecological Applications* **11**, 1800–1816.
- Wu A, Song Y, van Oosterom EJ, Hammer GL.** 2016. Connecting biochemical photosynthesis models with crop models to support crop improvement. *Frontiers in Plant Science* **7**, 1518.
- Yin X, Struik PC.** 2008. Applying modelling experiences from the past to shape crop systems biology: the need to converge crop physiology and functional genomics. *New Phytologist* **179**, 629–642.
- Yin X, Struik PC.** 2010. Modelling the crop: from system dynamics to systems biology. *Journal of Experimental Botany* **61**, 2171–2183.
- Zhu XG, Lynch JP, Lebauer DS, Millar AJ, Stitt M, Long SP.** 2016. *Plants in silico*: Why, why now and what?—An integrative platform for plant systems biology research. *Plant, Cell & Environment* **39**, 1049–1057.
- Zhu XG, Wang Y, Ort DR, Long SP.** 2013. e-Photosynthesis: a comprehensive dynamic mechanistic model of C₃ photosynthesis: from light capture to sucrose synthesis. *Plant, Cell & Environment* **36**, 1711–1727.