THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Multi-scale modelling to synergise plant systems biology and crop science

Citation for published version:

Chew, YH, Seaton, D & Millar, A 2016, 'Multi-scale modelling to synergise plant systems biology and crop science', *Field crops research*, vol. 202. <https://doi.org/10.1016/j.fcr.2016.02.012>

Digital Object Identifier (DOI):

[10.1016/j.fcr.2016.02.012](https://doi.org/10.1016/j.fcr.2016.02.012)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Field crops research

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Multi-scale modelling to synergise Plant Systems Biology and Crop Science

Yin Hoon Chew^{a1}, Daniel D. Seaton and Andrew J. Millar

SynthSys, C. H. Waddington Building, University of Edinburgh, King's Buildings, Edinburgh
EH9 3BF, UK

^aCorresponding author : Yin Hoon Chew

¹Present address : Icahn School of Medicine at Mount Sinai
Institute for Genomics and Multiscale Biology
1255 Fifth Avenue, Suite C2
New York, NY 10029
USA

Telephone number : +1 212 824 9673

Email : yinhood.chew@mssm.edu

18 **Abstract**

19 At the interface of the plant systems biology and crop modelling communities, a recurring
20 theme is the construction of an *in silico* plant that links across many levels of biological
21 organisation. These disciplines are not mutually exclusive; each has some elements of the other
22 and they have an overlapping goal in understanding and assisting crop improvement.
23 Therefore, we believe that synergies can be gained through knowledge exchange between the
24 two. Several modelling frameworks could support this aspiration. Our recent work on a
25 multiscale Arabidopsis Framework Model (FM) combined concepts from both systems biology
26 and crop modelling. We use the FM as a starting point to explore the potential benefits and
27 challenges of applying and extending such cross-disciplinary tools.

28

29 **Keywords:** systems biology; crop modelling; *in silico* plant; multiscale; climate change;
30 breeding

31

1.0 Introduction

The uncertain impact of climate change on crop yield has prompted various efforts to improve the predictive performance of crop models. Recent efforts included systematic comparisons of existing models and quantifying the uncertainty associated with each model, which has brought to fore the variety of model structures, assumptions and the approaches used in developing the models (Asseng et al., 2013; Rosenzweig et al., 2014). Studies have also extended beyond functional-structural plant models (FSPM) and process-based models (PBM), to consider organ level and genetic information that inform yield at the crop level (Chenu et al., 2009; Reymond et al., 2004).

The incorporation of genetic information that governs plant traits is not new. Gene actions were represented in earlier studies through linear estimates of effects on parameters to model different cultivars (Hoogenboom et al., 1997; White and Hoogenboom, 1996). More recent work linked crop models to quantitative trait loci (QTL), thus refining the representation to associated genomic regions (Reymond et al., 2003; Yin et al., 2005), though they are restricted to environmentally stable QTL (Chenu et al., 2009). Further refinement to molecular or gene network models has also been proposed (Hammer et al., 2006). The explosion of understanding in plant gene networks offers an opportunity to link the physiology of the plant and crop to mechanisms at the molecular level, and potentially thereafter to genome sequences.

Besides improving agricultural management, crop models have been useful in aiding crop breeding efforts (Chapman, 2008). Models can be used as preliminary screening tools to predict the performance of crossing existing cultivars before further tests are conducted, thus reducing the time taken to produce new cultivars. Recently, synthetic biology has been proposed as another alternative to conventional breeding, whereby varieties with desired traits are designed by modifying specific gene(s) through genetic engineering (Baltes and Voytas, 2015).

Evaluating alternative engineering strategies will often require quantitative models with explicit representation of the target gene networks, linked to their physiological functions. Moreover, the growth of genome sequences is expected in future to allow the mechanistic understanding of (some fraction of) the causal sequence variation between crop varieties. Again, linking the genomic data to quantitative mechanisms will require models that explicitly represent the functions of the relevant sequences.

Many physiological and genetic models have been developed for crops (Bogard et al., 2014; Gu et al., 2012), including a genetic network controlling wheat anthesis (Brown et al., 2013). However, quantitative models with gene networks, regulatory mechanisms and metabolisms are less common than in the model organism *Arabidopsis thaliana* (Lucas et al., 2011). One of the reasons is that crop models generally aim towards parsimony, i.e. capturing only the essential elements of environmental effects on plant performance. The availability of data was also uneven, as genetic tools to decipher molecular mechanisms have been relatively fewer in crop species; crop genomes are larger and more complex; and the timescales and/or facilities required for crop studies can be substantially greater. The majority of the plant gene functions discovered in the last decade have therefore been in *Arabidopsis*, making this the easiest species for mechanistic modelling, despite the considerable distance from crop models.

Emblematic examples of crop homologues for these *Arabidopsis* genes were quickly shown to underlie the high yield of crop varieties identified during the first Green Revolution (Ait-ali et al., 2003; Hedden and Kamiya, 1997; Peng et al., 1999; Sasaki et al., 2002). In recent years, many homologues of *Arabidopsis* genes have been found in crops (Chew and Halliday, 2011; Nakamichi, 2014). We are also seeing progressively more application of molecular and systems biology tools in crop studies (Kikuchi et al., 2003; Libault et al., 2010; Schmutz et al., 2010; Schnable et al., 2009). Different approaches have been suggested to make the best use of molecular plant science, systems biology and crop systems modelling in addressing the energy

crisis and food security issues (Hammer et al., 2004; Minorsky, 2003; Weckwerth, 2011; Yin and Struik, 2010). A recurring theme is to develop *in silico* or digital plants (Zhu et al., 2015), with integrative representation of gene functions at the molecular level while also linking across all biological levels of organisation, combining existing methods from crop dynamic modelling and the fast-emerging techniques in plant systems biology (Fig. 1).

This approach recognises the major opportunity for many more findings of basic plant research to support tangible crop improvement programmes. The considerable, financial and organisational constraints on the process have been reviewed elsewhere (RS, 2009). In this paper, we discuss whether and how multiscale models could provide a new technical avenue for research translation, with benefits for both fundamental and applied research. We propose the bridging of plant systems biology and crop systems modelling at various levels, where the different communities could complement one another (Fig. 1). As a case study, we discuss our recently published multiscale model of Arabidopsis growth (Chew et al., 2014) in relation to crop modelling, and where it stands relative to the *in silico* plant aspiration described above. We also present an example of the circadian clock gene network and its role in photoperiodism, because the molecular mechanisms are relatively well-understood, control important traits and have thus been studied in crops. Our aim here is not to propose this as the network of choice for genetic engineering, but to illustrate how systems biology models with molecular mechanisms may contribute towards crop designs for future breeding, including through synthetic biology. Similar principles could potentially be applied to many molecular processes that control crop traits.

2.0 The multiscale Arabidopsis Framework Model

The Arabidopsis Framework Model (FM) was developed using a modular approach by combining published models without modifying them. The FM consists of component models or modules that are characteristic of different research domains, such as molecular systems biology (gene regulatory network), crop science (functional-structural and source-sink relations at the organ level; phenology) and physiology (leaf-level photosynthesis; respiration, metabolism) (Fig. 1). Below, we briefly review each module, its links with crop science and translational potential.

2.1 The biological clock and photoperiodic response mechanisms

The module at the molecular level describes the gene circuit network of the circadian clock in Arabidopsis, which is one of the pervasive molecular networks regulating photosynthesis, metabolism and flowering time (Hotta et al., 2007). The circadian clock enables plants to perceive the duration of sunlight (photoperiod), an important cue for seasonal timing, so that plants can pace their daily biochemical reactions and developmental events to optimise growth and escape unfavourable environmental conditions (Millar, 2016; Simpson and Dean, 2002). The effects of photoperiod on developmental rates have been included in many crop models, even though no molecular basis was discovered until more recently (Nakamichi, 2014; Turner et al., 2005; Yano et al., 2000). Developmental transitions were related to seasonal changes, using a photoperiod-dependent scaling factor to modify the cumulative thermal time required for a developmental switch, measured in the photothermal units of phenology models (Robertson, 1968). For some crop species, photoperiod appears to be an obligatory signal; this is modelled using conditional functions where critical photoperiods must be exceeded before the model can proceed with successive events (Dingkuhn et al., 2008). In the case of the FM (Chew et al., 2014), the photoperiod-dependent scaling factor in the phenology module is a function of the expression level of a flowering gene (*FLOWERING LOCUS T*, *FT*). *FT* expression is linked to the clock gene circuit module, based on the well-characterised

photoperiod response network in Arabidopsis (Salazar et al., 2009). Here, gene expression is modelled using a set of ordinary differential equations (ODEs), where the change over time of each molecular component is simulated, along with their multiple responses to the light:dark cycle. This formulation not only provides a direct linkage between individual genes in the clock-photoperiodism pathways, environmental inputs and phenology, but also offers temporal resolution within the day, i.e. the scaling factor for photoperiodic response is continuously tuned to the dynamics of the flowering gene expression (Box 1). One advantage of this is that the changing photoperiod sensitivity within a day-night cycle can be captured, for example to compare the effects of different mutations for plants growing at different latitudes (Box 1). This capability may facilitate the understanding of Genotype x Environment interaction, specifically for crop varieties with genetic variation in the clock/photoperiodism genes (see Section 3.0) (Nakamichi, 2014).

2.2 Carbon metabolism

Metabolism is modelled in the FM in a simplified manner; only sugar (the transported carbon form) and starch (stored carbon) are considered (Gerakis et al., 2006). Sugar is transported from source organs to sink organs for growth and respiration, while part of the assimilated carbon is stored as starch during the day so that it can be converted into sugar at night when there is no photosynthesis. A recent study in Arabidopsis has revealed a circadian-clock control of starch consumption rate, which is paced so that there is enough starch to last the length of the night (Graf et al., 2010). This is currently represented in the FM using a simple division of total starch turnover by night period (Chew et al., 2014). More detailed models linking starch degradation to the clock gene network are also available (Pokhilko et al., 2014; Seaton et al., 2014). These models focus on a subset of the complex metabolic network: their scope could readily be broadened. In recent years, genome-scale metabolic models have been developed for several species, where the fluxes of all metabolites in a plant system are solved by

optimising within a set of constraints an objective function, e.g. maximising biomass or minimising total energy costs (Cheung et al., 2014; Grafahrend-Belau et al., 2013; Poolman et al., 2009). Such flux-balance modelling has been useful in identifying redundant or alternative pathways and their relative efficiencies, and can predict the shift in fluxes under different environmental conditions (Cheung et al., 2013; Williams et al., 2010), potentially with tissue resolution (Rolletschek et al., 2015). This level of detail is understandably not included in any crop models, nor is it usually considered in conventional breeding. Nevertheless, metabolic modelling can help identify target enzymes for breeding or genetic manipulation, to achieve desired metabolic outputs thus facilitating metabolic engineering (Schwender, 2008). A major limitation compared to single-cell organisms is that plant systems span a wide range of spatio-temporal scales. Changing a target enzyme may not be sufficient to translate metabolic gains into commensurate benefits at the whole-organism level (Leakey et al., 2012). The second opportunity for modelling is therefore to optimise the effect of modification at the genetic level on events at the whole-plant level and beyond. In this respect, models like the FM can provide a framework for linking metabolic models of increasing breadth to processes at the higher levels, to aid molecular breeding or synthetic biology.

2.3 Plant architecture

The FM also considers plant architecture, specifically shoot structure and its function in light interception for carbon acquisition. This is modelled using a functional-structural module where individual organ growth and leaf orientation are captured. The significance of plant architecture in crop yield has been explored; one good example is the extension of a 3D canopy model of wheat development to include mechanistic details such as organ-level photosynthesis and allocation of assimilate (Evers et al., 2010). Simulating canopy structure is useful in understanding competition for light, both inter-species and intra-species. Thus, crop models that consider plant form and function can be valuable for maximising crop yield per unit area

to optimise land use (Egli, 1988; Mao et al., 2014). In the case of Arabidopsis, this plant species has a simple rosette structure during the vegetative phase but upon switching to the reproductive phase, stem elongation and branching occur, forming a more complex inflorescence. The current version of the FM only models the simplest, vegetative phase. Including the developmental phases important for crops will allow future versions of the model to address the resource partitioning issues that are particularly relevant for translational research. An extension to include the functional-structural contribution of the inflorescence until fruit (and thus seed) formation is necessary (Christophe et al., 2008). These models would require comprehensive phenotypic and physiological data, some of which are available (Alonso-Blanco et al., 1999; Bennett et al., 2012; Diggle, 1997). Although the area seems under-explored in Arabidopsis relative to crop species, these studies have already revealed interesting correlations between branching and seed yield, which also depend on the plant varieties, and identified some related QTL regions. Indeed, a recent mechanistic model of sucrose transport in rice also highlighted the effect of grain arrangement on yield (Seki et al., 2015), suggesting that the genetic tools to manipulate carbon partitioning in Arabidopsis will allow future models to contribute further quantitative understanding of this area.

In contrast, plant development at the cellular, tissue and organ levels has been intensively studied in Arabidopsis. Many of the molecular mechanisms regulating cell geometry, tissue patterning and organ morphology have been elucidated using mathematical modelling (Roeder et al., 2010; Sauret-Gueto et al., 2013; Torii, 2012), which therefore provide opportunities for linking with the FM. As the FM already has a simple representation at the organ level, incorporating more detailed morphology will be a natural extension. Including geometry at the cellular level will require a major change in the resolution and hence in the modelling methods. Nevertheless, such extensions could provide a platform to further improve yield through altered

architecture, as a continuation of what has been achieved during the first Green Revolution (Khush, 2001).

2.4 Phenology

Vegetative development is simulated in the FM by a phenology module (Chew et al., 2012; Wilczek et al., 2009). This module utilises the thermal time concept, which is adapted from crop modelling approaches. Three components contribute: thermal time, photoperiod (see above) and vernalisation. Degree-days are integrated, scaled by the photoperiod and vernalisation factors, until a threshold value is exceeded to indicate the switch from the vegetative to the reproductive phase. Additionally, the FM also simulates the change in phyllochron, which is another phenological event within the vegetative phase, driven by thermal time. To model the whole life cycle, this module will need to include floral initiation, fruit formation, seed development and maturity, and plant senescence after which seeds can be harvested. Seed dormancy has recently been addressed (Springthorpe and Penfield, 2015).

Crop modelling techniques that already consider various developmental stages again offer valuable examples. It is worth noting here that ontogeny is actually a continuous process. The ‘discrete ontogeny’ of functional stages allows plant development to be mathematically represented as a sequence of states of finite automata (Komarov et al., 2003). Molecular events that govern plant ontogeny are also continuous, and are often modelled as such in systems biology, though hybrid discrete-continuous models have also been adopted such as for systems exhibiting multiscale dynamics (Lincoln and Tiwari, 2004). Indeed, this is the approach used in the FM, where the continuous gene regulatory module (see above and **Box 1**) was linked to the phenology module. We expect more such hybrid links in the future, especially when combining crop modelling concepts with those in systems biology.

The diverse technical approaches, which are natural in models from different research domains, can form a significant barrier to re-using large models. Model exchange standards such as SBML and CellML greatly facilitated progress in some areas of computational systems biology, through model sharing and dissemination (Chew et al., 2014; Le Novere, 2006; Le Novere, 2015). Further tools are required to increase computational efficiency and facilitate model analysis (Qiu et al., 2014), in addition to sharing experimental data (Leonelli et al., 2013). Furthermore, there are also platforms specific for the plant and agricultural research community (reviewed in Adam et al., 2012) like OpenAlea, APSIM and RECORD that support model composition (Bergez et al., 2013; Brown et al., 2014; Pradal et al., 2008). The FM used one such tool, Simile (Muetzelfeldt and Massheder, 2003), to refactor all four modules into a common format. These tools could mean that crop models and contemporary systems biology models can be more conveniently exchanged, reused and combined.

2.5 Photosynthetic engineering

One of the processes central to plant yield is photosynthesis. Crop models like APSIM, Sirius or CERES can effectively simulate growth based on organ demand or plant water status (Hammer et al., 2010; Hoogenboom et al., 1994; Jamieson et al., 1998). In these models, carbon assimilation (or dry mass accumulation) is explicitly modelled but crop growth is driven by phenological development and organ expansion. These approaches also offer the potential to link with systems biology, for as plant water relations (reviewed in (Tardieu et al., 2015); (Parent and Tardieu, 2014)) involve the intensively-studied guard cell signalling network, which was recently modelled in detail (Chen et al., 2012). The FM centres on carbon balancing with a photosynthetic representation using the classical Farquhar model (Farquhar et al., 1980) to simulate photosynthesis. The Farquhar model is commonly used in ecosystem models of carbon cycling, though it has also been adopted in some crop models (e.g. GECROS (Yin and van Laar, 2005)). This model and its later developments consider rate-limiting reactions like

carboxylation, electron transport and inorganic phosphate (Pi), using kinetic parameters measured at the leaf level (Farquhar and von Caemmerer, 1982; Sharkey, 1985). More recently, new tools in molecular, physiology, biochemistry and systems biology have facilitated the development of a more detailed model, e-Photosynthesis (Zhu et al., 2013), which has explicit description of individual reactions and major regulatory processes at the molecular level in the photosynthesis network. Models with such details have identified molecular components/pathways for improving photosynthetic capacity, which could serve as a blueprint for the bioengineering of crop species (Wang et al., 2014). The FM's modular approach provides an ideal platform to evaluate alternative modules, such as the coarse-grained Farquhar and fine-grained e-Photosynthesis models. A fine-grained model can explicitly represent the target enzymes and processes for engineering at the molecular level, providing the bridging platform to understand and optimise how candidate manipulations propagate into changes at the whole-plant level.

3.0 Classical breeding and directed design

The crops that we consume today have gone through centuries of human cultivation, during which one of the selection aims was to expand the regions where the plants can be grown and produce high yields. The mutations that were selected in the photoperiod pathways include variants in the crop homologues of several Arabidopsis circadian clock genes (see (Nakamichi, 2014) for a comprehensive review). Understanding in Arabidopsis has become the reference point for determining the functions of such mutated genes in crop cultivars (Box 1), such that the function of a growing number of genome sequence variants might in future be explained by reference to this system. A multiscale model such as the FM could extend the scope of such understanding, even if the other physiological processes reviewed above remain simply

represented in the model. The potential for synergy will grow as further modules gain mechanistic detail, such as the circadian regulation of starch metabolism noted above, especially if mechanisms identified and modelled in model species translate readily to crops.

As the demand for food and biofuel increases, the pressure of climate change could also restrict the temporal window for flowering (or other developmental events). A model-assisted breeding strategy, where crop cultivars optimised for specific regions are specially designed and bred or engineered, may open new avenues for local adaptation (Fernandez-Cornejo and Caswell, 2006). The mechanistic, molecular models would be a relatively small part of this process. Accurate climate models will be required to provide projected climate inputs (Gleckler et al., 2008; Knutti et al., 2010), though this is also an active research area (Duben et al., 2014; Refsgaard et al., 2014). Latitude-specific ‘maturity groups’ are well established in some crops, such as soybean (Boerma and Specht, 2004), but for other crops this would represent a major change in commercialisation. The *in silico* plant must therefore be considered within a broad research and translation agenda.

4.0 Concluding remarks

Our initial work on the Arabidopsis Framework Model is unusual in representing both plant systems biology and elements drawn from crop systems modelling. The FM might be seen as a ‘boundary object’ in sociological terms, at the interface of several research communities, potentially facilitating communication (Star and Griesemer, 1989). While the FM will clearly benefit from extension, it offers a medium to integrate the different types of understanding from fundamental plant research and crop models. Tools of this type also have significant potential to contribute to practical crop improvement, as one of several promising alternatives

that might be linked by umbrella structures such as the Joint Programming Initiative on Agriculture, Food Security and Climate Change (FACCE-JPI).

Acknowledgements

Our work was supported by European Commission FP7 collaborative project TiMet (contract no. 245143 to AJM and others).

Figure 1: *In silico* plant linking across biological levels of organisation. Plants are complex systems that cross multiple, interactive levels of organisation (column A). Each discipline (C) usually focuses on processes at specific level(s) (B). Shaded boxes indicate processes currently considered in our multiscale Arabidopsis Framework Model (Chew et al., 2014). Different modelling communities (D) could interact and work together at various levels, where synergy might be gained.

Box 1. Connecting molecular circuits to flowering phenotypes. The regulatory mechanisms underlying the sensitivity of flowering time to photoperiod have been characterised at a molecular level. This is illustrated in (A), showing a simplified representation of the regulatory interactions involved in the circadian clock (orange) and the photoperiod pathway (green), culminating in the regulation of *FT*. The dynamics of these networks have previously been described by mathematical models (Pokhilko et al., 2012; Salazar et al., 2009; Song et al., 2012), and were recently combined and updated in a single model (Seaton et al., 2015). This allows simulation of the behaviour of this network in different natural photoperiods at different times of year, for example those occurring in Halle, Germany (B). At different times of year,

the model simulates different levels of *FT* at this location (C). Thus, given a date of seedling emergence, the date of flowering can be predicted (D). Flowering times for different genotypes can also be predicted, e.g. in the case of the mutants *prp9;prp7* and *elf3* (knock-out mutants of circadian clock genes with orthologues in crop species (Staiger et al., 2013)). This allows simulation of a population of plants that show a distribution of dates of emergence. As shown in (E), the sensitivity of the wild-type plants to photoperiod means that flowering occurs over a narrower distribution of dates than either the late-flowering (*prp9;prp7*) or early-flowering (*elf3*) circadian clock mutants.

References

- Adam, M., Corbeels, M., Leffelaar, P.A., Van Keulen, H., Wery, J., Ewert, F., 2012. Building crop models within different crop modelling frameworks. *Agricultural Systems* 113, 57-63
- Ait-ali, T., Rands, C., Harberd, N.P., 2003. Flexible control of plant architecture and yield via switchable expression of *Arabidopsis* *gai*. *Plant Biotechnol J* 1, 337-343
- Alonso-Blanco, C., Blankestijn-de Vries, H., Hanhart, C.J., Koornneef, M., 1999. Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* 96, 4710-4717
- Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rotter, R.P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P.K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A.J., Doltra, J., Gayler, S., et al., 2013. Uncertainty in simulating wheat yields under climate change. *Nat Clim Change* 3, 827-832
- Baltes, N.J., Voytas, D.F., 2015. Enabling plant synthetic biology through genome engineering. *Trends Biotechnol* 33, 120-131
- Bennett, E., Roberts, J.A., Wagstaff, C., 2012. Manipulating resource allocation in plants. *J Exp Bot* 63, 3391-3400
- Bergez, J.E., Chabrier, P., Gary, C., Jeuffroy, M.N., Makowski, D., Quesnel, G., Ramat, E., Raynal, H., Rousse, N., Wallach, D., Debaeke, P., Durand, P., Duru, M., Dury, J., Faverdin, P., Gascuel-Oudoux, C., Garcia, F., 2013. An open platform to build, evaluate and simulate integrated models of farming and agro-ecosystems. *Environ Modell Softw* 39, 39-49
- Boerma, H.R., Specht, J.E. Soybeans: improvement, production and uses. (2004), American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc., Madison, Wisconsin, USA. pp. 453-454.
- Bogard, M., Ravel, C., Paux, E., Bordes, J., Balfourier, F., Chapman, S.C., Le Gouis, J., Allard, V., 2014. Predictions of heading date in bread wheat (*Triticum aestivum* L.) using QTL-based parameters of an ecophysiological model. *J Exp Bot* 65, 5849-5865
- Brown, H.E., Jamieson, P.D., Brooking, I.R., Moot, D.J., Huth, N.I., 2013. Integration of molecular and physiological models to explain time of anthesis in wheat. *Ann Bot-London* 112, 1683-1703

- Brown, H.E., Huth, N.I., Holzworth, D.P., Teixeira, E.I., Zyskowski, R.F., Hargreaves, J.N.G., Moot, D.J., 2014. Plant Modelling Framework: Software for building and running crop models on the APSIM platform. *Environ Modell Softw* 62, 385-398
- Chapman, S.C., 2008. Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials. *Euphytica* 161, 195-208
- Chen, Z.H., Hills, A., Batz, U., Amtmann, A., Lew, V.L., Blatt, M.R., 2012. Systems dynamic modeling of the stomatal guard cell predicts emergent behaviors in transport, signaling, and volume control. *Plant Physiol* 159, 1235-51
- Chenu, K., Chapman, S.C., Tardieu, F., McLean, G., Welcker, C., Hammer, G.L., 2009. Simulating the Yield Impacts of Organ-Level Quantitative Trait Loci Associated With Drought Response in Maize: A "Gene-to-Phenotype" Modeling Approach. *Genetics* 183, 1507-1523
- Cheung, C.Y.M., Poolman, M.G., Fell, D.A., Ratcliffe, R.G., Sweetlove, L.J., 2014. A Diel Flux Balance Model Captures Interactions between Light and Dark Metabolism during Day-Night Cycles in C-3 and Crassulacean Acid Metabolism Leaves. *Plant Physiol* 165, 917-929
- Cheung, C.Y.M., Williams, T.C.R., Poolman, M.G., Fell, D.A., Ratcliffe, R.G., Sweetlove, L.J., 2013. A method for accounting for maintenance costs in flux balance analysis improves the prediction of plant cell metabolic phenotypes under stress conditions. *Plant J* 75, 1050-1061
- Chew, Y.H., Halliday, K.J., 2011. A stress-free walk from Arabidopsis to crops. *Curr Opin Biotech* 22, 281-286
- Chew, Y.H., Wilczek, A.M., Williams, M., Welch, S.M., Schmitt, J., Halliday, K.J., 2012. An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytol* 194, 654-665
- 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, E4127-E4136
- Christophe, A., Letort, V., Hummel, I., Cournede, P.H., de Reffye, P., Lecoeur, J., 2008. A model-based analysis of the dynamics of carbon balance at the whole-plant level in Arabidopsis thaliana. *Funct Plant Biol* 35, 1147-1162
- Diggle, P.K., 1997. Ontogenetic contingency and floral morphology: The effects of architecture and resource limitation. *Int J Plant Sci* 158, S99-S107
- Dingkuhn, M., Kouressy, M., Vaksman, M., Clerget, B., Chantreau, J., 2008. A model of sorghum photoperiodism using the concept of threshold-lowering during prolonged appetite. *Eur J Agron* 28, 74-89
- Duben, P.D., McNamara, H., Palmer, T.N., 2014. The use of imprecise processing to improve accuracy in weather & climate prediction. *J Comput Phys* 271, 2-18
- Egli, D.B., 1988. Plant-Density and Soybean Yield. *Crop Sci* 28, 977-981
- Evers, J.B., Vos, J., Yin, X., Romero, P., van der Putten, P.E.L., Struik, P.C., 2010. Simulation of wheat growth and development based on organ-level photosynthesis and assimilate allocation. *J Exp Bot* 61, 2203-2216
- Farquhar, G.D., von Caemmerer, S., 1982. Modelling of photosynthetic response to environmental conditions. *Physiological plant ecology II* 12/B, 549 - 587
- Farquhar, G.D., Caemmerer, S.V., Berry, J.A., 1980. A Biochemical-Model of Photosynthetic Co2 Assimilation in Leaves of C-3 Species. *Planta* 149, 78-90
- Fernandez-Cornejo, J., Caswell, M. The First Decade of Genetically Engineered Crops in the United States. (2006), *Economic Information Bulletin, USDA*. pp. 36.
- Gerakis, A., Rasse, D.P., Kavdir, Y., Smucker, A.J.M., Katsalirou, I., Ritchie, J.T., 2006. Simulation of leaching losses in the nitrogen cycle. *Commun Soil Sci Plan* 37, 1973-1997
- Gleckler, P.J., Taylor, K.E., Doutriaux, C., 2008. Performance metrics for climate models. *J Geophys Res-Atmos* 113

409 Graf, A., Schlereth, A., Stitt, M., Smith, A.M., 2010. Circadian control of carbohydrate availability for
 410 growth in Arabidopsis plants at night. *Proceedings of the National Academy of Sciences of the*
 411 *United States of America* 107, 9458-9463
 412 Grafahrend-Belau, E., Junker, A., Eschenroder, A., Muller, J., Schreiber, F., Junker, B.H., 2013.
 413 Multiscale Metabolic Modeling: Dynamic Flux Balance Analysis on a Whole-Plant Scale. *Plant*
 414 *Physiol* 163, 637-647
 415 Gu, J.F., Yin, X.Y., Struik, P.C., Stomph, T.J., Wang, H.Q., 2012. Using chromosome introgression lines
 416 to map quantitative trait loci for photosynthesis parameters in rice (*Oryza sativa* L.) leaves
 417 under drought and well-watered field conditions. *J Exp Bot* 63, 455-469
 418 Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., Chapman, S., Podlich, D.,
 419 2006. Models for navigating biological complexity in breeding improved crop plants. *Trends*
 420 *Plant Sci* 11, 587-593
 421 Hammer, G.L., Sinclair, T.R., Chapman, S.C., van Oosterom, E., 2004. On systems thinking, systems
 422 biology, and the in silico plant. *Plant Physiol* 134, 909-911
 423 Hammer, G.L., van Oosterom, E., McLean, G., Chapman, S.C., Broad, I., Harland, P., Muchow, R.C.,
 424 2010. Adapting APSIM to model the physiology and genetics of complex adaptive traits in field
 425 crops. *J Exp Bot* 61, 2185-2202
 426 Hedden, P., Kamiya, Y., 1997. Gibberellin biosynthesis: Enzymes, genes and their regulation. *Annu Rev*
 427 *Plant Phys* 48, 431-460
 428 Hoogenboom, G., White, J.W., Jones, J.W., Boote, K.J., 1994. Beangro - a Process-Oriented Dry Bean
 429 Model with a Versatile User-Interface. *Agron J* 86, 182-190
 430 Hoogenboom, G., White, J.W., AcostaGallegos, J., Gaudiel, R.G., Myers, J.R., Silbernagel, M.J., 1997.
 431 Evaluation of a crop simulation model that incorporates gene action. *Agron J* 89, 613-620
 432 Hotta, C.T., Gardner, M.J., Hubbard, K.E., Baek, S.J., Dalchau, N., Suhita, D., Dodd, A.N., Webb, A.A.R.,
 433 2007. Modulation of environmental responses of plants by circadian clocks. *Plant Cell Environ*
 434 30, 333-349
 435 Jamieson, P.D., Semenov, M.A., Brooking, I.R., Francis, G.S., 1998. Sirius: a mechanistic model of wheat
 436 response to environmental variation. *Eur J Agron* 8, 161-179
 437 Khush, G.S., 2001. Green revolution: the way forward. *Nat Rev Genet* 2, 815-822
 438 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M.,
 439 Yamada, H., Ooka, H., Hotta, I., 2003. Collection, mapping, and annotation of over 28,000
 440 cDNA clones from japonica rice (vol 301, pg 376, 2003). *Science* 301, 1849-1849
 441 Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., Meehl, G.A., 2010. Challenges in Combining Projections
 442 from Multiple Climate Models. *J Climate* 23, 2739-2758
 443 Komarov, A.S., Palenova, M.M., Smirnova, O.V., 2003. The concept of discrete description of plant
 444 ontogenesis and cellular automata models of plant populations. *Ecol Model* 170, 427-439
 445 Le Novere, N., 2006. Model storage, exchange and integration. *Bmc Neurosci* 7
 446 Le Novere, N., 2015. Quantitative and logic modelling of molecular and gene networks. *Nat Rev Genet*
 447 16, 146-158
 448 Leakey, A.D.B., Ainsworth, E.A., Bernacchi, C.J., Zhu, X.G., Long, S.P., Ort, D.R., 2012. Photosynthesis in
 449 a CO₂-Rich Atmosphere. *Adv Photosynth Resp* 34, 733-768
 450 Leonelli, S., Smirnoff, N., Moore, J., Cook, C., Bastow, R., 2013. Making open data work for plant
 451 scientists. *J Exp Bot* 64, 4109-17
 452 Libault, M., Farmer, A., Joshi, T., Takahashi, K., Langley, R.J., Franklin, L.D., He, J., Xu, D., May, G.,
 453 Stacey, G., 2010. An integrated transcriptome atlas of the crop model Glycine max, and its use
 454 in comparative analyses in plants. *Plant J* 63, 86-99
 455 Lincoln, P., Tiwari, A., 2004. Symbolic systems biology: Hybrid modeling and analysis of biological
 456 networks. *Lect Notes Comput Sc* 2993, 660-672
 457 Lucas, M., Laplace, L., Bennett, M.J., 2011. Plant systems biology: network matters. *Plant Cell Environ*
 458 34, 535-53

459 Mao, L.L., Zhang, L.Z., Zhao, X.H., Liu, S.D., van der Werf, W., Zhang, S.P., Spiertz, H., Li, Z.H., 2014.
 460 Crop growth, light utilization and yield of relay intercropped cotton as affected by plant
 461 density and a plant growth regulator. *Field Crop Res* 155, 67-76

462 Millar, A.J., 2016. The intracellular dynamics of circadian clocks reach for the light of ecology and
 463 evolution. *Annual Review of Plant Biology* In Press

464 Minorsky, P.V., 2003. Achieving the in silico plant. *Systems biology and the future of plant biological*
 465 *research. Plant Physiol* 132, 404-409

466 Muetzelfeldt, R., Massheder, J., 2003. The Simile visual modelling environment. *European Journal of*
 467 *Agronomy* 18, 345-358

468 Nakamichi, N., 2014. Adaptation to local environment by modifications of photoperiod response in
 469 crops. *Plant and Cell Physiology*, 1 - 11

470 Parent, B., Tardieu, F., 2014. Can current crop models be used in the phenotyping era for predicting
 471 the genetic variability of yield of plants subjected to drought or high temperature? *J Exp Bot*
 472 65, 6179-89

473 Peng, J.R., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flintham, J.E., Beales, J., Fish, L.J.,
 474 Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D., Harberd, N.P.,
 475 1999. 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400,
 476 256-261

477 Pokhilko, A., Flis, A., Sulpice, R., Stitt, M., Ebenhoh, O., 2014. Adjustment of carbon fluxes to light
 478 conditions regulates the daily turnover of starch in plants: a computational model. *Mol Biosyst*
 479 10, 613-627

480 Pokhilko, A., Fernandez, A.P., Edwards, K.D., Southern, M.M., Halliday, K.J., Millar, A.J., 2012. The clock
 481 gene circuit in *Arabidopsis* includes a repressilator with additional feedback loops. *Mol Syst*
 482 *Biol* 8

483 Poolman, M.G., Miguet, L., Sweetlove, L.J., Fell, D.A., 2009. A Genome-Scale Metabolic Model of
 484 *Arabidopsis* and Some of Its Properties. *Plant Physiol* 151, 1570-1581

485 Pradal, C., Dufour-Kowalski, S., Boudon, F., Fournier, C., Godin, C., 2008. OpenAlea: a visual
 486 programming and component-based software platform for plant modelling. *Funct Plant Biol*
 487 35, 751-760

488 Qiu, J., Foster, I., Goble, C., 2014. Emerging Computational Methods for the Life Sciences Workshop
 489 2012. *Concurr Comp-Pract E* 26, 1231-1233

490 Refsgaard, J.C., Madsen, H., Andreassian, V., Arnbjerg-Nielsen, K., Davidson, T.A., Drews, M., Hamilton,
 491 D.P., Jeppesen, E., Kjellstrom, E., Olesen, J.E., Sonnenborg, T.O., Trolle, D., Willems, P.,
 492 Christensen, J.H., 2014. A framework for testing the ability of models to project climate change
 493 and its impacts. *Climatic Change* 122, 271-282

494 Reymond, M., Muller, B., Tardieu, F., 2004. Dealing with the genotypexenvironment interaction via a
 495 modelling approach: a comparison of QTLs of maize leaf length or width with QTLs of model
 496 parameters. *J Exp Bot* 55, 2461-2472

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

500 Robertson, G.W., 1968. A Biometeorological Time Scale for a Cereal Crop Involving Day and Night
 501 Temperatures and Photoperiod. *International Journal of Biometeorology* 12, 191 - 223

502 Roeder, A.H.K., Chickarmane, V., Cunha, A., Obara, B., Manjunath, B.S., Meyerowitz, E.M., 2010.
 503 Variability in the Control of Cell Division Underlies Sepal Epidermal Patterning in *Arabidopsis*
 504 *thaliana*. *Plos Biol* 8

505 Rolletschek, H., Grafahrend-Belau, E., Munz, E., Radchuk, V.V., Kartausch, R., Tschiersch, H., Melkus,
 506 G., Schreiber, F., Jakob, P.M., Borisjuk, L., 2015. Metabolic architecture of the cereal grain and
 507 its relevance to maximize carbon use efficiency. *Plant Physiol*

508 Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Muller, C., Arneth, A., Boote, K.J., Folberth, C.,
 509 Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T.A., Schmid, E., Stehfest, E., Yang,

H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences of the United States of America* 111, 3268-73

RS Reaping the benefits: Science and the sustainable intensification of global agriculture. (2009), The Royal Society, London.

Salazar, J.D., Saithong, T., Brown, P.E., Foreman, J., Locke, J.C.W., Halliday, K.J., Carre, I.A., Rand, D.A., Millar, A.J., 2009. Prediction of Photoperiodic Regulators from Quantitative Gene Circuit Models. *Cell* 139, 1170-1179

Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapan, D., Ishiyama, K., Saito, T., Kobayashi, M., Khush, G.S., Kitano, H., Matsuoka, M., 2002. Green revolution: A mutant gibberellin-synthesis gene in rice - New insight into the rice variant that helped to avert famine over thirty years ago. *Nature* 416, 701-702

Sauret-Gueto, S., Schiessl, K., Bangham, A., Sablowski, R., Coen, E., 2013. JAGGED Controls Arabidopsis Petal Growth and Shape by Interacting with a Divergent Polarity Field. *Plos Biol* 11

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J.X., Mitros, T., Nelson, W., Hyten, D.L., Song, Q.J., Thelen, J.J., Cheng, J.L., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., *et al.*, 2010. Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178-183

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F.S., Pasternak, S., Liang, C.Z., Zhang, J.W., Fulton, L., Graves, T.A., Minx, P., Reily, A.D., Courtney, L., Kruchowski, S.S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S.M., *et al.*, 2009. The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science* 326, 1112-1115

Schwender, J., 2008. Metabolic flux analysis as a tool in metabolic engineering of plants. *Curr Opin Biotech* 19, 131-137

Seaton, D.D., Ebenhoh, O., Millar, A.J., Pokhilko, A., 2014. Regulatory principles and experimental approaches to the circadian control of starch turnover. *J R Soc Interface* 11

Seaton, D.D., Smith, R.W., Song, Y.H., MacGregor, D.R., Stewart, K., Steel, G., Foreman, J., Penfield, S., Imaizumi, T., Millar, A.J., Halliday, K.J., 2015. Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. *Mol Syst Biol* 11

Seki, M., Feugier, F.G., Song, X.J., Ashikari, M., Nakamura, H., Ishiyama, K., Yamaya, T., Inari-Ikeda, M., Kitano, H., Satake, A., 2015. A Mathematical Model of Phloem Sucrose Transport as a New Tool for Designing Rice Panicle Structure for High Grain Yield. *Plant and Cell Physiology*, 1 - 15

Sharkey, T.D., 1985. Photosynthesis in Intact Leaves of C-3 Plants - Physics, Physiology and Rate Limitations. *Bot Rev* 51, 53-105

Simpson, G.G., Dean, C., 2002. Flowering - Arabidopsis, the rosetta stone of flowering time? *Science* 296, 285-289

Song, Y.H., Smith, R.W., To, B.J., Millar, A.J., Imaizumi, T., 2012. FKF1 Conveys Timing Information for CONSTANS Stabilization in Photoperiodic Flowering. *Science* 336, 1045-1049

Springthorpe, V., Penfield, S., 2015. Flowering time and seed dormancy control use external coincidence to generate life history strategy. *Elife* 4

Staiger, D., Shin, J., Johansson, M., Davis, S.J., 2013. The circadian clock goes genomic. *Genome Biol* 14

Star, S.L., Griesemer, J.R., 1989. Institutional Ecology, Translations and Boundary Objects - Amateurs and Professionals in Berkeleys-Museum-of-Vertebrate-Zoology, 1907-39. *Social Studies of Science* 19, 387 - 420

Tardieu, F., Simonneau, T., Parent, B., 2015. Modelling the coordination of the controls of stomatal aperture, transpiration, leaf growth, and abscisic acid: update and extension of the Tardieu-Davies model. *J Exp Bot* 66, 2227-37

Torii, K.U., 2012. Two-dimensional spatial patterning in developmental systems. *Trends Cell Biol* 22, 438-446

- Turner, A., Beales, J., Faure, S., Dunford, R.P., Laurie, D.A., 2005. The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science* 310, 1031-1034
- Wang, Y., Long, S.P., Zhu, X.G., 2014. Elements Required for an Efficient NADP-Malic Enzyme Type C4 Photosynthesis. *Plant Physiol* 164, 2231-2246
- Weckwerth, W., 2011. Green systems biology - From single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *J Proteomics* 75, 284-305
- White, J.W., Hoogenboom, G., 1996. Simulating effects of genes for physiological traits in a process-oriented crop model. *Agron J* 88, 416-422
- Wilczek, A.M., Roe, J.L., Knapp, M.C., Cooper, M.D., Lopez-Gallego, C., Martin, L.J., Muir, C.D., Sim, S., Walker, A., Anderson, J., Egan, J.F., Moyers, B.T., Petipas, R., Giakountis, A., Charbit, E., Coupland, G., Welch, S.M., Schmitt, J., 2009. Effects of Genetic Perturbation on Seasonal Life History Plasticity. *Science* 323, 930-934
- Williams, T.C.R., Poolman, M.G., Howden, A.J.M., Schwarzlander, M., Fell, D.A., Ratcliffe, R.G., Sweetlove, L.J., 2010. A Genome-Scale Metabolic Model Accurately Predicts Fluxes in Central Carbon Metabolism under Stress Conditions. *Plant Physiol* 154, 311-323
- Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., Baba, T., Yamamoto, K., Umehara, Y., Nagamura, Y., Sasaki, T., 2000. Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the arabidopsis flowering time gene CONSTANS. *Plant Cell* 12, 2473-2483
- Yin, X., van Laar, H.H. (2005) Crop systems dynamics: an ecophysiological simulation model for genotype-by-environment interactions Wageningen Publishers, Wageningen.
- Yin, X.Y., Struik, P.C., 2010. Modelling the crop: from system dynamics to systems biology. *J Exp Bot* 61, 2171-2183
- Yin, X.Y., Struik, P.C., van Eeuwijk, F.A., Stam, P., Tang, J.J., 2005. QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley. *J Exp Bot* 56, 967-976
- Zhu, X.G., Wang, Y., Ort, D.R., Long, S.P., 2013. e-photosynthesis: a comprehensive dynamic mechanistic model of C3 photosynthesis: from light capture to sucrose synthesis. *Plant Cell Environ* 36, 1711-1727
- Zhu, X.G., Lynch, J.P., LeBauer, D.S., Millar, A.J., Stitt, M., Long, S.P., 2015. Plants in silico: Why, Why Now and Framework - An integrative platform for plant systems biology research. *Plant Cell and Environment* In Press

Box 1. Connecting molecular circuits to flowering phenotypes. The regulatory mechanisms underlying the sensitivity of flowering time to photoperiod have been characterised at a molecular level. This is illustrated in (A), showing a simplified representation of the regulatory interactions involved in the circadian clock (orange) and the photoperiod pathway (green), culminating in the regulation of *FT*. The dynamics of these networks have previously been described by mathematical models (Salazar *et al.*, 2009; Pokhilko *et al.*, 2012; Song *et al.*, 2012), and were recently combined and updated in a single model (Seaton *et al.*, 2015). This allows simulation of the behaviour of this network in different natural photoperiods at different times of year, for example those occurring in Halle, Germany (B). At different times of year, the model simulates different levels of *FT* at this location (C). Thus, given a date of seedling emergence, the date of flowering can be predicted (D). Flowering times for different genotypes can also be predicted, e.g. in the case of the mutants *prp9;prp7* and *elf3* (knock-out mutants of circadian clock genes with orthologs in crop species (Staiger *et al.*, 2013)). This allows simulation of a population of plants that show a distribution of dates of emergence. As shown in (E), the sensitivity of the wild-type plants to photoperiod means that flowering occurs over a narrower distribution of dates than either the late-flowering (*prp9;prp7*) or early-flowering (*elf3*) circadian clock mutants.