Field Crops Research 202 (2017) 77–83

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

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

Yin Hoon Chew*, Daniel D. Seaton, Andrew J. Millar

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

ARTICLE INFO

Article history:
Received 18 June 2015
Received in revised form
24 December 2015
Accepted 12 February 2016
Available online 26 February 2016

Keywords:
Systems biology
Crop modelling
In silico plant
Multiscale
Climate change
Breeding

ABSTRACT

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

1. 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,

---

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

E-mail address: yinhoon.chew@mssm.edu (Y.H. Chew).

http://dx.doi.org/10.1016/j.fcr.2016.02.012
0378-4290/

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. 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