in silico Plants
OXFORD UNIVERSITY PRESS

Opinion

Biological reality and parsimony in crop models—why we need both in crop improvement!

Graeme Hammer*[1], Charlie Messina[2], Alex Wu[1] and Mark Cooper[1]

[1]Queensland Alliance for Agriculture and Food Innovation, Queensland Bioscience Precinct, The University of Queensland, 306 Carmody Road, St Lucia, Queensland 4067, Australia
[2]Corteva Agriscience, 8503 NW 62nd Avenue, Johnston, IA 50131, USA

Received: 6 August 2019 **Accepted:** 2 October 2019

Abstract. The potential to add significant value to the rapid advances in plant breeding technologies associated with statistical whole-genome prediction methods is a new frontier for crop physiology and modelling. Yield advance by genetic improvement continues to require prediction of phenotype based on genotype, and this remains challenging for complex traits despite recent advances in genotyping and phenotyping. Crop models that capture physiological knowledge and can robustly predict phenotypic consequences of genotype-by-environment-by-management (G×E×M) interactions have demonstrated potential as an integrating tool. But does this biological reality come with a degree of complexity that restricts applicability in crop improvement? Simple, high-speed, parsimonious models are required for dealing with the thousands of genotypes and environment combinations in modern breeding programs utilizing genomic prediction technologies. In contrast, it is often considered that greater model complexity is needed to evaluate potential of putative variation in specific traits in target environments as knowledge on their underpinning biology advances. Is this a contradiction leading to divergent futures? Here it is argued that biological reality and parsimony do not need to be independent and perhaps should not be. Models structured to readily allow variation in the biological level of process algorithms, while using coding and computational advances to facilitate high-speed simulation, could well provide the structure needed for the next generation of crop models needed to support and enhance advances in crop improvement technologies. Beyond that, the trans-scale and transdisciplinary dialogue among scientists that will be required to construct such models effectively is considered to be at least as important as the models.

Keywords: Crop improvement; crop model; genomic prediction; phenotypic prediction; plant breeding.

Introduction

Based on current rates of yield improvement for major crops, comprehensive analyses have shown that by 2050, there will be a significant shortfall in global food production capacity (Mueller *et al.* 2012; Ray *et al.* 2013; Fischer *et al.* 2014). It is imperative that we hasten yield advance. The world population can no longer take for granted an ability to feed itself in 2050 and beyond. This will necessarily involve the sustainable intensification of production systems requiring combinations of agronomic and breeding interventions.

Dynamic crop growth and development models (CGMs) have the capacity to explore consequences of potential agronomic and breeding interventions in

*Corresponding authors' e-mail address: g.hammer@uq.edu.au

Citation: Hammer G, Messina C, Wu A, Cooper M. 2019. Biological reality and parsimony in crop models—why we need both in crop improvement! *In Silico Plants* 2019: diz010; doi: 10.1093/insilicoplants/diz010

design of crops for production systems (Hammer *et al.* 2014, 2016*a*; Chenu *et al.* 2017). Using simplified mathematical representations of the interacting biological and environmental components of the dynamic soil–plant–environment system, they can reliably predict trajectories of crop attributes through the crop life cycle (e.g. Soufizadeh *et al.* 2018). Environmental (E), genetic (G), and management (M) influences, and hence their dynamic interactions (G×E×M), can be incorporated via the nature and coefficients of the response and control equations in the model and aspects of its initialization. Messina *et al.* (2009) and Cooper *et al.* (2014*b*) highlight the potential role of crop models in crop improvement, particularly the G aspects via plant breeding, which will be our focus here.

Plant breeding provides the genetic foundation for achieving realized crop improvement for a defined target set of environments, commonly referred to as the target population of environments (TPEs). Plant breeders focus on predicting differences among genotypes for trait phenotypic performance to support selection of superior entries. While the details of individual breeding programs differ among crop species, and they will depend on the resources available to the plant breeder, there are common elements that describe the core structure of a breeding program (Fig. 1). Each cycle involves:

(1) testing of individuals sampled or selected from the reference population of genotypes for their trait values in a relevant sample of environments, referred to as multi-environment trials (METs),
(2) selection of individuals, at a defined selection pressure, based on their predicted superior phenotype from analyses of METs and
(3) controlled mating of the selected individuals to produce a new generation of individuals.

The progeny obtained from the mating of the selected individuals form the base reference population of genotypes for the next cycle. The genotypes selected to have superior trait performance are retained and advanced through the stages of the breeding program with some ultimately released for commercial use.

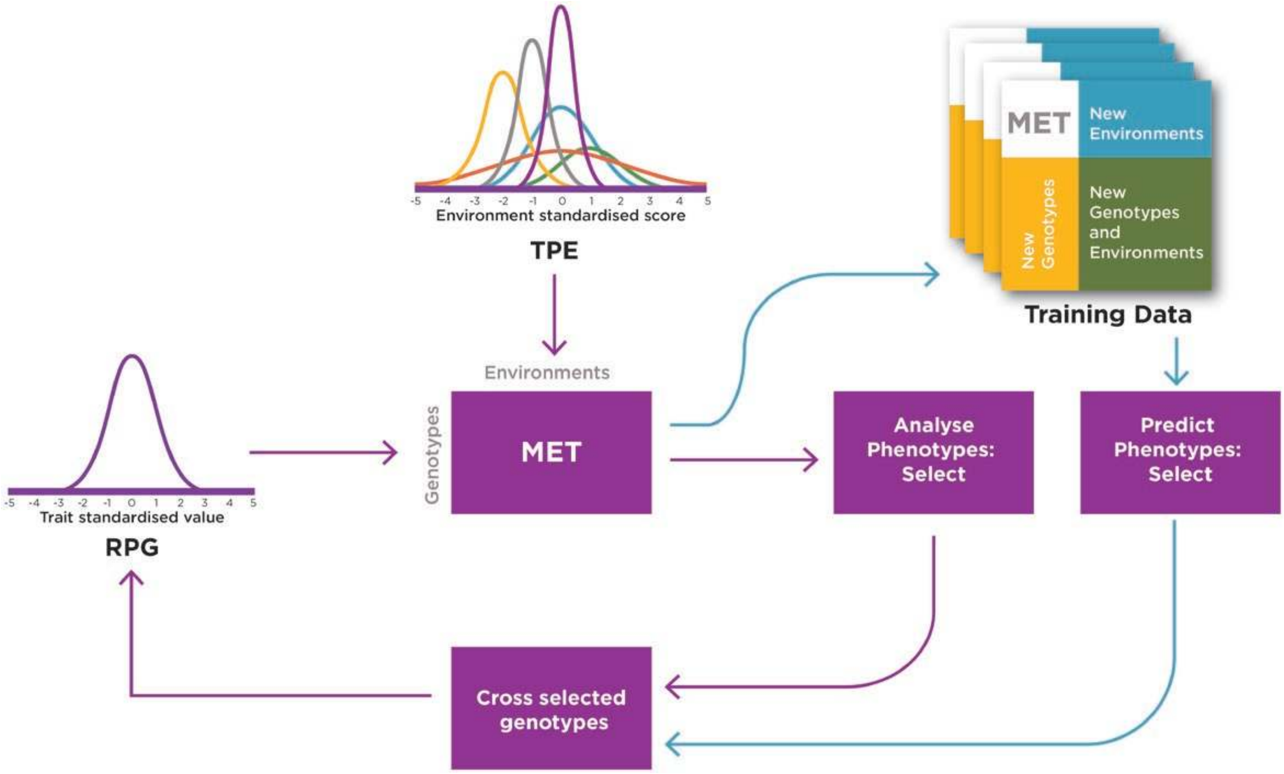


Figure 1. Schematic of the common elements of a breeding program cycle: Individuals are sampled from a reference population of genotypes (RPGs) for testing within a sample of environments that represents the Target Population of Environments (TPE)-the Multi-Environment Trial (MET). Trait phenotypes are measured on the individuals within the MET and the trait phenotypes are analysed. If Single Nucleotide Polymorphisms (SNPs) have been assayed for the genotypes, the trait phenotypes from the MET can be used as a training data set to enable genomic prediction for all individuals in the RPG that have a SNP genome fingerprint. Based on the results of the analyses, the individuals are sorted into either a select or a reject group. The selected individuals are retained and used in a planned crossing scheme to create progeny. In this way the alleles of the genes that determine the trait phenotypes that are possessed by the selected individuals are passed on to the progeny of the next generation. The progeny become the individuals of the next cycle of the RPG.

Applying the conventions of quantitative genetics the relationship between the phenotype (*p*) and genotype (*g*) of an individual *i* in environment (*e*) *j* is represented by the linear model:

$$p_{ijk} = e_j + g_i + (ge)_{ij} + \varepsilon_{ijk} \qquad (1)$$

where p_{ijk} is the measured trait phenotype for observation *k* on genotype *i* in environment *j*, e_j is the effect of environment *j*, g_i is the effect of genotype *i*, $(ge)_{ij}$ is the genotype-by-environment interaction effect for genotype *i* in environment *j*, and ε_{ijk} is the residual (error) effect associated with observation *k* on genotype *i* in environment *j*. The plant breeder seeks to create new genotypes within the breeding program, test them for trait phenotypic performance, and identify and select those with superior trait performance relative to the commercial genotypes already used by farmers. Testing, conducted to measure trait phenotypes for the genotypes, is combined with selection methodology to sort the genotypes into a group predicted to have superior phenotypes.

The genetic principles on which crop breeding programs are designed and their operational details have been well documented (e.g. Fehr 1987; Hallauer *et al.* 1988; Comstock 1996; Allard 1999; Bernardo 2002). For quantitative traits controlled by many genes, it is assumed that genetic improvement of the traits will be achieved by selection operating over multiple cycles of the breeding program. Selection theory provides a quantitative framework for genetic improvement of traits by breeding as represented in Fig. 1. Within each cycle, the alleles of the genes that determine the superior phenotypes of the selected individuals are passed to their progeny to form the reference population of genotypes for the next cycle. In this way, response to selection is associated with the increase in frequency of the favourable alleles over cycles of breeding. Selection theory based on these principles has had a major influence on the design of breeding programs and the key elements are represented in the 'breeder's equation', which in its most basic form can be written as

$$\Delta G_p = ih^2\sigma_p/t \qquad (2)$$

where ΔG_p is the measure of genetic gain from a cycle of the breeding program, *i* is a measure of the selection pressure applied to identify the superior individuals from among those tested, h^2 is the heritability of the target trait, which is defined as the ratio of the genotypic to phenotypic variance from analysis of METs (σ^2_g/σ^2_p), σ_p is the standard deviation of the trait phenotypes in the reference population and *t* is the time it takes to complete one cycle of the breeding program. Details and refinements for alternative breeding program designs and more advanced treatments of the basic breeder's equation can be found in Falconer and Mackay (1996), Comstock (1996), Lynch and Walsh (1998), Holland *et al.* (2003) and Walsh and Lynch (2018).

Recently, the availability of low-cost genome sequencing technologies that have enabled high-throughput measurement of DNA sequence polymorphisms among the individuals created within a breeding reference population has opened up new opportunities to design and upscale breeding programs (Heffner *et al.* 2009; Ramstein *et al.* 2019). The phenotypic data collected in METs can be used in combination with the single nucleotide polymorphism (SNP) sequence data for the genotypes tested in the METs to develop a whole-genome sequence-based prediction model based on the SNPs to predict the trait phenotypic variation (Meuwissen *et al.* 2001). Once the model is established in an appropriate training data set, the trait phenotypes can then be predicted for individuals in the reference population that have been genotyped with an appropriate set of SNPs without directly measuring the trait phenotypes of those individuals (Fig. 1). For such prediction-based breeding strategies, the Breeder's equation (2) can be redefined in terms of the prediction accuracy that can be achieved for the target trait phenotypes based on the SNPs:

$$\Delta G_A = ir_A\sigma_A/t \qquad (3)$$

where ΔG_A is genetic gain in terms of breeding value, *i* is a measure of selection pressure as for equation (2), r_A is the predictive accuracy, defined as the correlation between the estimated breeding value in the training data set and their corresponding true values in the TPE, σ_A is the additive genetic standard deviation of the target trait for the selection unit in the reference population and *t* is the time to complete one cycle of the breeding program.

Motivations for applying genomic selection include increasing the scale of the breeding program when genotyping of individuals is cost effective relative to phenotyping, decreasing the time to complete a cycle of the breeding program, and enabling the prediction of trait phenotypic performance for new genotypes or new environments expected in the TPE. Following the development of genomic prediction methodology (Meuwissen *et al.* 2001) and its successful application to animal breeding (Garcia-Ruiz *et al.* 2016; Garner *et al.* 2016), successful applications to crop breeding programs have followed (Cooper *et al.* 2014*a*, *b*).

It can be argued that there is little attention given to biological reality in the models of trait genetic architecture that are the foundation of quantitative genetics and genomic selection in plant breeding. The situation is similar for more recent concepts on design of plant varieties based on use of DNA sequence data with high-throughput phenotyping images and machine learning algorithms (Ramstein *et al.* 2019). Nevertheless, the basic and more recent models of quantitative traits have and continue to work for predicting the effects of natural and artificial selection and generating genetic gain. However, it remains that context dependencies limit the predictive power of these quantitative statistical methods. For example, the estimate of phenotypic variance for differences among genotypes (equation 1)] is confounded by genotype-by-environment (*G*×*E*) interaction effects. Voss-Fels *et al.* (2019) reviewed the application of genomic selection methodology for plant breeding and considered some of the challenges. A key challenge was developing adequate levels of prediction accuracy for complex traits where strong non-additive effects restricted the extent of associations between the SNPs and the trait phenotypes in the training data sets. This limited the broader application of the genomic prediction models to the genotypes and TPE beyond the training data set. These situations can occur in the presence of strong non-additivity associated with G×E×M interactions.

It has been suggested (Chapman *et al.* 2003; Hammer *et al.* 2006; Cooper *et al.* 2014*b*; Messina *et al.* 2018) that introducing biological insight via using dynamic CGMs to bridge the genotype-to-phenotype predictability gap offers an avenue to address such context dependencies and generate significant advance in possible rates of genetic gain. Technow *et al.* (2015) demonstrated how a process-based crop growth model could be used in combination with genomic prediction to improve the prediction accuracy achieved in the presence of non-additivity associated with G×E×M interactions. Cooper *et al.* (2016) applied this methodology to enable prediction for grain yield of maize in drought environments and Messina *et al.* (2018) demonstrated extensions of the methodology to the challenge of predicting new genotypes in new environments in the presence of trait-by-trait interactions and G×E interactions influencing yield of maize in drought and favourable environments.

While these advances are encouraging, there remains a range of views on the overall approach and on the nature and level of complexity of the CGM needed to achieve the predictive capability required to support crop improvement (Tardieu 2003; Hammer *et al.* 2004, 2006, 2010; Yin and Struik 2010; Marshall-Colon *et al.* 2017). The community more focussed on guiding crop improvement through plant engineering from gene network and metabolic pathway level are seeking plant models based on mechanistic details at lower levels of biological organization (Marshall-Colon *et al.* 2017). Although framed in a different manner, it is not clear if that perspective on the development of plant/crop modelling is incompatible with that arising from the quantitative genetics and genetic gain perspective. Both require the prediction of phenotype in the TPE associated with variation at gene level to be effective in crop improvement, but the degree of parsimony and level of biological reality needed are often viewed as conflicting—but are they? Hence, the objective of this paper is to consider and discuss the level of detail in plant/crop modelling needed to enhance the next generation of crop improvement. To reach this objective, we briefly review approaches to phenotypic prediction, the role of plant/crop scale models in crop improvement, integrating CGM with whole-genome prediction (WGP) for enhanced genetic gain and the role of physiological process level plant models in crop design. This underpins a discussion of the implications this raises for developing the next generation of crop models needed to advance crop improvement.

Phenotypic Prediction

Phenotypic prediction remains at the core of genetic improvement of crops, which can be considered as a search strategy on the complex adaptation or fitness landscape formed from the myriad of phenotypic outcomes associated with all possible G×E×M combinations (Cooper and Hammer 1996; Messina *et al.* 2011; Hammer *et al.* 2014). Plant breeders are interested in using genetic information to predict potential new genotypes with superior yield performance, usually employing statistical quantitative genetics approaches to span the levels of biological organization (Fig. 2). Their interests in predicting phenotype focus on predicting differences among genotypes for trait phenotypic performance for selection rather than on the accurate prediction of trait phenotypes for individual genotypes within environments. Equations (1–3) above capture the essence of this approach. It is important to note that the variance among the environments for traits, which is usually large for traits such as grain yield, does not contribute directly to the phenotypic variance for comparisons among the individuals or to the heritability of the trait phenotypes. Therefore, gene-to-phenotype (G2P) models that focus on accurate prediction of phenotypes for one or a few genotypes across many environments will not be suitable for the prediction needs of the plant breeder unless they can be extended to predict trait phenotypic differences between genotypes at the scale of the numbers of genotypes created and tested within the breeding program.

Alternatively, with the rapid progress in molecular technologies for genome sequencing and functional genomics, it had been widely expected that a

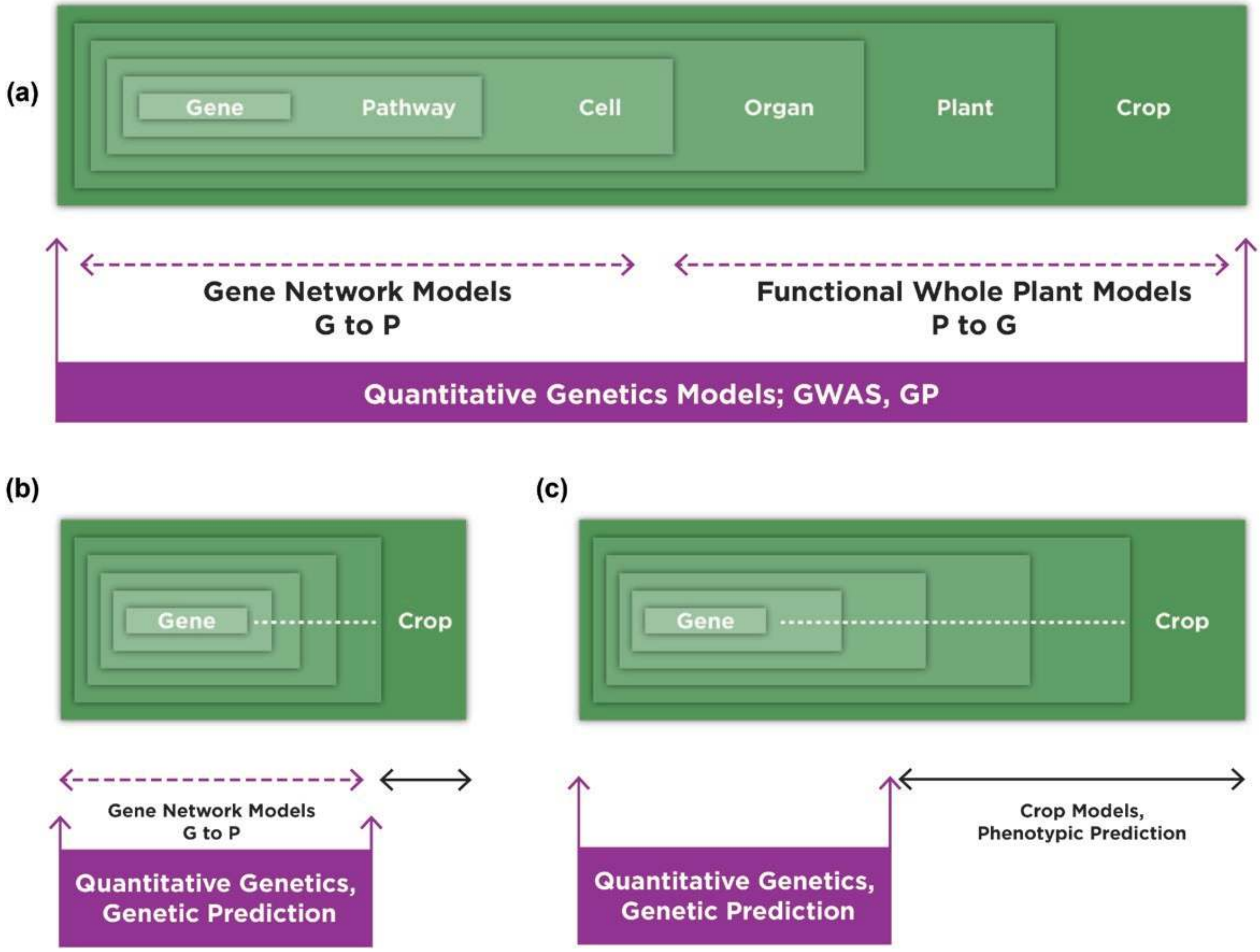


Figure 2. (A) Approaches to G-to-P prediction, their association with levels of biological organization, and the concepts of (B) 'short' phenotypic distance where traits scale directly from molecular to organism scale and there is a likely greater role for genetic prediction and gene network models, and (C) 'long' phenotypic distance where traits do not scale readily from molecular to whole organism level and there is likely a greater role for ecophysiology and crop models (Hammer *et al.* 2016*b*: adapted with permission from Springer Publishing).

gene-by-gene engineering approach would enable enhanced efficiency in crop improvement. The prevailing paradigm was one of developing gene network or metabolic pathway models to span from gene sequence to crop phenotype (Minorsky 2003). There has been considerable development of component process models with the advancing knowledge in plant science, such as for predicting transition to flowering (Salazar *et al.* 2009), sucrose synthesis from photosynthesis (Zhu *et al.* 2013) and the consequences of some pathway manipulations. Indeed, at crop scale, there have been notable successes in developing plants that better resist pests or tolerate herbicides in field production environments (Edgerton 2009). Those cases involved single-gene transformations where plant phenotypic response scaled directly from the level of molecular action, making the response highly predictable. That is, there was a close connection of gene expression to phenotypic expression at plant/crop scale or a short 'phenotypic distance' (Fig. 2B). There have been other examples where the engineering approach has generated promising leads, such as with modifying the photosynthetic pathway (Salesse-Smith *et al.* 2018; South *et al.* 2019), although claims of likely impact on field crop yield are often overstated. A recent detailed cross-scale modelling analysis of photosynthetic manipulations suggested significant, but far more modest, likely impacts on crop yield (Wu *et al.* 2019) due to confounding effects of environmental and crop scale interactions (Wu *et al.* 2019). Nonetheless, with the advances in gene editing technology (Doudna and Charpentier 2014), notions of plant design by gene and metabolic network manipulation provide renewed impetus for constructing models from the gene upward (Marshall-Colon *et al.* 2017).

It remains, however, that little of the promise of the engineering approach has been realized in practice for key complex traits where relationships among components and their genetic controls involve quantitative multi-gene interactions. There is a long 'phenotypic distance' due to the extent of the biological integration required from the causal polymorphisms at genome scale to the phenotype of interest at plant and crop scale (e.g. Sinclair *et al.* 2004) (Fig. 2C). Integrating gene effects across scales of biological organization for phenotypic prediction in such situations is not straightforward (Hammer *et al.* 2006). The predictive power of associations of genes/genomic regions with phenotypic responses for complex traits remains constrained by context-dependent gene/trait effects attributed to gene–gene and gene–environment interactions associated with the underpinning gene networks (Cooper *et al.* 2005).

As noted earlier, the same issue restricts the extent of advance in plant improvement via approaches based on quantitative genetics and genomic prediction technologies (Voss-Fels *et al.* 2019). Continuing advances in genotyping and crop genomics (Heffner *et al.* 2009; Morrell *et al.* 2011; Morris *et al.* 2013) have now facilitated association mapping approaches that assess correlation of phenotype with genotype in populations or panels of unrelated individuals. Such genome-wide association studies rely on advanced statistical procedures to identify associations between a phenotype and a genomic marker profile. Genomic selection involves the use of phenotypic prediction equations based on profiles of marker data from a training set of genotypes, which have been phenotyped. The predictions are then applied across breeding materials that are genotyped extensively but not phenotyped. This offers considerable potential for more rapid genetic gain in breeding. However, for complex traits, the procedure still suffers from context-dependent effects and the 'phenotypic distance' problem (Fig. 2). Association mapping and genomic selection rely on the stability of the relationship between a phenotype and the set of genomic markers found in the training set, which is strongly dependent on the relevance of the genotypes and environments sampled.

Biological and ecophysiological insight contained in dynamic crop models provides a means to alleviate context dependencies associated with the long phenotypic distance issue that limits prediction in situations where traits do not scale readily from molecular to whole organism level (Fig. 2). The crop model captures some of the trait–trait and trait–environment effects associated with dynamic interactions of simpler component traits and the growing environment. The complex trait becomes an emergent consequence of these interactions among component traits with shorter phenotypic distance that are likely more highly heritable. For example, van Oosterom *et al.* (2011) present a framework for crop scale adaptation to post-anthesis drought stress in sorghum based on consequential effects of traits that influence water use and availability through the crop life cycle, such as maturity, canopy development, tillering, root architecture and transpiration efficiency (TE). They demonstrate how variation in these traits can influence the dynamics of the crop water balance and ultimately the rate of post-anthesis leaf senescence or 'stay-green' under terminal stress conditions. Stay-green is a complex trait known for its positive connection to adaptation to water-limited environments in sorghum breeding (Jordan *et al.* 2012). Borrell *et al.* (2014*a*) utilize this water balance framework to dissect the stay-green trait and show how positive effects of stay-green quantitative trait loci (QTLs) on grain yield under drought were explained as emergent consequences of their effects on temporal and spatial water use patterns that result from changes in these component traits. In essence, the biological insight of the trait and environment interactions inherent in the dynamic framework unpacks the context dependencies related to expression of the complex stay-green phenotype.

Biological insight contained in more detailed gene network/metabolic pathway models provides a means to target specific gene interventions (or inform genomic prediction) in cases of short phenotypic distance where traits scale more directly from gene expression to organism phenotypic expression with less effect of context dependencies (Fig. 2B). However, prediction at this level requires detailed knowledge of regulatory mechanisms operating on expression dynamics of gene networks. For example, the transition to flowering is known to be regulated by expression of FLOWERING LOCUS T (FT) (Corbesier *et al.* 2007). However, FT expression is linked to gene circuits regulated by the circadian clock, photoperiod and temperature (Salazar *et al.* 2009). The dynamics of these networks have been studied in detail in *Arabidopsis* and synthesized in a mathematical model (Seaton *et al.* 2015) that enables simulation of network behaviour and prediction of flowering for different environments and genotypes (Chew *et al.* 2017). This provides an avenue to quantify how specific gene manipulations might influence timing of floral induction (Dong *et al.* 2012).

The Role of Plant/Crop Scale Models in Crop Improvement

CGMs that are structured to capture the dynamic interactions of the physiological determinants of crop growth and development (Fig. 3) can be employed to analyse and aid understanding of the physiological

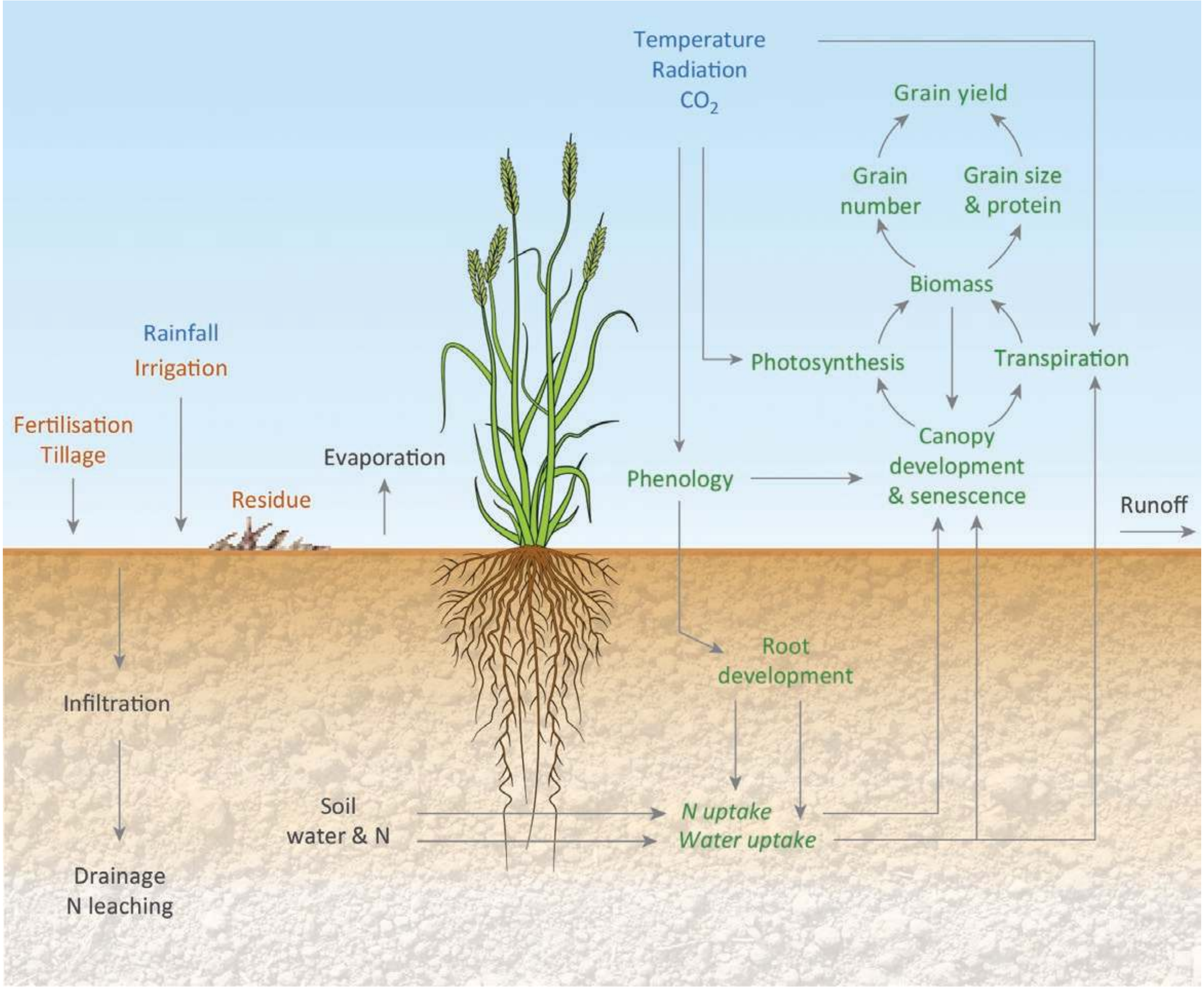


Figure 3. Crop model schematic showing the key processes involved in crop growth and development and their interactions with the crop system (Chenu *et al.* 2017: with permission from Elsevier).

and genetic basis of trait variation (Messina *et al.* 2009; Hammer *et al.* 2016*a*). Crop performance, and hence the trait phenotypes we measure and map, is an emergent consequence of the dynamic interplay of crop growth and development processes within an environmental context over the crop life cycle. CGMs can thus both inform phenotyping for QTL detection and predict consequences of G×E×M interactions to assist crop improvement in general and molecular breeding in particular (Hammer *et al.* 2005; Cooper *et al.* 2009; Voss-Fels *et al.* 2019). The value of the framework to assist plant breeding depends on the extent to which the algorithms included in the CGM adequately capture the physiological determinants of genetic variation for adaptive traits of interest to the breeder (Tardieu 2003; Cooper *et al.* 2009; Messina *et al.* 2009; Tardieu and Tuberosa 2010) and alleviate context dependencies. Chenu *et al.* (2009) integrated organ-level QTL effects on leaf expansion to form a G2P CGM capable of simulating impact of the QTLs at the whole plant level and predicting their importance in terms of yield in various environments. The leaf submodel developed (Chenu *et al.* 2008) was based on environmental responses of leaf elongation rate that were characterized by stable QTL (Reymond *et al.* 2003, 2004). This approach avoided QTL–environment interactions often observed for more integrated traits such as whole plant leaf area production, which was empirically related directly to environmental conditions in early crop models (Sinclair 1986; Hammer *et al.* 1993). Hence, fundamental plant/crop physiological and genetic studies are often necessary to improve CGM architecture for G2P studies (Tardieu

2003; Messina *et al.* 2006; Chenu *et al.* 2008; Hammer *et al.* 2009; Messina *et al.* 2009; Bertin *et al.* 2010; Yin and Struik 2010). The issue of balancing biological reality and parsimony in the level of detail needed for CGMs to be able to integrate processes across levels of organization while predicting emergent functional consequences for the organism is the focus of later discussion in this paper.

Two examples are presented to outline the role of plant/crop scale models in crop improvement:

(i)Modelling the physiology and genetics underpinning the stay-green trait in sorghum

As outlined earlier, stay-green is a complex trait known for its positive connection to adaptation to water-limited environments in sorghum breeding (Jordan *et al.* 2012). Sorghum hybrids that 'stay-green' for longer during grain-filling tend to have superior yield under terminal drought conditions. While this trait has been known to breeders for many years (Rosenow 1977; Rosenow *et al.* 1983; Henzell *et al.* 1992), and phenotyped by them in METs, the underpinning genetics is not simple and the physiological mechanisms not well understood. Hence, while phenotypic selection for the trait has proceeded and resulted in improved yield in water-limited environments (Jordan *et al.* 2012), enhanced understanding of the physiological and genetic underpinning of the trait offers the potential to advance rates of genetic gain via molecular breeding and/or genomic prediction. Researchers have mapped the QTLs that contribute to stay-green in a range of populations (e.g. Xu *et al.* 2000), many of which were derived from crosses with BTx642, a derivative of an Ethiopian durra landrace in which the trait was discovered.

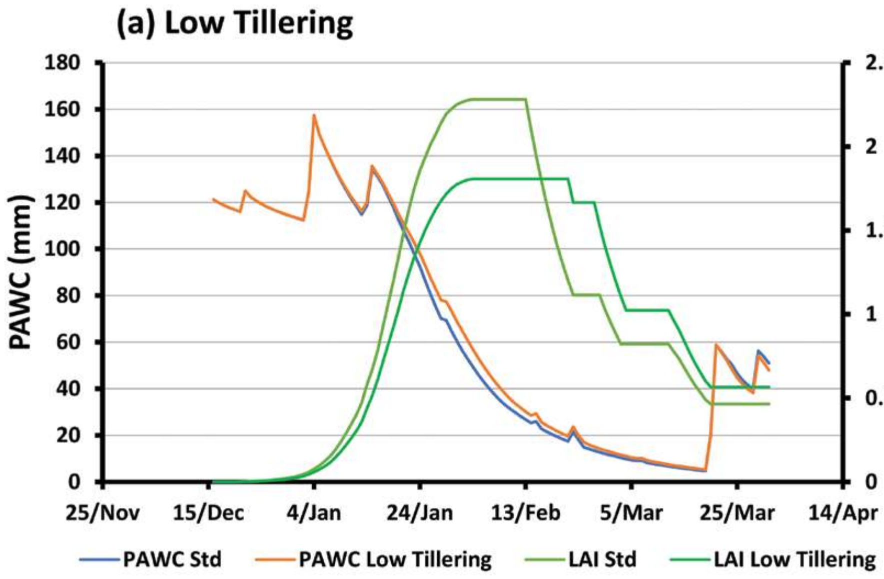
The ecophysiological framework set out by van Oosterom *et al.* (2011) suggests that traits moderating the water balance through the crop cycle so it is more favourable during grain-filling could underpin the presence of stay-green in terminal drought environments. There are a number of traits that could generate stay-green by influencing the temporal dynamics of the crop water balance. If soil water is limited, water availability at anthesis can be increased by either increasing water accessibility to the crop or restricting pre-anthesis water use. Accessibility of water can be manipulated through root architecture. Studies in sorghum have identified variation in seedling nodal root angle as an indicator of capacity for the root system to be better distributed at depth and increase the amount of water extractable from deep soil layers via improved soil occupancy (Singh *et al.* 2011, 2012) as also found in wheat (Manschadi *et al.* 2006). Mace *et al.* (2012) found that the QTL associated with genetic variation for nodal root angle in sorghum co-located with known QTL for stay-green.

The amount of crop water use pre-anthesis can be reduced through early anthesis or restriction of canopy size, which can be manipulated via genetic regulation of leaf size and tillering. Turner (2004) and Hammer *et al.* (2006) highlight the concept of 'shifting water' from pre- to post-anthesis by regulating potential canopy size. Tiller production depends on internal plant competition for assimilates (Bos and Neuteboom 1998; Lafarge and Hammer 2002; Kim *et al.* 2010*b*) and genotypic differences in tillering of sorghum have been associated with differences in leaf appearance and width, which affected vigour of the main shoot and restrict tillering (Kim *et al.* 2010*a*; van Oosterom *et al.* 2011). Alam *et al.* (2014*a*) presented a quantitative framework that explained the genetic and environmental regulation of tillering in sorghum breeding populations by incorporating a measure of surplus assimilate status of the plant and a genetic propensity to tiller. A subsequent QTL analysis using this framework (Alam *et al.* 2014*b*) on data from multiple sorghum populations identified QTL associated with these two main driving factors. QTL for propensity to tiller co-located with genes involved with hormonal control of tiller bud outgrowth, such as the biosynthesis of strigolactones (Beveridge and Kyozuka 2010). Other QTL co-located with factors affecting plant assimilate status, which is consistent with the hypothesis that availability of assimilate beyond the requirement of existing culms regulates tillering (Bos and Neuteboom 1998; Lafarge and Hammer 2002; Kim *et al.* 2010*a*, *b*). This concept at whole plant scale pre-empts and supports the more recent findings from molecular studies on the important role of sugar signalling in bud release (Mason *et al.* 2014).

Pre-anthesis crop water use can also be reduced by limited maximum transpiration, whereby stomatal conductance is reduced under conditions of high evaporative demand and, hence, plant TE is increased. Genetic variation in this trait has been reported in sorghum (Gholipoor *et al.* 2010) and other species (Sinclair *et al.* 2010). Variation in plant hydraulic conductivity associated with the role of aquaporins in water transport has been identified as a likely underpinning mechanism (Vadez *et al.* 2014). Preliminary simulation studies in sorghum (Sinclair *et al.* 2004) have indicated the potential value of this trait for crop adaptation in water-limited situations and this has been reinforced with comprehensive studies for maize in the USA (Messina *et al.* 2015).

Borrell *et al.* (2014*b*) utilize this water balance framework to dissect the stay-green trait and show how positive effects of stay-green QTLs on grain yield under drought were explained as emergent consequences of their effects on temporal and spatial water use patterns that result from changes in these component traits. In essence, the biological insight of the trait and environment interactions inherent in the dynamic framework unpacks the context dependencies related to expression of the complex stay-green phenotype.

Hammer *et al.* (2016*a*) reported a simulation study with the APSIM-sorghum model in Australia that quantified consequences of varying two component traits affecting stay-green—tillering and TE. Kholová *et al.* (2014) reported a similar study using APSIM-sorghum to model the effect of plant water use traits on yield and stay-green expression in post-rainy season sorghum in India. In the Australian study, the tillering routine was adapted to generate a reduced tillering type when compared with a standard hybrid, and a limited maximum transpiration rate was introduced to restrict water use in the middle of the day. The latter required invoking an hourly calculation and used the routine developed by Hammer *et al.* (2009), with transpiration limits imposed as set out by Sinclair *et al.* (2004). The simulated results for a single year (Fig. 4), chosen because of its terminal


(a) Low Tillering

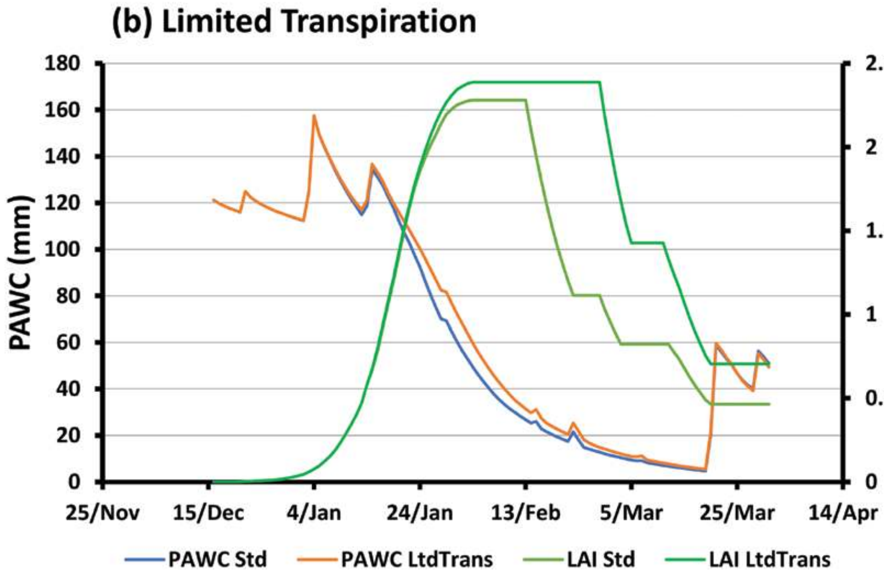

(b) Limited Transpiration

Figure 4. Simulated plant available soil water content (PAWC) and crop LAI through the crop life cycle for a terminal drought year from a sorghum crop simulation at Emerald (NE Australia) for a standard hybrid (blue and orange lines) and a contrasting hybrid (red and green lines) with either (A) reduced tillering or (B) a limited maximum transpiration rate (Hammer *et al.* 2016*a*).

moisture stress pattern, indicated the capacity of the model to generate stay-green as an emergent consequence in this type of situation. The leaf area index (LAI) associated with reduced tillering type or limited maximum transpiration remained high for longer into the post-anthesis period of the crop cycle. This was due to reduced canopy size (LAI) causing reduced pre-anthesis water use in the case of reduced tillering, or enhanced TE causing reduced pre-anthesis water use in the case of limited maximum transpiration, despite similar canopy size. In both of these situations, there was increased simulated yield associated with the modified type (data not shown). In the case of reduced tillering, this was not linked to any increase in total biomass at maturity, but rather to increased harvest index as a consequence of the increased water availability and crop growth post-anthesis. In the case of limited maximum transpiration rate, both total and grain biomass were simulated to increase as a result of the enhanced TE this effect generated. When viewed over the long-term climate for the location used in the study, there was considerable yield advantage to both reduced tillering and limited maximum transpiration for seasons where yield of the standard hybrid was below about 4.5 t ha^{-1} (Hammer *et al.* 2016*a*). This is consistent with observations from breeding trials for the stay-green trait (Jordan *et al.* 2012) and reflects the majority of farmer field situations in NE Australia, where the long-term average yield level is around 3 t ha^{-1} (Potgieter *et al.* 2016). However, there was yield reduction in high-yielding seasons associated with the conservative growth associated with these adaptations in seasons where water is not limiting.

This result highlights the value of these component traits and suggests phenotyping strategies to target them directly could be advantageous. This exemplifies the concept of dissecting a complex trait to component targets that might be more tightly linked to genetic control (higher heritability) and thus offer potential to enhance genetic gain if targeted and effective phenotyping and analysis systems are invoked (e.g. Chenu *et al.* 2018; van Eeuwijk *et al.* 2019).

(ii)Developing drought-tolerant maize hybrids for the US Corn-Belt

The development of the AQUAmax drought-tolerant maize hybrids (Cooper *et al.* 2014*a*) was built on the legacy of long-term genetic improvement of maize by breeding (Duvick *et al.* 2004; Smith *et al.* 2014). There is a long history of yield improvement of maize for the diverse range of environments of the US Corn-Belt and the G×E×M interactions that have emerged as crop management practices have evolved; e.g. the transition from manual to mechanized harvesting, the long-term increases in plant population that followed mechanized planting and harvesting and genetic improvement for standability. The long-term commercial breeding for yield of maize had resulted in genetic improvement of maize yield potential, drought tolerance, disease and pest resistance and agronomics for the environments and management strategies of US farming systems. Therefore, in context with the breeding program schematic (Fig. 1), multiple cycles of breeding for increased yield in the TPE had been completed (Duvick *et al.* 2004) prior to the commencement of the recent focused efforts to improve the drought tolerance of maize hybrids (Campos *et al.* 2004; Cooper *et al.* 2014*a*). Thus, the challenge was to further improve on the levels of drought tolerance that had already been achieved without penalizing the yield potential of maize in the favourable high-input environments of the US Corn-Belt. Based on a series of experimental investigations into the historical improvements that had been achieved through breeding, multiple hypotheses were proposed for potential pathways to further improve levels of drought tolerance (Campos *et al.* 2004; Barker *et al.* 2005).

Process-based CGMs, and the crop ecophysiological principles on which they are based, in combination with experimental investigations made important contributions to the testing of the hypotheses for pathways to enhanced levels of drought tolerance (e.g. Campos *et al.* 2004, 2006; Hammer *et al.* 2009; Messina *et al.* 2011, 2015; Reyes *et al.* 2015; van Oosterom *et al.* 2016). A key result that emerged from these studies was that for modern elite hybrids, there were important physiological pathways to improved yield under drought conditions that did not require access to and acquisition of greater quantities of total available soil water (e.g. Reyes *et al.* 2015; van Oosterom *et al.* 2016). These results suggested that the rate and patterns of water use during crop growth and development were important components that were contributing to the higher yield of some maize hybrids under important field drought conditions within the TPE. Following these findings, a series of side-by-side field comparisons between drought-tolerant and sensitive hybrids with similar maturity and development patterns revealed different rates and patterns of water use from the total soil profile, with the drought-tolerant hybrids frequently extracting water from the soil profile at a slightly lower rate, resulting in greater measured total available soil water for the tolerant hybrids during the most sensitive developmental stages of flowering and early grain-filling. In environments where these differences were expressed and there were no rainfall events during the flowering period, the higher available soil water for the tolerant hybrids translated into a higher grain yield (Cooper *et al.*

2014*a*). Predictably in drought-prone environments, the yield effect of the differences in available soil water between the tolerant and sensitive hybrids was greater with increased plant population and conversely could be reduced by decreasing plant population (Cooper *et al.* 2014*a*). Decreasing plant population to conserve soil water is a common agronomic practice in the more drought-prone environments of the Western region of the US Corn-Belt (Lyon *et al.* 2003; Grassini *et al.* 2009).

Following the identification of improved yield in drought-prone environments that was associated with reduced rates of water use, investigations were undertaken to identify mechanistic bases. A plant-level mechanism identified to be important was the presence of limited maximum transpiration under high atmospheric vapour pressure deficit (VPD) (Yang *et al.* 2012; Gholipoor *et al.* 2013; Choudhary *et al.* 2014; Shekoofa *et al.* 2015), with a potential role of aquaporins (Choudhary *et al.* 2015). The identification of a putative mechanistic basis for differences in hybrid yield under drought that was associated with the observed differences in water conservation provided direction for extensions to a process-based CGM for maize and modelling of the yield impact of trait variation for the limited maximum transpiration trait within the context of the TPE of the US Corn-Belt (Messina *et al.* 2015). While the limited-transpiration trait was identified as one mechanism contributing to the yield differences associated with water conservation it was not the only mechanism considered, others included canopy size, which was itself associated with differences in leaf number and leaf size (Messina *et al.* 2018), and root system efficiency (van Oosterom *et al.* 2016).

Given the potential involvement of multiple trait mechanisms contributing to the observed hybrid variation for patterns of water use and yield in drought-prone environments, it was necessary to design a breeding strategy that allowed for the recognized and as of yet unknown multiple workable solutions (Messina *et al.* 2011) to achieve further improvements for drought tolerance without compromising yield potential. A core component of the breeding strategy was to augment the METs conducted in the TPE (e.g. Fig. 1) with specifically designed field-based, managed-environments that allowed high-throughput precise phenotyping for yield under relevant drought conditions (Cooper *et al.* 2014*a*). When correctly designed, the managed-environments allowed consistent identification, phenotyping and modelling of multiple trait combinations that contributed to water conservation and yield variation within all stages of a breeding program. Further, complementary managed-environments that revealed yield variation for trait combinations associated with enhanced access to deep soil moisture were also established (Cooper *et al.* 2014*a*). The resulting MET data sets from the combined phenotyping in the managed-environments and the TPE provided the necessary data to establish suitable training data sets to apply genomic prediction to accelerate breeding for enhanced drought tolerance with no reductions in yield potential.

Industry-scale comparisons between AQUAmax and non-AQUAmax hybrids within the US Corn-Belt over thousands of environment–management combinations over a period of 6 years (2008–13) demonstrated the higher yield of the tolerant hybrids in comparison to a large number of different non-AQUAmax hybrids used throughout the USA (Gaffney *et al.* 2015). Other studies based on smaller numbers of hybrid comparisons and small samples of environments and limited sets of management practices (e.g. 2–10 hybrids in 1–10 environments) have either demonstrated advantages of the drought-tolerant hybrids (Hao *et al.* 2015*a*, *b*, 2016; Mounce *et al.* 2016) or did not identify the differences (Roth *et al.* 2013) that were observed when larger numbers of comparisons were included in the studies.

Epilogue. The two examples detailed indicate that crop design supported and/or enabled by a CGM is able to deal effectively with complex traits and does not require adoption of a classical ideotype breeding approach to direct the breeding outcomes. If the CGM is used to identify a functional target within the physiological framework of resource capture × resource use efficiency × partitioning, then there may be, and it is likely that there are, multiple workable solutions to achieve the functional target. To clarify this point, consider an example. If the functional target is to make an extra 25 mm of water available during reproductive development and/or grain-filling then as noted in both case studies above this can be achieved through multiple pathways. Thus, any one ideotype could represent one of the multiple workable solutions. Thus, alternative workable solutions (ideotypes) could be ranked on suitable criteria and for the purposes of maintaining genetic and physiological diversity in the germplasm pool of the breeding program. To sustain long-term gain, we can prefer to design a breeding strategy that focuses on selecting multiple workable solutions for the functional target. This interpretation is consistent with both the sorghum stay-green and maize AQUAmax cases. For the AQUAmax case, specific steps were (i) identifying the functional target (an extra 25 mm of water post-flowering), (ii) achieving this target through different trait combinations and (iii) selection conducted for yield in managed environments where drought was managed to reveal yield differences that were associated with that target through any of many trait combinations that were possible given the trait diversity used in the breeding program. Some trait combinations were likely more reliable in achieving the 25 mm water shifting target and it seems that limited maximum transpiration was likely a frequent component, but reduced plant size was also exploited to some extent. The small plant stature pathway of conserving water may have been more risky overall in that it may have limited yield more severely in the favourable environments, as found for reduced tillering in sorghum. There are likely other pathways we still have to discover, such as root system efficiency (van Oosterom *et al.* 2016).

Integrating CGMs With WGP for Enhanced Genetic Gain

Recent developments (Technow *et al.* 2015; Cooper *et al.* 2016; Messina *et al.* 2018) have suggested a pathway to progress beyond the examples above, which have a direct focus on specific target traits, modelling and prediction of their effects, and subsequent phenotyping and genetic analysis to link with crop improvement. Technow *et al.* (2015) merged the biophysical capability of a CGM with an approach to WGP using a Bayesian framework (Meuwissen *et al.* 2001). The resulting CGM-WGP methodology provided a one-step approach to enhancing predictive power and capacity to deal with G×E context dependencies in breeding systems. Cooper *et al.* (2016) applied this methodology to enable prediction for grain yield of maize in drought environments and Messina *et al.* (2018) demonstrated extensions of the methodology to the challenge of predicting new genotypes in new environments in the presence of trait-by-trait interactions and G×E interactions influencing yield of maize in drought and favourable environments.

This CGM-WGP one-step approach incorporates biological insight directly within genomic prediction algorithms, in this case a Bayesian methodology, to train models using data from multiple environments. Messina *et al.* (2018) used virtual and experimental data from a maize drought breeding program to evaluate whether the CGM-WGP methodology enabled improved phenotypic prediction when G×E interactions were an important determinant of performance. In the simulation study, they generated a virtual population of 1,000 double haploid (DH) lines with dense genotyping and yield tested in two water-limited (WL) and one non-water-limited (NWL) environment. Table 1 shows the reported correlations of predicted with observed (simulated) in each prediction environment for both CGM-WGP and the more conventional BayesA methods given various sets of environments used for training the models. The highlighted contrasts indicate the enhanced prediction capacity of the CGM-WGP approach in environments outside the training set. For example, the CGM-WGP approach is able to maintain high correlation for the NWL environment when trained only on the two WL environments. This was not the case for the BayesA method. The latter relies on an additive linear model of genetic effects, which cannot capture this G×E effect. In contrast, the CGM-WGP approach, by connecting the genetic profiles via key model coefficients, is able to capture the G×E effects as they emerge from the non-linear dynamics of the CGM. In this study, the CGM was optimized for four key factors that influence both potential yield and stress responsiveness in maize—maximum leaf (and hence canopy) size, radiation use efficiency (RUE), initial assimilate partitioning to the developing ear and limited maximum transpiration rate (and hence variation in TE). The ability of the CGM-WGP approach to better capture the G×E effects in the simulation study was also reflected in the results of the empirical breeding experiment (Messina *et al.* 2018).

Two recent comprehensive reviews (van Eeuwijk *et al.* 2019; Voss-Fels *et al.* 2019) have explored avenues for accelerating crop genetic gains by integrating advances in genomic selection, modelling and phenotyping

Table 1. Mean prediction accuracy for yield, estimated by the correlation coefficient *r* (±SD), for CGM-WGP prediction methodology and the reference BayesA genomic prediction method for combinations of estimation and prediction environments.

Estimation environment	**Prediction environment**					
	WL1 (1988)	**WL2 (2012)**	**NWL (2010)**	**WL1 (1988)**	**WL2 (2012)**	**NWL (2010)**
		CGM-WGP			**BayesA**	
WL1 + WL2	0.79 ± 0.04	0.75 ± 0.05	0.67 ± 0.08	0.77 ± 0.04	0.51 ± 0.13	−0.50 ± 0.15
WL2 + NWL	0.72 ± 0.09	0.75 ± 0.06	0.81 ± 0.04	−0.39 ± 0.16	0.26 ± 0.25	0.75 ± 0.07
WL1 + WL2 + NWL	0.76 ± 0.05	0.71 ± 0.06	0.77 ± 0.04	0.43 ± 0.13	0.53 ± 0.11	−0.01 ± 0.2

WL denotes water-limited environments and NWL is for the non-water-limited environment (Messina *et al.* 2018: with permission from Elsevier).

techniques. Both identify the significant potential of the approach outlined in the CGM-WGP study above to extend classical approaches and provide opportunities for the redesign of current plant breeding programmes. For example, the integral role of a CGM in the phenotypic prediction framework is identified as an avenue to focus field-based and high-throughput phenotyping effort, enhance characterization of target production environments along with choice of the model training set of environments (Rincent *et al.* 2017) and facilitate design strategies to access trait genetic diversity that would otherwise be difficult to access if the focus is restricted to additive genetic effects.

The Role of Physiological Process Level Plant Models in Crop Design

A finer level of granularity in modelling than found in advanced crop-level growth and development models is needed to predict consequences of manipulation at biochemical/metabolic pathway or more detailed process level. For example, Zhu *et al.* (2007) developed a complete dynamic model of the C3 photosynthetic carbon metabolism pathway to explore whether photosynthetic rate might be increased by altered partitioning of resources among the enzymes of carbon metabolism. By linking the model with an optimizing algorithm, and allowing partitioning between enzymes of a fixed total amount of protein-nitrogen to vary, they identified instances where the manipulation of partitioning could likely greatly increase carbon gain without any increase in the total protein-nitrogen investment in the apparatus for photosynthetic carbon metabolism.

Zhu *et al.* (2013) extended this model and described a detailed model of photosynthetic metabolism that includes each discrete process from light capture to carbohydrate synthesis in the cell. They demonstrated that the e-photosynthesis model effectively mimicked the typical kinetics observed in intact C3 leaves of leaf CO_2 uptake, O_2 evolution and other photosynthetic dynamics following perturbations in light, $[CO_2]$ and $[O_2]$. Song *et al.* (2017) extended this metabolic model to canopy level by linking it to a ray tracing model and aspects of canopy architecture to simulate light penetration into the canopy at any instant. They explored the consequences of manipulating leaf chlorophyll concentration ([Chl]) on canopy photosynthesis (A_c) over a diurnal period and identified putative avenues to increase A_c. The modelling revealed a modest ~3 % increase in A_c when [Chl] was reduced by 60 %. However, if the leaf nitrogen conserved by this decrease in leaf [Chl] were to be optimally allocated to other components of photosynthesis, A_c was increased by over 30 %.

While these process level modelling studies at metabolic pathway, leaf and canopy level suggest opportunities for significant advance in crop-level biomass and yield associated with photosynthetic manipulation, which is often asserted in these studies, cross-scale modelling analyses have identified many interactions that confound this simplistic extrapolation when extended to crop scale (Gu *et al.* 2014; Wu *et al.* 2019). Wu *et al.* (2019) identified effects of photosynthetic manipulation on seasonal patterns of resource (e.g. water and nitrogen) capture and use on plant water and nitrogen status, along with effects on temporal dynamics of crop growth and partitioning to grain, that confounded effects of photosynthetic manipulation on biomass and yield. The consequences were dependent on the nature of the photosynthetic manipulation and the field production environment involved.

The cross-scale model of Wu *et al.* (2019), constructed in the APSIM crop modelling framework, used the partition of the canopy into sunlit and shade leaves based on canopy architecture and solar geometry and quality of the indecent radiation (Hammer and Wright 1994; de Pury and Farquhar 1997). This was combined with the steady state assimilation-internal leaf CO_2 ($A–C_i$) relationship of the functional FvCB photosynthesis model (Farquhar *et al.* 1980), and consideration of diffusion of atmospheric CO_2 through stomatal pores into the intercellular space (C_i) and to the Rubisco carboxylation site inside chloroplasts. Comprehensive tests on a wide range of field crop experiments with this cross-scale model indicated good levels of prediction of crop biomass and yield. This suggests the use of the FvCB model combined with a biophysical model of gas (CO_2 and water vapour) exchange and the simple and robust canopy upscaling approach as an effective avenue to integrate and explore the consequences of photosynthetic manipulation at crop scale (Wu *et al.* 2019). The link provides an avenue for assessing putative changes in photosynthetic pathways via their effects on the $A–C_i$ curve and gas exchange attributes.

In C_3 crops, such as wheat, photorespiration is a contributor to loss of assimilated carbon and has been highlighted as a target for photosynthetic enhancement (Long *et al.* 2006, 2015). Photorespiration occurs when Rubisco oxygenates RuBP and generates by-products that require energy to process with concomitant releases of assimilated carbon. Recent studies with transgenic plants have demonstrated enhancements in leaf photosynthetic CO_2 assimilation (South *et al.* 2019) in transgenic tobacco plants engineered to use an alternative photorespiration pathway. Assuming that the observed net effects on the $A–C_i$ curves of the transgenic plants could be attributed to the alternative photorespiration pathway, it is possible to simulate the anticipated consequences on crop biomass and yield using the cross-scale model of Wu *et al.* (2019).

The A–C_i curve reported for the transgenic tobacco plants by South *et al.* (2019) shows a 10 % decrease in the chloroplastic CO_2 compensation point (Γ*), and a 20 % increase in the maximum quantum efficiency (QE). However, there was also a concomitant 35 % increase in the maximum Rubisco carboxylation rate at 25 °C (V_{cmax25}). It is possible to simulate the consequences of these effects on wheat yield if we assume that the same relative effects were to occur in a current elite crop cultivar. A simulation study was designed to quantify how the three aspects of the modified A–C_i curve, when added sequentially, might affect wheat yield in a standard production environment. Wheat growth and yield simulations were conducted using standard agronomic management practices and the long-term historic (1900–2015) weather data for Dalby, in the Australian wheat belt. This approach allowed sampling of a wide range of environmental conditions (i.e. solar radiation, temperature, rainfall) associated with annual climatic variability at this location. Simulations were conducted for each year of the climate record assuming a quick maturing variety was planted on May 15 every year with a plant density of 150 plants m^{-2} and a row spacing of 40 cm. The soil was a representative 180 cm deep black vertosol as found in the region. Hundred millimetres of plant available water were assumed at planting and initial soil N status and applied fertilizer N were set to simulate non-limiting nitrogen conditions.

Simulated yields with standard photosynthetic attributes for wheat varied from 1 to 7 tonnes per hectare (t/ha) with most years having yield of <4 t/ha (Fig. 5), which reflects the common incidence of water limitation in this production environment. Incorporating the decrease in the chloroplastic CO_2 compensation point (Γ*) associated with reduced photorespiration into the cross-scale model had only a small effect (~1.5 %) on the simulated yield (Fig. 5A). This was consistent with the experimental finding that Γ* alone was not the main contributor to anticipated improved crop performance (South *et al.* 2019). It was also consistent with a theoretical analysis of modified photorespiration using an extended metabolic model based on Zhu *et al.* (2007), which showed only small changes in the A–C_i curve (Xin *et al.* 2015). The further addition of the increase in the maximum QE (+20 %) due to the lower energetic requirement of the photorespiration modified plants (Fig. 5B) resulted in a greater level of improved yield (~4 %), especially at yield levels >4 t/ha. The modelling suggested that this effect enhanced the electron-transport limited rate of leaf CO_2 assimilation and contributed to canopy biomass growth and crop yield via that effect. This was also consistent with the theoretical photosynthetic light response analysis of photorespiratory modification by Xin *et al.* (2015). Adding the experimentally observed increase in V_{cmax25} (+35 %) had a slight further effect on yield (Fig. 5C), which is consistent with our previous findings that wheat canopy photosynthesis and crop biomass growth is predominantly limited by electron transport rate (Wu *et al.* 2018, 2019). Overall, the simulated change in yield differed among the reported effects on the A–C_i curve, ranging from a slight yield penalty to the greatest increase of over 0.5 t/ha improvement in some high-yielding (>4 t/ha) environments when the complete set of A–C_i attributes associated with the photorespiration transgenic lines was incorporated (Fig. 5C).

The simulation study suggests that the degree of yield improvement associated with changes to the A–C_i curve observed from the engineering targeting photorespiration depends on the type of photosynthetic enhancement and the environmental context of the crop.

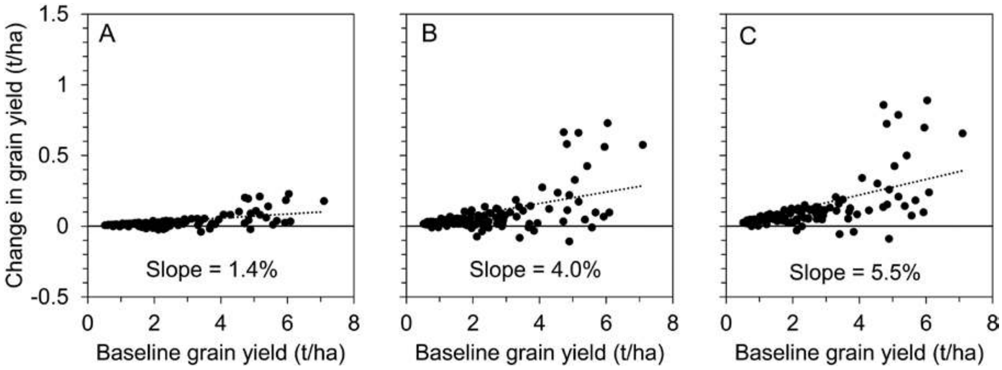


Figure 5. Simulated change in wheat yield with sequential addition of effects on assimilation associated with photorespiratory pathway modifications described by South *et al.* (2019). (A–C) Scenarios indicating effects on simulated grain yield due to the additive effects of (A) Γ*, the chloroplastic CO_2 compensation point, (B) QE, the maximum quantum efficiency and (C) V_{cmax25}, the maximum Rubisco carboxylation rate at 25 °C. The slope of the dotted line in each panel indicates the average percentage change in yield for that combination.

Greater yield improvement occurred with enhancements in the QE and only in high-yielding conditions. While there is a wide range of potential outcomes, significant yield advance is predicted on some occasions. This result highlights the value of linking detailed understanding and modelling of biochemical/metabolic pathways to algorithms that can estimate the likely value at crop yield level of genetic manipulation of the pathway. Where this is possible, even with some degree of uncertainty, it can add some realistic grounding to discussion in improvement programs contemplating use of genetic modification approaches.

Implications for Next-Generation Crop Modelling

The examples detailed above indicate a diversity of potential roles for crop modelling in supporting advance in crop improvement. Crop models have potential to unravel environmental and genetic context dependencies in pursuing more rapid genetic gain from quantitative genomic prediction via the development of CGM-WGP approaches. They have capacity to dissect complex adaptive traits to more amenable determinants and thus inform phenotyping for genetic analysis as well as providing the integrative framework to predict consequences of G×E×M interactions in production environments (Hammer *et al.* 2016*a*, *b*; Chenu *et al.* 2017). This links with the ability to inform targets for crop design at both the plant trait level for quantitative approaches and at the gene network and metabolic pathway level for possibilities with plant engineering (Fig. 2).

So what does this mean for developing the next generation of crop models needed to advance crop improvement? Do we require a range of models that reflect the varying levels of biological organization at play in these approaches to crop improvement? Future advances in crop improvement will be underpinned by significant increases in genomic and phenomic data, and advances in analytical technologies (Ramstein *et al.* 2019), as well as advance in process understanding at various levels of biological organization. While the notion of linking trait genetic architecture with crop growth models to facilitate G-to-P prediction has already been significantly developed (Yin *et al.* 2005; Letort *et al.* 2008; Uptmoor *et al.* 2008; Chenu *et al.* 2009; Boote *et al.* 2016; Yin *et al.* 2016), there is some concern that model structure is not adequate for this purpose. At the outset of the putative use of crop models in breeding, Hammer *et al.* (2002) and Yin *et al.* (2004) noted the need for enhanced plant/crop model structure if models were to play an integral role. There has been some progress towards redesign of crop models so that they incorporate more physiological feedback features and structural rigour so that processes can be better linked to genetics (Hammer *et al.* 2010; Boote *et al.* 2016), but it is largely accepted that more is needed (Yin *et al.* 2016; Wang *et al.* 2019). However, there is also the need for parsimony and speed if models are to form integral components of the quantitative approaches embedded in modern plant breeding programs (Messina *et al.* 2018). Given that changes in yield associated with genetic gain in elite breeding populations are relatively small and often of uncertain origin, the issue becomes not one of predicting yield of each individual within that population, but rather of linking observed phenotypic variation with plausible differences in potentially causal traits using a dynamic CGM.

In addition to the rapid advance in quantitative genomic prediction technologies, there has also been a commensurate development in the understanding of metabolic pathways and plant gene networks at the molecular level. This offers an opportunity to link the physiology of the plant to mechanisms at the molecular level, and potentially thereafter to genome sequences (Chew *et al.* 2017). For example, for photosynthesis, Zhu *et al.* (2007, 2013) have captured this enhanced process understanding in a detailed model of fluxes and used it to consider approaches to pathway optimization. Much of this advance in biological realism has occurred at the (sub)cellular level, although perspectives to develop integrative modelling platforms to facilitate whole plant simulations by linking gene networks, protein synthesis, metabolic pathways, physiology and growth are emerging (Marshall-Colon *et al.* 2017). However, there has also been advance in biological realism associated with representation of ecophysiological mechanisms at plant/crop level (Tardieu 2010). But attempts to connect the detailed (sub)cellular models with less complex mechanistic crop models at the supra-cellular scale remain limited. Hence, the downward control from the organ or plant level on the processes taking place within the cell, which is regulated via hormonal or other signalling cascades, is not well captured (Poorter *et al.* 2013).

Here, we argue that crop modelling for biological realism and parsimony do not need to be independent and perhaps should not be. Biological realism in crop modelling requires formalisms based on insights on ecophysiological mechanisms at plant/crop scale as well as on insights on metabolic processes at cellular scale. Parsimony in crop modelling requires frugality of assumptions and detail in order to achieve robust predictions of crop growth and yield—as simple as possible but no simpler—across diverse genotypes and environments. Multiscale models that operate effectively across levels of biological organization provide an avenue for advance. Chew *et al.* (2014) set out a multiscale model for *Arabidopsis* that uses a modular approach to link component modules that are characteristic of different research domains, such as molecular systems biology (gene regulatory network), crop science (phenology, organ source-sink relations) and physiology (photosynthesis, respiration, metabolism). Wu *et al.* (2019) set out a cross-scale model for major crop species that they used to assess effects of photosynthesis manipulation on crop yield across diverse production environments. These approaches capture the feed-forward and feedback interactions across levels of biological organization in a crop system that determine its functional dynamics.

Models structured to readily utilize algorithms operating at varying levels of biological organization, while using coding and computational advances to facilitate high-speed simulation, could well provide the next generation of crop models needed to support and enhance advances in crop improvement technologies. Hierarchical algorithm nesting is a means to link approaches operating at differing levels of complexity and biological organization while retaining biological realism at all levels. For example, the common use of RUE to drive crop growth in many crop models operating on a daily time step can be expanded to incorporate photosynthesis-light and A–C_i responses of sunlit and shaded canopy components throughout the day (Hammer and Wright 1994; Wu *et al.* 2019) or further to incorporate detailed light ray tracing and canopy architecture (Song *et al.* 2017). The finer resolution adds enhanced realism, and hence an ability to inform the simpler approaches, but it also adds complexity and difficulty in parameterization that limits predictive capacity and utility. Models that integrate understanding of processes across levels of biological organization, providing insight into key phenomena and responses that emerge at higher levels, will likely be most valuable. Messina *et al.* (2019) detail a dynamic model for cohorting of reproductive structures along the ear in maize while accounting for carbon and water supply and demand balances. While the model operates at an organ level, it can generate well-known emergent phenotypes such as the relationship between plant growth, anthesis-silking interval, kernel number and yield, as well as ear phenotypes under drought (e.g. tip kernel abortion). Hence, in addition to advancing our understanding of maize reproductive biology, this predictive capacity supports avenues for breeding and increasing productivity in maize. This interdependence between models formulated at different levels of complexity exemplifies what is required to link biological reality and parsimony in modelling. Simpler parsimonious models require the ecophysiological rigour to incorporate, or generate, observed phenotypic responses at plant/crop scale. Models that are more complex should be able to generate, or inform, the structure and parameterization of the simpler models. The dialogue generated across levels of scale provides the basis for co-learning and the opportunity for evolution of component algorithms as knowledge advances at both molecular and ecophysiological levels. Chew *et al.* (2017) discuss how multiscale modelling might provide a bridge between plant systems biology and crop systems modelling at various levels with benefits for both fundamental and applied research.

Advanced software platforms capable of supporting this approach to crop modelling are emerging. Brown *et al.* (2014) set out the design of the generic plant modelling framework (PMF), now operational in the APSIM cropping systems modelling platform (Holzworth *et al.* 2014), that utilizes a modular approach to maximize code reuse for model component processes and allow flexibility in model structure. The PMF can thus be used to develop models of different complexities and cater for the hierarchical algorithm nesting concept in achieving that. For model developers operating at different levels of complexity to have confidence in using such interdependent models for crop improvement, it must be demonstrable that the model is robust, accurate and reliable. Rigorous validation and testing of predictive acumen need to be part of an ongoing model development and release process. Brown *et al.* (2018) and Wu *et al.* (2019) exemplify what is required using detailed examples for crop model improvement in the APSIM platform. The Arabidopsis Framework Model (Chew *et al.* 2017) was also developed using a modular approach that enabled evaluation of components developed at differing levels of resolution. Computational frameworks, such as OpenAlea (Pradal *et al.* 2008) and yggdrasil (Lang 2019), which have been specifically designed to assist with integrating model components for plant models via user-friendly interfaces, provide an avenue to explore and compare modelling approaches.

Beyond the modelling platforms capable of operating technically across levels of biological organization, most likely the trans-scale and transdisciplinary dialogue engendered in achieving them will drive the effective use of crop modelling in crop improvement. This requires a culture of connectivity, trust and mutual respect among diverse operatives operating at differing biological levels and in different disciplines. Sinclair *et al.* (2004) noted that the successful efforts up to that time to use physiology to improve crop yield, all involved multidisciplinary effort. Chew *et al.* (2017) noted that their integrative modelling operating at the interface of several research communities had the potential to facilitate communication and draw together the different types of understanding from fundamental plant research and crop models. Poorter *et al.* (2013) identified the need for connection between supra-cellular (phenotype down) and sub-cellular (gene network up) systems biology approaches to facilitate advance. Yin and Struik (2008) have advocated the need to operate across levels of biological organization to shape their 'crop systems biology'. The need for effective transdisciplinary dialogue and connectivity is clear. Committed teams with shared vision and effective leadership targeting the building of cross-scale models with a clear purpose provide a means to achieve this. However, innovation in the funding and evaluation of scientific research will be required to support any move to operate with a true cross-disciplinary culture of connectivity.

Sources of Funding

This work was supported by Australian Research Council, Corteva Agriscience and Grains Research and Development Corporation.

Conflict of Interest

None declared.

Acknowledgements

This paper draws on research collaborations and discussions with a diverse range of people, which would not have been possible without the long-term and ongoing support of their respective organizations. Additional financial support from a number of research funding agencies/collaborators, including the Australian Research Council, Corteva Agriscience and Grains Research and Development Corporation, is gratefully acknowledged. The authors acknowledge Springer Publishing and Elsevier for permissions to reuse figures and tables.

Literature Cited

Alam MM, Hammer GL, van Oosterom EJ, Cruickshank AW, Hunt CH, Jordan DR. 2014*a*. A physiological framework to explain genetic and environmental regulation of tillering in sorghum. *The New Phytologist* **203**:155–167.

Alam MM, Mace ES, van Oosterom EJ, Cruickshank A, Hunt CH, Hammer GL, Jordan DR. 2014*b*. QTL analysis in multiple sorghum populations facilitates the dissection of the genetic and physiological control of tillering. *Theoretical and Applied Genetics* **127**:2253–2266.

Allard RW. 1999. *Principles of plant breeding*, 2nd edn. New York: John Wiley & Sons, Inc.

Barker T, Campos H, Cooper M, Dolan D, Edmeades G, Habben J, Schussler J, Wright D, Zinselmeier C. 2005. Improving drought tolerance in maize. *Plant Breeding Reviews* **25**:173–253.

Bernardo R. 2002. *Breeding for quantitative traits in plants*. Woodbury, MN: Stemma Press.

Bertin N, Martre P, Génard M, Quilot B, Salon C. 2010. Under what circumstances can process-based simulation models link genotype to phenotype for complex traits? Case-study of fruit and grain quality traits. *Journal of Experimental Botany* **61**:955–967.

Beveridge CA, Kyozuka J. 2010. New genes in the strigolactone-related shoot branching pathway. *Current Opinion in Plant Biology* **13**:34–39.

Boote KJ, Vallejos CE, Jones, JW, Correll MJ. 2016. Crop modeling approaches for predicting phenotype of grain legumes with linkage to genetic information. In: Yin X, Struik P, eds. *Crop systems biology: narrowing the gap between genotype and phenotype*. Cham, Switzerland: Springer International Publishing, 163–192.

Borrell AK, Mullet JE, George-Jaeggli B, van Oosterom EJ, Hammer GL, Klein PE, Jordan DR. 2014*a*. Drought adaptation of stay-green sorghum is associated with canopy development, leaf anatomy, root growth, and water uptake. *Journal of Experimental Botany* **65**:6251–6263.

Borrell AK, van Oosterom EJ, Mullet JE, George-Jaeggli B, Jordan DR, Klein PE, Hammer GL. 2014*b*. Stay-green alleles individually enhance grain yield in sorghum under drought by modifying canopy development and water uptake patterns. *The New Phytologist* **203**:817–830.

Bos HJ, Neuteboom JH. 1998. Morphological analysis of leaf and tiller number dynamics of wheat (*Triticum aestivum* L.): responses to temperature and light intensity. *Annals of Botany* **81**:131–139.

Brown HE, Huth N, Holzworth D. 2018. Crop model improvement in APSIM: using wheat as a case study. *European Journal of Agronomy* **100**:141–150.

Brown HE, Huth NI, Holzworth DP, Teixeira EI, Zyskowski RF, Hargreaves JNG, Moot DJ. 2014. Plant modelling framework: software for building and running crop models on the APSIM platform. *Environmental Modelling & Software* **62**:385–398.

Campos H, Cooper M, Edmeades GO, Löffler C, Schussler JR, Ibañez M. 2006. Changes in drought tolerance in maize associated with fifty years of breeding for yield in the U.S. Corn Belt. *Maydica* **51**:369–381.

Campos H, Cooper M, Habben JE, Edmeades GO, Schussler JR. 2004. Improving drought tolerance in maize: a view from industry. *Field Crops Research* **90**:19–34.

Chapman SC, Cooper M, Podlich D, Hammer GL. 2003. Evaluating plant breeding strategies by simulating gene action and dryland environment effects. *Agronomy Journal* **95**:99–113.

Chenu K, Chapman SC, Hammer GL, McLean G, Salah HB, Tardieu F. 2008. Short-term responses of leaf growth rate to water deficit scale up to whole-plant and crop levels: an integrated modelling approach in maize. *Plant, Cell & Environment* **31**:378–391.

Chenu K, Chapman SC, Tardieu F, McLean G, Welcker C, Hammer GL. 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.

Chenu K, Porter JR, Martre P, Basso B, Chapman SC, Ewert F, Bindi M, Asseng S. 2017. Contribution of crop models to adaptation in wheat. *Trends in Plant Science* **22**:472–490.

Chenu K, Van Oosterom EJ, McLean G, Deifel KS, Fletcher A, Geetika G, Tirfessa A, Mace ES, Jordan DR, Sulman R, Hammer GL. 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, Millar AJ. 2017. Multi-scale modelling to synergise plant systems biology and crop science. *Field Crops Research* **202**:77–83.

Chew YH, Wenden B, Flis A, Mengin V, Taylor J, Davey CL, Tindal C, Thomas H, Ougham HJ, de Reffye P, Stitt M, Williams M, Muetzelfeldt R, Halliday KJ, Millar AJ. 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.

Choudhary S, Sinclair TR, Messina CD, Cai W, Warner D, Cooper M. 2015. Inhibitor screen for limited-transpiration trait among maize hybrids. *Environmental and Experimental Botany* **109**:161–167.

Choudhary S, Sinclair TR, Messina CD, Cooper M. 2014. Hydraulic conductance of maize hybrids differing in transpiration response to vapour pressure deficit. *Crop Science* **54**:1147–1152.

Comstock RE. 1996. *Quantitative genetics with special reference to plant and animal breeding*. Ames, IA: Iowa State University Press.

Cooper M and Hammer GL. 1996. Synthesis of strategies for crop improvement. In: Cooper M, Hammer GL, eds. *Plant adaptation and crop improvement*. Wallingford, UK: CAB International, ICRISAT & IRRI, 591–623.

Cooper M, Gho C, Leafgren R, Tang T, Messina C. 2014*a*. Breeding drought-tolerant maize hybrids for the US Corn-Belt: discovery to product. *Journal of Experimental Botany* **65**:6191–6204.

Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, Wright D, Graham G. 2014*b*. Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop and Pasture Science* **65**:311–336.

Cooper M, Podlich DW, Smith OS. 2005. Gene-to-phenotype models and complex trait genetics. *Australian Journal of Agricultural Research* **56**:895–918.

Cooper M, Technow F, Messina C, Gho C, Totir LR. 2016. Use of crop growth models with whole-genome prediction. Application to a maize multi-environment trial. *Crop Sci* **56**:2141.

Cooper M, van Eeuwijk FA, Hammer GL, Podlich DW, Messina C. 2009. Modeling QTL for complex traits: detection and context for plant breeding. *Current Opinion in Plant Biology* **12**:231–240.

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G. 2007. FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science (New York, N.Y.)* **316**:1030–1033.

de Pury DGG, Farquhar GD. 1997. Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. *Plant, Cell, & Environment* **20**:537–557.

Dong Z, Danilevskaya O, Abadie T, Messina C, Coles N, Cooper M. 2012. A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. *PLoS ONE* **7**:e43450.

Doudna JA, Charpentier E. 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science (New York, N.Y.)* **346**:1258096.

Duvick DN, Smith JSC, Cooper M. 2004. Long-term selection in a commercial hybrid maize breeding program. *Plant Breeding Reviews* **24**:109–151.

Edgerton MD. 2009. Increasing crop productivity to meet global needs for feed, food, and fuel. *Plant Physiology* **149**:7–13.

Falconer DS, Mackay TFC. 1996. *Introduction to quantitative genetics*, 4th edn. Harlow, UK: Longman Limited.

Farquhar GD, von Caemmerer S, Berry JA. 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. *Planta* **149**:78–90.

Fehr WR. 1987. *Principles of cultivar development (Volume 1). Theory and technique*. New York: Macmillan Publishing Company.

Fischer T, Byerlee D, Edmeades G. 2014. *Crop yields and global food security: will yield increase continue to feed the world?* ACIAR Monograph No. 158. Canberra, Australia: Australian Centre for International Agricultural Research, xxii + 634.

Gaffney J, Schussler J, Löffler C, Cai W, Paszkiewicz S, Messina C, Groeteke J, Keaschall J, Cooper M. 2015. Industry-scale evaluation of maize hybrids selected for increased yield in drought-stress conditions of the US Corn Belt. *Crop Science* **55**:1608–1618.

Garcia-Ruiz A, Cole JB, VanRaden PM, Wiggans GR, Ruiz-López FJ, Van Tassell CP. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proceedings of the National Academy of Sciences of the United States of America* **113**:E3995–E4004.

Garner JB, Douglas ML, Williams SR, Wales WJ, Marett LC, Nguyen TT, Reich CM, Hayes BJ. 2016. Genomic selection improves heat tolerance in dairy cattle. *Scientific Reports* **6**:34114.

Gholipoor M, Choudhary S, Sinclair TR, Messina CD, Cooper M. 2013. Transpiration response of maize hybrids to atmospheric vapour pressure deficit. *Journal of Agronomy and Crop Science* **199**:155–160.

Gholipoor M, Prasad PVV, Mutava RN, Sinclair TR. 2010. Genetic variability of transpiration response to vapor pressure deficit among sorghum genotypes. *Field Crops Research* **119**:85–90.

Grassini P, Yang H, Cassman KG. 2009. Limits to maize productivity in Western Corn-Belt: a simulation analysis for fully irrigated and rainfed conditions. *Agricultural and Forest Meteorology* **149**:1254–1265.

Gu J, Yin X, Stomph TJ, Struik PC. 2014. Can exploiting natural genetic variation in leaf photosynthesis contribute to increasing rice productivity? A simulation analysis. *Plant, Cell & Environment* **37**:22–34.

Hallauer AR, Miranda Filho JB. 1988. *Quantitative genetics in maize breeding*, 2nd edn. Ames, IA: Iowa State University Press.

Hammer GL, Carberry PS, Muchow RC. 1993. Modelling genotypic and environmental control of leaf area dynamics in grain sorghum. I. Whole plant level. *Field Crops Research* **33**:293–310.

Hammer G, Chapman S, van Oosterom E, Podlich D. 2005. Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems. *Australian Journal of Agricultural Research* **56**:947–960.

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.

Hammer GL, Dong Z, McLean G, Doherty A, Messina C, Schussler J, Zinselmeier C, Paszkiewicz S, Cooper M. 2009. Can changes in canopy and/or root systems architecture explain historical maize yield trends in the U.S. Corn Belt? *Crop Science* **49**:299–312.

Hammer G, McLean G, Doherty A, van Oosterom E, Chapman S. 2016*a*. Sorghum crop modelling and its utility in agronomy and breeding. In: Prasad V, Ciampitti I, eds. *Sorghum: state of the art and future perspectives*. Madison, WI: ASA and CSSA. Agronomy Monographs 58.

Hammer G, Messina C, van Oosterom E, Chapman C, Singh V, Borrell A, Jordan D, Cooper M. 2016*b*. Molecular breeding for complex adaptive traits: how integrating crop ecophysiology and modelling can enhance efficiency. In: Yin X, Struik P, eds. *Crop systems biology: narrowing the gap between genotype and phenotype*. Cham, Switzerland: Springer International Publishing, 147–162.

Hammer GL, Kropff MJ, Sinclair TR, Porter JR. 2002. Future contributions of crop modelling—from heuristics and supporting decision-making to understanding genetic regulation and aiding crop improvement. *European Journal of Agronomy* **18**:15–31.

Hammer GL, McLean G, Chapman S, Zheng B, Doherty A, Harrison MT, van Oosterom E, Jordan D. 2014. Crop design for specific adaptation in variable dryland production environments. *Crop and Pasture Science* **65**:614–626.

Hammer GL, Sinclair TR, Chapman SC, van Oosterom E. 2004. On systems thinking, systems biology, and the in silico plant. *Plant Physiology* **134**:909–911.

Hammer GL, van Oosterom E, McLean G, Chapman SC, Broad I, Harland P, Muchow RC. 2010. Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. *Journal of Experimental Botany* **61**:2185–2202.

Hammer GL, Wright GC. 1994. A theoretical-analysis of nitrogen and radiation effects on radiation use efficiency in peanut. *Australian Journal of Agricultural Research* **45**:575–589.

Hao B, Xue Q, Marek TH, Jessup KE, Hou X, Xu W, Bynum ED, Bean BW. 2015*a*. Soil water extraction, water use, and grain yield by drought-tolerant maize on the Texas high plains. *Agricultural Water Management* **155**:11–21.

Hao B, Xue Q, Marek TH, Jessup KE, Becker J, Hou X, Xu W, Bynum ED, Bean BW, Colaizzi PD, Howell TA. 2015*b*. Water use and grain yield in drought-tolerant corn in the Texas high plains. *Agronomy Journal* **107**:1922–1930.

Hao B, Xue Q, Marek TH, Jessup KE, Hou X, Xu W, Bynum ED, Bean BW. 2016. Radiation-use efficiency, biomass production, and grain yield in two maize hybrids differing in drought tolerance. *Journal of Agronomy and Crop Science* **202**:269–280.

Heffner EL, Sorrells ME, Jannink J-L. 2009. Genomic selection for crop improvement. *Crop Science* **49**:1–12.

Henzell RG, Brengman R, Fletcher D, McCosker T. 1992. Relationships between yield and non-senescence ('stay-green') in some grain sorghum hybrids grown under terminal drought stress. In: Foale MA, Henzell RG, Vance PN, eds. Second Australian Sorghum Conference, Gatton, Queensland, Australia. Melbourne, Victoria, Australia: Australian Institute of Agricultural Science, 355–359.

Holland JB, Nyquist W, Cervantes-Martínez CT. 2003. Estimating and interpreting heritability for plant breeding: an update. *Plant Breeding Review* **22**:9–112.

Holzworth DP, Huth NI, deVoil PG, Zurcher EJ, Herrmann NI, McLean G, Chenu K, van Oosterom EJ, Snow V, Murphy C, Moore AD, Brown H, Whish JPM, Verrall S, Fainges J, Bell LW, Peake AS, Poulton PL, Hochman Z, Thorburn PJ, Gaydon DS, Dalgliesh NP, Rodriguez D, Cox H, Chapman S, Doherty A, Teixeira E, Sharp J, Cichota R, Vogeler I, Li FY, Wang E, Hammer GL, Robertson MJ, Dimes JP, Whitbread AM, Hunt J, van Rees H, McClelland T, Carberry PS, Hargreaves JNG, MacLeod N, McDonald C, Harsdorf J, Wedgwood S, Keating BA. 2014. APSIM-evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software* **62**:327–350.

Jordan DR, Hunt CH, Cruickshank AW, Borrell AK, Henzell RG. 2012. The relationship between the stay-green trait and grain yield in elite sorghum hybrids grown in a range of environments. *Crop Science* **52**:1153–1161.

Kholová J, Murugesan T, Kaliamoorthy S, Malayee S, Baddam R, Hammer GL, McLean G, Deshpande S, Hash CT, Craufurd PQ, Vadez V. 2014. Modelling the effect of plant water use traits on yield and stay-green expression in sorghum. *Functional Plant Biology* **41**:1019–1034.

Kim HK, Luquet D, van Oosterom E, Dingkuhn M, Hammer G. 2010*a*. Regulation of tillering in sorghum: genotypic effects. *Annals of Botany* **106**:69–78.

Kim HK, van Oosterom E, Dingkuhn M, Luquet D, Hammer G. 2010*b*. Regulation of tillering in sorghum: environmental effects. *Annals of Botany* **106**:57–67.

Lafarge TA, Hammer GL. 2002. Tillering in grain sorghum over a wide range of population densities: modelling dynamics of tiller fertility. *Annals of Botany* **90**:99–110.

Lang M. 2019. Yggdrasil: a Python package for integrating computational models across languages and scales. *Silico Plants* **2019**:diz001.

Letort V, Mahe P, Cournède PH, de Reffye P, Courtois B. 2008. Quantitative genetics and functional-structural plant growth models: simulation of quantitative trait loci detection for model parameters and application to potential yield optimization. *Annals of Botany* **101**:1243–1254.

Long SP, Marshall-Colon A, Zhu XG. 2015. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161**:56–66.

Long SP, Zhu XG, Naidu SL, Ort DR. 2006. Can improvement in photosynthesis increase crop yields? *Plant, Cell & Environment* **29**:315–330.

Lynch M, Walsh B. 1998. *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer Associates, Inc.

Lyon DJ, Hammer GL, McLean GB, Blumenthal JM. 2003. Simulation supplements field studies to determine no-till dryland corn population recommendations for semiarid western Nebraska. *Agronomy Journal* **95**:884–891.

Mace ES, Singh V, Van Oosterom EJ, Hammer GL, Hunt CH, Jordan DR. 2012. QTL for nodal root angle in sorghum (*Sorghum bicolor* L. Moench) co-locate with QTL for traits associated with drought adaptation. *Theoretical and Applied Genetics* **124**:97–109.

Manschadi AM, Christopher J, deVoil P, Hammer GL. 2006. The role of root architectural traits in adaptation of wheat to water-limited environments. *Functional Plant Biology* **33**:823–837.

Marshall-Colon A, Long SP, Allen DK, Allen G, Beard DA, Benes B, von Caemmerer S, Christensen AJ, Cox DJ, Hart JC, Hirst PM, Kannan K, Katz DS, Lynch JP, Millar AJ, Panneerselvam B, Price ND, Prusinkiewicz P, Raila D, Shekar RG, Shrivastava S, Shukla D, Srinivasan V, Stitt M, Turk MJ, Voit EO, Wang Y, Yin X, Zhu XG. 2017. Crops *in silico*: generating virtual crops using an integrative and multi-scale modeling platform. *Frontiers in Plant Science* **8**:786.

Mason MG, Ross JJ, Babst BA, Wienclaw BN, Beveridge CA. 2014. Sugar demand, not auxin, is the initial regulator of apical dominance. *Proceedings of the National Academy of Sciences of the United States of America* **111**:6092–6097.

Messina C, Hammer G, Dong Z, Podlich D, Cooper M. 2009. Modelling crop improvement in a G*E*M framework via gene-trait-phenotype relationships. In: Sadras VO, Calderini D, eds. *Crop physiology: applications for genetic improvement and agronomy*. Amsterdam, The Netherlands: Acad. Press, 235–265.

Messina CD, Hammer GL, McLean G, Cooper M, van Oosterom EJ, Tardieu F, Chapman SC, Doherty A, Gho C. 2019. On the dynamic determinants of reproductive failure under drought in maize. *Silico Plants* **2019**:diz003.

Messina CD, Jones JW, Boote KJ, Vallejos CE. 2006. A gene-based model to simulate soybean development and yield responses to environment. *Crop Science* **46**:456–466.

Messina CD, Podlich D, Dong Z, Samples M, Cooper M. 2011. Yield-trait performance landscapes: from theory to application in breeding maize for drought tolerance. *Journal of Experimental Botany* **62**:855–868.

Messina CD, Sinclair TR, Hammer GL, Curan D, Thompson J, Oler Z, Gho C, Cooper M. 2015. Limited-transpiration trait may increase maize drought tolerance in the US Corn Belt. *Agronomy Journal* **107**:1978.

Messina CD, Technow F, Tang T, Totir R, Gho C, Cooper M. 2018. Leveraging biological insight and environmental variation to improve phenotypic prediction: integrating crop growth models (CGM) with whole genome prediction (WGP). *European Journal of Agronomy* **100**:151–162.

Meuwissen TH, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**:1819–1829.

Minorsky PV. 2003. Achieving the *in silico* plant: systems biology and the future of plant biological research. *Plant Physiology* **132**:404–409.

Morrell PL, Buckler ES, Ross-Ibarra J. 2011. Crop genomics: advances and applications. *Nature Reviews Genetics* **13**:85–96.

Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES, Kresovich S. 2013. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences of the United States of America* **110**:453–458.

Mounce RB, O'Shaughnessy SA, Blaser BC, Colaizzi PD, Evett SR. 2016. Crop response of drought-tolerant and conventional maize hybrids in a semiarid environment. *Irrigation Science* **34**:231–244.

Mueller ND, Gerber JS, Johnston M, Ray DK, Ramankutty N, Foley JA. 2012. Closing yield gaps through nutrient and water management. *Nature* **490**:254–257.

Poorter H, Anten NP, Marcelis LF. 2013. Physiological mechanisms in plant growth models: do we need a supra-cellular systems biology approach? *Plant, Cell & Environment* **36**:1673–1690.

Potgieter AB, Lobell DB, Hammer GL, Jordan DR, Davis P, Brider J. 2016. Yield trends under varying environmental conditions for sorghum and wheat across Australia. *Agricultural and Forest Meteorology* **228**:276–285.

Pradal C, Dufour-Kowalski S, Boudon F, Fournier C, Godin C. 2008. OpenAlea: a visual programming and component-based software platform for plant modelling. *Functional Plant Biology* **35**:751–760.

Ramstein GP, Jensen SE, Buckler ES. 2019. Breaking the curse of dimensionality to identify causal variants in breeding 4. *Theoretical and Applied Genetics* **132**:559–567.

Ray DK, Mueller ND, West PC, Foley JA. 2013. Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* **8**:e66428.

Reyes A, Messina CD, Hammer GL, Liu L, van Oosterom E, Lafitte R, Cooper M. 2015. Soil water capture trends over 50 years of single-cross maize (*Zea mays* L.) breeding in the US Corn-Belt. *Journal of Experimental Botany* **66**:7339–7346.

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

Reymond M, Muller B, Tardieu F. 2004. Dealing with the genotype × environment interaction via a modelling approach: a comparison of QTLs of maize leaf length or width with QTLs of model parameters. *Journal of Experimental Botany* **55**:2461–2472.

Rincent R, Kuhn E, Monod H, Oury FX, Rousset M, Allard V, Le Gouis J. 2017. Optimization of multi-environment trials for genomic selection based on crop models. *Theoretical and Applied Genetics* **130**:1735–1752.

Rosenow DT. 1977. Breeding for lodging resistance in sorghum. In: Loden HD, Wilkinson D, eds. Proceedings of the 32nd Annual Corn and Sorghum Research Conference. Washington, DC, USA: American Seed Trade Association, 171–185.

Rosenow DT, Quisenberry JE, Wendt CW, Clark LE. 1983. Drought tolerant sorghum and cotton germplasm. *Agricultural Water Management* **7**:207–222.

Roth JA, Ciampitti IA, Vyn TJ. 2013. Physiological evaluations of recent drought-tolerant maize hybrids at varying stress levels. *Agronomy Journal* **105**:1129–1141.

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.

Salesse-Smith CE, Sharwood RE, Busch FA, Kromdijk J, Bardal V, Stern DB. 2018. Overexpression of Rubisco subunits with RAF1 increases Rubisco content in maize. *Nature Plants* **4**:802–810.

Seaton DD, Smith RW, Song YH, MacGregor DR, Stewart K, Steel G, Foreman J, Penfield S, Imaizumi T, Millar AJ, Halliday KJ. 2015. Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. *Molecular Systems Biology* **11**:776.

Shekoofa A, Sinclair TR, Messina CD, Cooper M. 2015. Variation among maize hybrids in response to high vapour pressure deficit at high temperatures. *Crop Science* **55**:1–5.

Sinclair TR. 1986. Water and nitrogen limitations in soybean grain production. I. Model development. *Field Crops Research* **15**:125–141.

Sinclair TR, Purcell LC, Sneller CH. 2004. Crop transformation and the challenge to increase yield potential. *Trends in Plant Science* **9**:70–75.

Sinclair TR, Zwieniecki MA, Holbrook NM. 2010. Low leaf hydraulic conductance associated with drought tolerance in soybean. *Physiologia Plantarum* **132**:446–451.

Singh V, van Oosterom EJ, Jordan DR, Hammer GL. 2011. Genetic variability and control of root architecture in sorghum. *Crop Science* **51**:2011–2020.

Singh, V, van Oosterom EJ, Jordan DR, Hammer GL. 2012. Genetic control of nodal root angle in sorghum and its implications on water extraction. *European Journal of Agronomy* **42**:3–10.

Smith S, Cooper M, Gogerty J, Löffler C, Borcherding D, Wright K. 2014. Maize. In: Smith S, Diers B, Specht J, Carver B, eds. *Yield gains in major U.S. field crops*. Madison, WI: CSSA Special Publication 33, 125–171.

Song Q, Wang Y, Qu M, Ort DR, Zhu XG. 2017. The impact of modifying photosystem antenna size on canopy photosynthetic efficiency—development of a new canopy photosynthesis model scaling from metabolism to canopy level processes. *Plant, Cell & Environment* **40**:2946–2957.

Soufizadeh S, Munaro E, McLean G, Massignam A, van Oosterom EJ, Chapman SC, Messina C, Cooper M, Hammer GL. 2018. Modelling the nitrogen dynamics of maize crops—enhancing the APSIM maize model. *European Journal of Agronomy* **100**:118–131.

South PF, Cavanagh AP, Liu HW, Ort DR. 2019. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science* **363**:eaat9077.

Tardieu F. 2003. Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. *Trends in Plant Science* **8**:9–14.

Tardieu F. 2010. Why work and discuss the basic principles of plant modelling 50 years after the first plant models? *Journal of Experimental Botany* **61**:2039–2041.

Tardieu F, Tuberosa R. 2010. Dissection and modelling of abiotic stress tolerance in plants. *Current Opinion in Plant Biology* **13**:206–212.

Technow F, Messina CD, Totir LR, Cooper M. 2015. Integrating crop growth models with whole genome prediction through approximate Bayesian computation. *PLoS One* **10**:e0130855.

Turner NC. 2004. Agronomic options for improving rainfall-use efficiency of crops in dryland farming systems. *Journal of Experimental Botany* **55**:2413–2425.

Uptmoor R, Schrag T, Stutzel H, Esch E. 2008. Crop model based QTL analysis across environments and QTL based estimation of time to floral induction and flowering in *Brassica oleracea*. *Molecular Breeding* **21**: 205–216.

Vadez V, Kholova J, Medina S, Kakkera A, Anderberg H. 2014. Transpiration efficiency: new insights into an old story. *Journal of Experimental Botany* **65**:6141–6153.

van Eeuwijk FA, Bustos-Korts D, Millet EJ, Boer MP, Kruijer W, Thompson A, Malosetti M, Iwata H, Quiroz R, Kuppe C, Muller O, Blazakis KN, Yu K, Tardieu F, Chapman SC. 2019. Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. *Plant Science* **282**:23–39.

van Oosterom EJ, Borrell AK, Deifel K, Hammer GL. 2011. Does increased leaf appearance rate enhance adaptation to post-anthesis drought stress in sorghum? *Crop Science* **51**:2728–2740.

van Oosterom EJ, Yang Z, Zhang F, Deifel KS, Cooper M, Messina CD, Hammer GL. 2016. Hybrid variation for root system efficiency in maize: potential links to drought adaptation. *Functional Plant Biology* **43**:502–511.

Voss-Fels KP, Cooper M, Hayes BJ. 2019. Accelerating crop genetic gains with genomic selection. *Theoretical and Applied Genetics* **132**:669–686.

Walsh B, Lynch M. 2018. *Evolution and selection of quantitative traits*. Oxford: Sinauer Associates, Oxford University Press.

Wang E, Brown HE, Rebetzke GJ, Zhao Z, Zheng B, Chapman SC. 2019. Improving process-based crop models to better capture genotype×environment×management interactions. *Journal of Experimental Botany* **70**:2389–2401.

Wu A, Doherty A, Farquhar GD, Hammer GL. 2018. Simulating daily field crop canopy photosynthesis: an integrated software package. *Functional Plant Biology* **45**:362–377.

Wu A, Hammer GL, Doherty A, von Caemmerer S, Farquhar GD. 2019. Quantifying impacts of enhancing photosynthesis on crop yield. *Nature Plants* **5**:380–388.

Xin CP, Tholen D, Devloo V, Zhu XG. 2015. The benefits of photorespiratory bypasses: how can they work? *Plant Physiology* **167**:574–585.

Xu W, Subudhi PK, Crasta OR, Rosenow DT, Mullet JE, Nguyen HT. 2000. Molecular mapping of QTLs conferring stay-green in grain sorghum (*Sorghum bicolor* L. Moench). *Genome* **43**:461–469.

Yang Z, Sinclair TR, Zhuc M, Messina CD, Cooper M, Hammer GL. 2012. Temperature effect on transpiration response of maize plants to vapour pressure deficit. *Environmental and Experimental Botany* **78**:157–162.

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. *The 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.

Yin X, Struik PC, Gu J, Wang H. 2016. Modelling QTL-trait-crop relationships: past experiences and future prospects. In: Yin X, Struik P, eds. *Crop systems biology: narrowing the gap between genotype and phenotype*. Cham, Switzerland: Springer International Publishing. pp. 193–218.

Yin X, Struik PC, Kropff MJ. 2004. Role of crop physiology in predicting gene-to-phenotype relationships. *Trends in Plant Science* **9**:426–432.

Yin X, Struik PC, van Eeuwijk FA, Stam P, Tang J. 2005. QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley. *Journal of Experimental Botany* **56**:967–976.

Zhu XG, de Sturler E, Long SP. 2007. Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm. *Plant Physiology* **145**:513–526.

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.