Chapter 10
Crop Systems Biology: Where Are We and Where to Go?

Xinyou Yin and Paul C. Struik

Abstract The preceding chapters outline approaches in systems biology, genetic mapping and crop modelling, and have shown whether and how these approaches could potentially be integrated to form an effective 'crop systems biology' approach in support of crop improvement. To fulfil the great expectations from the integrated modelling, crop models should be improved based on understandings at lower organizational levels, in the meanwhile ensuring that model-input parameters can be easily phenotyped. The '*crop systems biology*' approach is believed ultimately to realize the expected roles of modelling in narrowing genotype-phenotype gaps and predicting the phenotype from genomic data. Such an approach could be an important tool to solve some imminent food-, feed-, and energy-related, 'real-world' problems.

10.1 Why Crop Systems Biology?

Ecophysiological crop modelling has gradually become a research method and discipline since the first plant modelling paper was published by de Wit (1959). For understanding of a system, models appear good tools for heuristics, for example, to make explicit the importance of properties of system elements in the context of the whole system. For applications, the modelling approach has been predominantly devoted to higher aggregation levels (e.g., optimising agronomic management actions, assessing the impact of climate change on agroecosystems, designing sustaining cropping and farming systems, analysing global yield gaps, etc.). Modelling applications at lower aggregation levels such as designing crop ideotypes and cultivars based on analysing genotype-phenotype relationships have progressed slowly (Jackson et al. 1996; Boote et al. 2013), although use of models as a tool to design crop ideotypes has long been recognised (see review of Loomis et al. 1979).

X. Yin (✉) • P.C. Struik
Department of Plant Sciences, Centre for Crop Systems Analysis, Wageningen University,
P.O. Box 430, 6700 AK Wageningen, The Netherlands
e-mail: Xinyou.Yin@wur.nl

X. Yin, P.C. Struik (eds.), *Crop Systems Biology*,
DOI 10.1007/978-3-319-20562-5_10

However, applications for higher and lower aggregation levels cannot be separated absolutely. For example, there have been calls for more mechanistic models to estimate the impact of global CO_2 fertilisation (Yin 2013; Sun et al. 2014). Also, breeding for better crops is essential to improve global food security in the face of climate change. To better address the issue at the higher aggregation levels, developing modelling approaches to study genotype-phenotype relationships becomes increasingly important (Chenu et al. 2009). This latter issue has been addressed internationally either in conference sessions (e.g., Weiss 2003; Cooper and Hammer 2005) or in special symposia or workshops (e.g., Spiertz et al. 2007).

Despite the great effort that has been made, current crop models are mostly crude, in terms of their ability of treating gene-trait-crop relationships. Systems simulation modelling has long been suggested as a powerful tool to understand crop yield formation and to support crop improvement (Loomis et al. 1979). Expectations for modelling in support of modern breeding are high (Dwivedi et al. 2007). However, according to Lawlor (2002), the lack of truly ‘mechanistic’ crop simulation models (which make use of biochemical information) is a major constraint to advance the understanding of crop yield traits. Boote et al. (2013) also emphasised the needs for more mechanisms in crop models when used for characterising genotype-phenotype relationships. Such a need has been underlined in several chapters of this book.

The modelling studies at the crop level using some knowledge of fundamental plant biology (e.g., biochemistry) are currently sporadic, modelling results published so far to analyse yield traits are inconsistent, and some models are based on untested hypotheses. We, therefore, have proposed a more systematic modelling approach – ‘crop systems biology’ (Yin and Struik 2007, 2008, 2010), to analyse complex traits at the crop level, not only with the aim of establishing close links with understanding at the gene or genome level, but also in terms of its comprehensive reliance on the whole-metabolism biochemistry and physiology. Therefore, the proposed crop systems biology is a crop-level approach to modelling complex crop traits relevant to global food production and energy supply, via establishing the links between ‘omics’-level information, underlying biochemical understanding, and physiological component processes. Crop systems biology, as a research realm, has both fundamental and applied features. For fundamental aspects, crop systems biology models should provide biological interpretation of those phenomena such as genotype-by-environment (G×E) interactions, epistasis, and pleiotropy that prove recalcitrant in genetics. In terms of applications, the goal of crop systems biology is to become a robust tool in support of crop improvement programmes (Yin and Struik 2008).

10.2 Roadmap to Crop Systems Biology; Where Are We?

Development of crop systems biology models certainly depends on what trait a researcher wants to target. Although other traits have also been modelled (e.g., Martre et al. 2003), crop yield is a complex trait that most existing crop simulation models have attempted to predict. It may not be surprising that simulation of yield formation should be a first major focus trait in crop systems biology. In addition, if crop yield can be modelled accurately, underlying mechanisms for traits related to resource use efficiencies (such as water use efficiency or nitrogen use efficiency) can be analysed accordingly.

This book collected papers from several leading groups, where some initial ideas to develop crop systems biology models have been explored or examples to apply such models in analysing topical crop physiological and breeding questions typically for manipulating crop yield or related traits were described. As introductory material, Chap. 1 by Baldazzi et al. provides some fundamentals of quantitative genetics (particularly mapping of quantitative trait loci QTL), approaches in systems biology for modelling cellular, gene regulatory and metabolic networks, and challenges in integrating these networks into plant or crop models. Chapter 2 by Xu and Buck-Sorlin describes a new, morphologically explicit modelling approach called FSPM (Functional-Structural Plant Modelling) and its potential applications in breeding. Some crop modellers argue that detailed morphological properties have also been captured in classical crop models. Nevertheless, probably because FSPM can create virtual plants visualised in dynamic 3-D pictures, there has been a high level of enthusiasm for applying this FSPM approach, from students to professors, in various aspects of research and education in plant and crop science. Xu and Buck-Sorlin describe how FSPM was linked with QTL analysis, which would assist to breed for various plant traits (morphological and architectural traits in particular). Both systems biology and crop modelling rely on bioinformatics and biometrics or statistics in analysing and interpreting either measured or simulated data. Chapter 3 by Bustos et al. provides statistical approaches (linear mixed models as the default model class) in the context of G×E interactions. While quantitative genetics developed from statistical approaches is fundamental for guiding classical breeding, the factorial regression as discussed in Chap. 3 is a statistical approach closest to the concept of crop growth modelling in capturing the response of physiological and agronomic traits in response to environmental variables. As discussed by Bustos et al., statistical models and crop growth models complement each other. For example, the factorial regression cannot generate tempo-spatial profiles of the trait under study and its associated components. In contrast, crop model simulated responses can be analysed in the context of adapting the crop to the changing environment, allowing the virtual profiling of plants and analysis of how processes interact when crops are perturbed by one or several changes. Chapter 4 by Génard et al. showed how this knowledge generated through *in silico* profiling can be used to decipher G×E interactions so as to build genotypes adapted to particular conditions, i.e., plant ideotypes. Similar line of reasoning is continued in Chap. 5, where Luquet et al. attempted an *in silico* prediction of margins for genetic improvement of rice using Ecomeristem, a model that seems to lie in between FSPM and classical crop models. The target was to analyse the trade-off between early vigour and drought tolerance, and to design rice ideotypes that combine the two traits. One of the traits associated with drought response is the restricted transpiration. In Chap. 6, Sinclair et al. described the steps of modelling-physiology-transcriptomics-genetic screening they followed in developing soybean cultivars with the desired trait that restricts transpiration. Also based on this restricted transpiration and other examples, Hammer al. argued in Chap. 7 that crop ecophysiology and functional modelling can provide an effective link between molecular and organism scales and can enhance molecular breeding by adding value to genetic prediction approaches. A physiological framework that facilitates dissection and modelling of complex traits can inform phenotyping methods for marker/gene detection and underpin prediction of likely phenotypic consequences of trait and genetic variation in target environments. This is further consolidated in Chap. 8 by Boote et al., who showed model-based approaches revealing that manipulating trait values is beneficial in one environment but not in another environment. They also showed how to link model input parameters with allelic effects of several known genes to establish gene-based modelling of growth and seed yield in common bean, based on the framework of White and Hoogenboom (1996). Perhaps, few model-input parameters are controlled only by pleiotropic effects of a few well characterised major genes; a more likely scenario is that like other quantitative traits, each model-input parameter has own specific polygenes underlying its phenotypic variation. This is a basis of the most active line in this research realm over the last 15 years, i.e., QTL-based crop modelling, and experiences and future prospects are comprehensively reviewed in Chap. 9 by Yin et al.

How far do these states-of-the-art described in the preceding chapters of this book reach the high expectations for crop systems biology? In our judgement, these are just in the juvenile phase of the first of the two-step roadmap that Yin and Struik (2007) proposed for crop systems biology, i.e., the prototype stage and the advanced stage. For the first, a widely used framework or concepts in many existing crop simulation models including processes such as photosynthesis, respiration and assimilate partitioning can still be used. At the level of these processes, there are rich physiological and biochemical data and therefore the understandings are of the highest confidence. For this first step, *crop systems biology* models may not be necessarily more complex than existing crop simulation models in structure, nor is their additional input requirement. The latter is important, and model-input parameters should include those close to the traits breeders score for selection. We should also seek opportunities to derive model-input parameters in parallel with the development of high-throughput phenotyping (White et al. 2012; Parent and Tardieu 2014). However, model algorithms for individual processes are supposed to be more mechanistic than those used in existing crop models. In many cases, a summary form of a detailed biochemical model – e.g., the photosynthesis model of Farquhar et al. (1980) coupled to CO_2 diffusion algorithms (Yin and Struik 2009) – can be incorporated as a sub-model, and this has been incorporated into the crop model GECROS (Yin 2013; Gu et al. 2014). In other cases, direct results or stoichiometries from biochemical studies (e.g., examination of the biochemical pathways for production of proteins, carbohydrates and lipids from glucose by Penning de Vries et al. 1974) can be utilized. A prototype of crop systems biology models needs to be made from this first step, in which physiological and biochemical information at the process level is assembled and then scaled up to the crop level in a way similar to temporal and spatial integrations as practised in conventional crop simulation models. In relation to crop improvement, a key element of the first step would be to identify the parts of mechanisms that are conservative in energy and water transfer, and in carbon and nitrogen metabolisms, and the parts of mechanisms that show genetic variation and are potentially amenable to selection. In case of grain yield, the prototype models should allow identification and quantitative assessment of specific parts of processes, which could be altered to achieve improvement of yield potential. The parts showing genetic variation can be identified by genetic analysis. For example, in Chap. 7, Hammer et al. indicated that the crop model APSIM has been recently upgraded to structure a generic cereal template for more explanatory approaches to modelling the hierarchy of physiological determinants of crop growth and development. They showcased the stay-green phenotype in sorghum, which was generated as an emergent consequence of canopy nitrogen dynamics associated with genetic differences in dwarfing. Taller genotypes required more nitrogen for structural stem tissue, leaving less available for leaves, which was more rapidly diminished by translocation to grain during grain-filling. Hence, "stay-green" was generated as an emergent consequence in the shorter genotypes in response to genetic differences in plant height.

10.3 Roadmap to Crop Systems Biology; Where to Go?

Perhaps in parallel with the first step, crop systems biology modelling could move to the second step as progresses at the 'omics'-level understanding are being made, towards reaching down to lower organizational levels. For this, it is necessary to map the organization levels and the communication systems between these levels for the different key processes (Struik et al. 2007). Modelling for reaching down to the lower levels is most likely to be done in a manner of one-process-at-a-time; and in this respect, a modular design of the model is important to ensure that changes of a sub-model will not affect other parts of the model. Welch et al. (2003) have already developed a neural network model of *Arabidopsis* flowering time control, based on studies on qualitative, genetic characterization of major flowering time genes in this model plant species. Wilczek et al. (2009) continued the work, using the concept of dynamic simulation as commonly used in crop modelling, by linking individual model coefficients to the activities of specific genes and their regulators involved in the transitions to flowering in *Arabidopsis thaliana*. Similar modelling studies could be performed for phenology of crop species (see an example for maize, Dong et al. 2012, and for wheat, Brown et al. 2013). Further, existing modelling of metabolisms, such as the Benson-Calvin cycle of photosynthesis and the photorespiratory cycle (cf. Giersch 2000) and nitrogen assimilation in relation to the activity of key enzymes (e.g., nitrate reductase and glutamine synthetase), could also be added. With the rapid development of functional genomics in the wake of high-throughput technologies, combined studies of physiological components with gene expression profiles should illustrate the function of genes, biochemical pathways and cellular processes that are affected in a coordinated manner (Stitt and Fernie 2003). Such studies should lay the groundwork for elucidating regulatory networks and causal linkages among gene products, biochemistry and whole-plant physiology. Summary models for a particular metabolism or process are expected to increasingly become available. Sometimes, models capable of assessing the impact of altered biochemical pathways are not necessarily too much more complex than existing models. For example, von Caemmerer (2013) showed that the models that can simulate the photorespiratory bypass and the bicarbonate pumps (both have been explored as targets to reduce the CO_2 compensation point in C_3 photosynthesis) are only slightly more complex than the standard photosynthesis model of Farquhar et al. (1980). Such summary models can potentially be embedded into crop systems biology models.

Clearly, different temporal, spatial and structural scales are required for different components, pathways, and processes of the model; and this has been showcased in the recent multiscale model of Chew et al. (2014) for *Arabidopsis* that integrates gene dynamics, carbon partitioning, organ architecture, and development response to endogenous and environmental signals. Chapter 1 of this book by Baldazzi et al. discussed the challenges in the multiscale modelling by combining information from molecular biology and genetics with crop models in relation to environmental factors and agricultural practices. Ultimately, crop systems biology may develop into a highly computer-intensive discipline. Such coupled models should inform how and where those recalcitrant genetic phenomena (G×E interactions, epistasis and pleiotropy) come about. They should also allow identification of specific parts of metabolic pathways and processes, which could be altered via genetic engineering to achieve improvement of crop yield potential (Zhu et al. 2011). These specific parts should be amenable to the analysis by the 'omics' approach in terms of the expressions of specific genes, proteins or enzyme activities. For example, gene expression of aquaporins, the putative proteins involved in regulation of water and CO_2 diffusion inside leaves (see also Chap. 6 of this book by Sinclair et al.) have been found to explain most of the variation of stomatal and mesophyll conductance during water stress and recovery in olive (Perez-Martin et al. 2014). In short, these models should ultimately enable *in silico* assessment of crop response to genetic fine-tuning under defined environmental scenarios, thereby being powerful tools in supporting breeding or genetic engineering for complex crop traits. Again, the parsimony rule, especially in terms of the number of required parameters, also applies to models for navigating biological complexity across scales (Hammer et al. 2006), as estimating many parameters in any model is a daunting task (see also Chap. 8 by Boote et al.), even when using advanced bioinformatics or data mining tools (Martin et al. 2015). After all, model sophistication should not be achieved at the cost of model heuristics.

10.4 Crop Systems Biology Needs Cross-Discipline Efforts in Concert

Manipulation of a relatively small number of genes (notably, dwarfing and photoperiod-insensitivity genes in many crops) had resulted in the first 'Green Revolution'. For the next 'Green Revolution' to happen, we have to deal with many genes so that they work in concert. Advances in genetics and genomics, when integrated via systems biology approaches, can offer unprecedented opportunities to examine crop genetic variation and utilize this variability for breeding purposes in different target environments. However, alterations made at the genome level, although substantial, could have little effect on the crop-level phenotypes (Sinclair et al. 2004; Yin and Struik 2008). Systems biology should not be the privilege of only those working on molecular, sub-cellular or cellular levels. To allow genomics and systems biology to have significant impact, the information from fundamental plant biology should reach up to the crop level, and 'crop systems biology' should be established to deal with complex 'gene-trait-crop' relationships and to enhancing the prediction of the phenotype from genomic information. Recent work of Chew et al. (2014) has shown the promise of such a multiscale approach, based on understandings for the model species *Arabidopsis*. For that to happen for crop species, it is necessary to have the long-term, multi-disciplinary efforts to build the links between geneticists, systems biologists, breeders and crop ecophysiologists towards the next 'Green Revolution' to solve some imminent food-, feed-, and energy-related, 'real-world' problems. By then, the chain as envisaged in Chap. 6 by Sinclair et al., i.e., 'from model to phenotype to genotype to cultivar', can become a reality more than ever. However, at this moment, as expressed in Chap. 8 by Boote et al., "the disciplines have diverged so much that geneticists are not well connected with the field phenotyping, and crop modellers are not connected with the geneticists".

We hope that the publication of this book on crop systems biology promotes cross fertilization between disciplines and can catalyse some joint efforts from the science community to correct that divergence.

References

Boote KJ, Jones JW, White JW, Asseng S, Lisaso JI (2013) Putting mechanisms into crop production models. Plant Cell Environ 36:1658–1672

Brown HE, Jamieson PD, Brooking IR, Moot DJ, Huth NI (2013) Integration of molecular and physiological models to explain time of anthesis in wheat. Ann Bot 112:1683–1703

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

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. Proc Natl Acad Sci USA 111(39):E4127–E4136

Cooper M, Hammer GL (2005) Complex traits and plant breeding – can we understand the complexities of gene-to-phenotype relationships and use such knowledge to enhance plant breeding outcomes? Aust J Agric Res 56:869–872

de Wit CT (1959) Potential photosynthesis of crop surfaces. Neth J Agric Sci 7:141–149

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

Dwivedi SL, Crouch JH, Mackill DJ, Xu Y, Blair MW, Ragot M, Upadhyaya HD, Ortiz R (2007) The molecularization of public sector crop breeding: progress, problems and prospects. Adv Agron 95:163–318

Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. Planta 149:78–90

Giersch C (2000) Mathematical modelling of metabolism. Curr Opin Plant Biol 3:249–253

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 Environ 37:22–34

Hammer G, Cooper M, Tardieu F, Welch S, Walch B, van Eeuwijk F, Chapman S, Podlich D (2006) Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci 11:587–593

Jackson P, Robertson M, Cooper M, Hammer G (1996) The role of physiological understanding in plant breeding, from a breeding perspective. Field Crops Res 49:11–37

Lawlor DW (2002) Carbon and nitrogen assimilation in relation to yield: mechanisms are the key to understanding production systems. J Exp Bot 53:773–787

Loomis RS, Rabbinge R, Ng E (1979) Explanatory models in crop physiology. Annu Rev Plant Physiol 30:339–367

Martin L, Cook C, Matasci N, Williams J, Bastow R (2015) Data mining with iPlant: a meeting report from the 2013 GARNet workshop, *Data mining with iPlant*. J Exp Bot 66:1–6

Martre P, Porter JR, Jamieson PD, Triboï E (2003) Modeling grain nitrogen accumulation and protein composition to understand the sink/source regulations of nitrogen remobilization for wheat. Plant Physiol 133:1959–1967

Parent B, Tardieu F (2014) Can current crop models be used in the phenotyping era for predicting the genetic variability of yield of plants subjected to drought or high temperature? J Exp Bot 65:6179–6189

Penning de Vries FWT, Brunsting AHM, van Laar HH (1974) Products, requirements and efficiency of biosynthesis: a quantitative approach. J Theor Biol 45:339–377

Perez-Martin A, Michelazzo C, Torres-Ruiz JM, Flexas J, Fernandez JE, Sebastiani L, Diaz-Espejo A (2014) Regulation of photosynthesis and stomatal and mesophyll conductance under water stress and recovery in olive trees: correlation with gene expression of carbonic anhydrase and aquaporins. J Exp Bot 65:3143–3156

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

Spiertz JHJ, Struik PC, van Laar HH (2007) Scale and complexity in plant systems research: gene-plant-crop relations, vol 21, Wageningen UR frontis series. Springer, Dordrecht, 329 pp

Stitt M, Fernie AR (2003) From measurements of metabolites to metabolomics: an 'on the fly' perspective illustrated by recent studies of carbon–nitrogen interactions. Curr Opin Biotechnol 14:136–144

Struik PC, Cassman KG, Koornneef M (2007) A dialogue on interdisciplinary collaboration to bridge the gap between plant genomics and crop sciences. In: Spiertz JHJ, Struik PC, van Laar HH (eds) Scale and complexity in plant systems research: gene-plant-crop relations. Springer, Dordrecht, pp 319–328

Sun Y, Gu L, Dickinson RE, Norby RJ, Pallardy SG, Hoffman FM (2014) Impact of mesophyll diffusion on estimated global land CO_2 fertilization. Proc Natl Acad Sci U S A 111:15774–15779

von Caemmerer S (2013) Steady-state models of photosynthesis. Plant Cell Environ 36:1617–1630

Weiss A (2003) Introduction. Agron J 95:1–3

Welch SM, Roe JL, Dong Z (2003) Genetic neural network model of flowering time control in *Arabidopsis thaliana*. Agron J 95:71–81

White JW, Hoogenboom G (1996) Simulating effects of genes for physiological traits in a process-oriented crop model. Agron J 88:416–422

White JW, Andrade-Sanchez P, Gore MA, Bronson KF, Coffelt TA, Conley MM, Feldmann KA, French AN, Heun JT, Hunsaker DJ, Jenk MA, Kimball BA, Roth RL, Strand RJ, Thorp KR, Wall GW, Wang G (2012) Field-based phenomics for plant genetics research. Field Crop Res 133:101–112

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J, Egan JF, Moyers BT, Petipas R, Giakountis A, Charbi E, Coupland G, Welch SM, Schmitt J (2009) Effects of genetic perturbation on seasonal life history plasticity. Science 323:930–934

Yin X (2013) Improving ecophysiological simulation models to predict the impact of elevated atmospheric CO_2 concentration on crop productivity. Ann Bot 112:465–475

Yin X, Struik PC (2007) Crop systems biology: an approach to connect functional genomics with crop modelling. In: Spiertz JHJ, Struik PC, van Laar HH (eds) Scale and complexity in plant systems research: gene-plant-crop relations. Springer, Dordrecht, pp 61–71

Yin X, Struik PC (2008) Applying modelling experiences from the past to shape crop systems biology: the need to converge crop physiology and functional genomics. New Phytol 179:629–642

Yin X, Struik PC (2009) C_3 and C_4 photosynthesis models: an overview from the perspective of crop modelling. NJAS Wageningen J Life Sci 57:27–38

Yin X, Struik PC (2010) Modelling the crop: from system dynamics to systems biology. J Exp Bot 61:2171–2183

Zhu XG, Zhang GL, Tholen D, Wang Y, Xin CP, Song QF (2011) The next generation models for crops and agro-ecosystems. Sci China Inf Sci 54:589–597