Chapter 9
Modelling QTL-Trait-Crop Relationships: Past Experiences and Future Prospects

Xinyou Yin, Paul C. Struik, Junfei Gu, and Huaqi Wang

Abstract Ecophysiological crop models have long been used to understand crop responses to environmental factors and to crop management practices, by integrating quantitative functional relationships for various physiological processes. In view of the potential added value of robust crop modelling to classical quantitative genetics, model-input parameters are increasingly considered to represent 'genetic coefficients', which are environment-independent and amenable to selection. Likewise, modern molecular genetics can enhance applications of ecophysiological modelling in breeding design by elucidating the genetic basis of model-input parameters. A number of case studies, in which the effects of quantitative trait loci (QTL) have been incorporated into existing ecophysiological models to replace model-input parameters, have shown promise of using these QTL-based models in analysing genotype-phenotype relationships of more complex crop traits. In this chapter, we will review recent research achievements and express our opinions on perspectives for QTL-based modelling of genotype-by-environment interactions and even epistasis on complex traits at crop level.

9.1 Introduction

A major challenge in field- and greenhouse-crop production today is breeding for genotypes and realizing their potential in given (often stressful) environments to produce sufficient high-quality products while maintaining the sustainability of production systems and resource use. This goal can be achieved via creating

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

J. Gu
Key Laboratory of Crop Genetics and Physiology of Jiangsu Province, Yangzhou University, Yangzhou, Jiangsu, China

H. Wang
Plant Breeding and Genetics, China Agricultural University, 100193 Beijing, China

phenotypes of complex traits at the level of the crop – the community of mutually interacting plants, usually of the same species. A thorough insight into gene-trait-crop relationships is therefore crucial. Currently, there is an increasing recognition amongst geneticists and breeders (e.g., Tuberosa and Salvi 2006; Dwivedi et al. 2007; Langridge and Fleury 2011; Messina et al. 2011) and physiologists (e.g., Chenu et al. 2009; Zhu et al. 2011) of immediate need for physiological and computational tools to assist breeders in more effectively analysing, interpreting, translating, and integrating the outputs from high-throughput genomics research, and to help resolving genotype-by-environment interactions (G×E) efficiently and selecting the best technology interventions and associated breeding systems for their target traits and target environments.

Actually, decades ago, process-based physiological models of crop growth have already been suggested to be useful tools in supporting breeding (e.g., Loomis et al. 1979; Spitters and Schapendonk 1990). These models quantify causality between relevant physiological processes and responses of these processes to environmental variables, and, therefore, allow predictions of crop yields not restricted to the environments in which the model parameters have been derived. Crop models require environmental inputs (i.e., weather variables and management options) and physiological inputs. The latter inputs are used as model parameters for characterizing genotypic differences. These genotype-specific parameters, acronymised as GSP by Boote et al. (see Chap. 8 of this book), are also referred to as '*genetic coefficients*' (White and Hoogenboom 1996; Mavromatis et al. 2001) or 'model-input traits' (Yin et al. 2000a), implying that model-input parameters might be (at least partly) under genetic control. As model parameters can reflect certain genetic characteristics, crop modelling has long been considered a useful computational tool to assist breeding (Loomis et al. 1979; Spitters and Schapendonk 1990; Boote et al. 2001). Shorter et al. (1991) already proposed collaborative efforts between breeders, physiologists and modellers, using models as a framework to integrate physiology with breeding.

Given the common experience that crop models based on physiologically sound mechanisms can quantify and integrate responses of crop yield to both genetic and environmental factors, crop physiologists, breeders and modellers have explored the potential of using crop models in various aspects of breeding. These activities include: (1) identifying main yield-determining traits, both under poor and conducive environments for crop growth (Semenov and Halford 2009; Yin et al. 2000b; Heuvelink et al. 2007), (2) defining optimum selection environments in order to maximize selection progress (Aggarwal et al. 1997), (3) optimizing single trait values (Boote and Tollenaar 1994; Setter et al. 1995; Yin et al. 1997), (4) designing ideotypes in which trade-offs between conflicting crop traits are properly evaluated (Spitters and Schapendonk 1990; Penning de Vries 1991; Dingkuhn et al. 1993; Kropff et al. 1995; Haverkort and Kooman 1997), and (5) assisting multi-location testing (Dua et al. 1990) and explaining G×E (Mavromatis et al. 2001; van Eeuwijk et al. 2005; Bertin et al. 2010). Some of these explorative activities were summarised by Boote et al. (2001).

However, crop physiology has not contributed much to breeding (Jackson et al. 1996). All the above-mentioned studies, based on model simulations, are to give suggestions that breeders may use. Stam (1998) and Koornneef and Stam (2001), from a geneticist's and breeder's point of view, expressed their concerns about this model-based approach that ignores the inheritance of the model-input traits. For example, for designing ideotypes by modelling, it is assumed, either tacitly or explicitly, that these traits can be combined at will in a single genotype. Such an assumption ignores the possible existence of constraints, feedback mechanisms and correlations among the traits. Constraints might be imposed simply by the fact that little genetic variation exists in the genetic material available for selection. Thus, models may not identify those traits for which gain via breeding may be easiest (Jackson et al. 1996). Correlations between the traits, due either to a tight linkage between genes or to a single gene that affects multiple traits (pleiotropy), may seriously hamper the realization of an ideotype (for example, an early-maturing potato cultivar with high resistance against late blight; Visker et al. 2003; Struik 2010). Knowledge of the genetic basis of phenotypic variation, even described in terms of model-input traits, is crucial for a successful breeding programme (Stam 1998). Therefore, understanding the inheritance of the model parameters within the framework is required (Stam 1998). To assist the development of efficient breeding strategies, crop modelling requires quantitative understanding of the inheritance of the model-input parameters.

Largely to that end, there have been a growing number of studies that combine crop modelling with modern genetic approaches. In this chapter, we review recent research experiences on elucidating the QTL-trait-crop relationships by integrating crop systems modelling and genetic QTL mapping. Other roles of crop modelling in genetics and breeding will also be explored. Future prospects in this research line are discussed in the context of assisting crop improvement programmes.

9.2 Complementarity of Crop Modelling and Genetic Mapping

In genetics, complex crop traits can be unravelled into the effects of individual QTL – quantitative trait loci (Paterson et al. 1988), commonly using the materials of a segregating population derived from a bi-parental cross. A common result of QTL analysis of complex crop traits is that QTL expression is usually conditional on the environment and this greatly impedes the application of QTL-mapping information for manipulating complex traits (Stratton 1998).

Crop models can potentially be of help in this respect to better address genotype-phenotype relationships, provided that model-input parameters can be easily measured (Yin et al. 2004) and vary little with environmental conditions (Reymond et al. 2003; Tardieu 2003). Model-input parameters (or 'genetic coefficients') reflect effects of genetic origin in the way that one set of parameters represents one
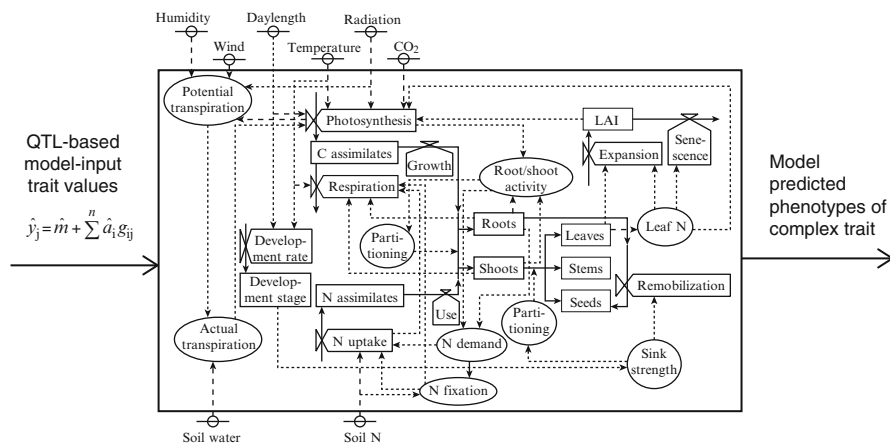
Fig. 9.1 Illustration of QTL(quantitative trait loci)-based crop modelling to predict complex phenotypes. Values of model-input traits were estimated from the effects of their identified QTL to replace the original values of the traits. QTL-based trait values $\hat{y}_j$ were calculated using a simplest genetic model, in which $\hat{m}$ is the estimated intercept, g_{ij} is the genetic (additive effect) predictor of the i-th QTL genotype for the j-th individual of the mapping population, $\hat{a}_i$ is the estimated additive effect of the i-th QTL on the trait ($i = 1, 2, \ldots, n$). This figure predominantly illustrates the structure of a typical crop model using the Forrester's symbols. In such a crop growth model, climatic factors (e.g., radiation, temperature, …), soil conditions, and crop parameters are inputs, and model simulation gives output of complex traits such as crop yield

genotype (Tardieu 2003). Hence, the models manifest that the crop phenotype is achieved through nonlinear interactive and ontogenetic responses of component processes to multiple environmental factors. Such an approach has added value to classical genetics, since geneticists often ignore or overlook competition, density, nutrient supply, morphology, physiology and plasticity, lumping such matters vaguely under the 'G×E' term or introducing simple response functions in their statistical models (e.g., van Eeuwijk et al. 2005).

To enhance applications of ecophysiological modelling in genetic analysis and breeding, understanding the genetic basis of model-input parameters is essential. Yin et al. (1999a, b) first showed that the QTL approach can be applied to model-input parameters to elucidate their inheritance. Such attempts have resulted in an integrated approach of so-called 'QTL-based ecophysiological modelling' (Fig. 9.1), which links crop modelling with genetics, focusing on the G×E problem and genotype-phenotype relationships. The QTL-based models can be used to predict performance of any genotype in any environment.

This approach of QTL-based modelling was first illustrated to predict a very complex trait – the grain yield of barley (*Hordeum vulgare*) by Yin et al. (1999a, b, 2000a). The same approach for QTL-based modelling analyses was applied to crop traits such as leaf elongation rate in maize (Reymond et al. 2003) and flowering time in barley (Yin et al. 2005), rice (Nakagawa et al. 2005) and *Brassica oleracea* (Uptmoor et al. 2008, 2012), fruit quality in peach (Quilot et al. 2004, 2005) and

tomato (Bertin et al. 2010; Prudent et al. 2011), maize kernel number (Amelong et al. 2015), and Arabidopsis stomatal conductance (Reuning et al. 2015), all relatively simple traits with well-defined influences of some dominant environmental factors (such as vapour pressure deficit, soil moisture content, temperature and photoperiod). In the domain of morphological traits, the phenotypic effects of QTL for culm length, grain number, and grain size have been simulated using morphologically explicit models in barley (Buck-Sorlin 2002) and in rice (Xu et al. 2011; see Chap. 2 of this book by Xu and Buck-Sorlin). These studies on relatively simple developmental, morphology-related or growth traits demonstrate that the approach can unravel G×E, and highlight the potential to analyse more complex traits manifested through season-long growth dynamics (see Gu et al. 2014b).

In short, genetic mapping dissects a quantitative trait into various genetic factors – QTL (Paterson et al. 1988), but it can only predict the trait phenotype in independent new environmental conditions to a limited extent (Stratton 1998). Ecophysiological modelling can reveal how G×E comes about (Tardieu 2003), but it does not consider the genetic basis of model parameters that describe genotypic differences. Combining ecophysiological modelling and genetic mapping can dissect complex traits into component traits, integrate effects of QTL of the component traits over time and space at the whole-crop level, and predict yield performance of various genetic make-ups under different environmental conditions.

9.3 Roles of Modelling in Assisting Genetic Analysis and Breeding

There is *in silico* evidence that this combined modelling and genetic approach can facilitate translating the QTL mapping into more efficient marker-assisted breeding strategies (Hammer et al. 2006). To this end, more accurate crop models would facilitate the improvement of efficiencies of combined model- and marker-assisted breeding. In this section we summarise, in abroad sense, the applications of modelling in support of genetic analysis and breeding programme.

9.3.1 Models Can Support Phenotyping for the QTL Mapping

A pre-requisite of the proper use of phenotypic data for quantitative genetic analysis is that the phenotypic data of the different genotypes should be collected under the same environmental conditions and at the same plant developmental stage. On the other hand, quantitative genetic analysis requires screening of a large population to realize the required genetic resolution based on high power of the analyses. Complicated statistical analyses and experimental designs were often used to remove environmental errors, for example, caused by heterogeneity in the

experimental field. But for highly sensitive traits (such as photosynthesis), microclimate fluctuations could also obscure the genetic effects existing in the population. Ecophysiological models based on solid physiological knowledge could be useful tools to standardize the measurements (Gu et al. 2012a). Using model-based standardization, several QTL related to photosynthesis were found under fluctuating field conditions, and were confirmed in independent greenhouse environments. Ecophysiological models can thus play a role in improving the quality of data on traits that are sensitive to environmental changes.

Another example was reported by Yin et al. (1999a), who mapped specific leaf area (SLA) in a barley recombinant inbred lines population. After adjusting SLA values measured at the same chronological time to values at the same physiological age, the effect on SLA from the *denso* gene was no longer significant. The effect of the *denso* gene detected at the same chronological time was therefore the consequence of its direct effect on flowering time. An ecophysiological model can thus indeed assist QTL analysis by removing either environmental noise or indirect effects from other traits.

Breeders often have a crude method of phenotyping. Modelling can help to upgrade their phenotyping activities. Khan (2012) used several expressions to describe phasic development curves of canopy cover dynamics in potato. Not only the overall area under the curve but also individual model parameters were found to vary among individuals of a mapping population, and the parameters most related to the area under the curve were identified, providing the trait components selectable for improving canopy light interception and biomass yield.

9.3.2 Models Can Dissect Complex Traits into Physiological Components

Physiological modelling can dissect complex traits (e.g., photosynthesis or yield) into physiological component traits. Gu et al. (2012b) used a photosynthesis model to dissect photosynthesis into: (1) stomatal conductance g_s, (2) mesophyll conductance g_m, and (3) electron transport capacity J_{max} and Rubisco carboxylation capacity V_{cmax}. Using the crop growth model GECROS, yield was connected to, and dissected into seven physiological input parameters (Gu et al. 2014b). By dissecting complex traits into physiologically meaningful component traits, it is possible to assess genetic variation for each component trait and evaluate its relative importance by sensitive analyses or regression analyses. For example, genetic variation in light-saturated photosynthesis and transpiration efficiency was found to be mainly associated with variation in g_s and g_m (Gu et al. 2012b). The physiological input trait 'total crop nitrogen uptake at maturity' was found to have the most significant effect on yield (Gu et al. 2014b). Similarly, Prudent et al. (2011) combining an

ecophysiological modelling and QTL analysis, identified key elementary processes and genetic factors underlying tomato fruit sugar concentration. All these results show that the physiological model could be helpful to decide on priority targets for breeding, although possible impact remains to be validated through actual breeding and field testing.

9.3.3 Models Can Integrate and Project Single Organ Level Genetic Variation to Crop Level

Modelling not only can dissect complex traits into physiological relevant components, but can also integrate effects of QTL of the component traits over time and space, and predict complex traits at the whole-crop level of various genetic make-ups under different environmental conditions (Yin and Struik 2010). This could be useful to evaluate the effect of changes in a single trait or single trait-related QTL on a crop, while keeping other traits constant to avoid the confounding effects from other physiological processes, which is not plausible in a 'real' experiment. For example, as stated earlier, improving photosynthesis is generally thought crucial for improving plant production, but often no correlation or even negative correlations between photosynthesis and plant production were observed (Evans and Dunstone 1970; Teng et al. 2004; Zhao et al. 2008; Jahn et al. 2011; Gu et al. 2014b). The reason for this discrepancy could be that plants differed genetically in many respects other than photosynthesis. Hence, Gu et al. (2014a) used the crop model GECROS, and found that the natural genetic variation in leaf photosynthesis within our experimental mapping population would result in equivalent differences in production when scaled up to crop level. The ability of integration and upscaling can also help evaluate impacts of QTL for a specific organ-level trait at crop level in a different environment. Using the CROPGRO-soybean model, Messina et al. (2006) estimated the effects of QTL markers from a set of near-isogenic lines and satisfactorily predicted the variation of yield across five years and eight sites among an independent set of soybean cultivars. Chenu et al. (2009), using the crop model APSIM-Maize, evaluated a QTL accelerating leaf elongation on maize yield. This QTL could cause a yield increase in an environment with water deficit before flowering, but reduced yield under terminal drought stress. This information could be used in breeding for specific environments or for facing the challenges caused by climate change. Most importantly, the feature of integration could allow for designing ideotypes of various genetic make-ups underlying physiological processes. Based on the genetic variation and resulting QTL for each physiological component in photosynthesis, it was shown that the ideotype for leaf-level photosynthesis and transpiration efficiency (TE) could potentially be improved by 17.0 % and 25.1 %, respectively (Gu et al. 2012b).

9.3.4 Ecophysiological Model May Help to Resolve Genetic Complexities

A simple genetic model can be assumed for QTL analysis of the component traits, but more sophisticated genetic control (G×E, and epistasis, i.e., the interaction between genes) on the complex trait per se can be manifested when QTL-based parameter values are fed-back to the ecophysiological model. As discussed earlier, use of ecophysiological models to predict and interpret G×E has been widely recognised and exemplified (Reymond et al. 2003; Yin et al. 2005; Messina et al. 2006; see Chaps. 3, 4, 5, 6, 7, and 8 of this book). The use of the models to interpret epistasis is less recognised. Epistasis is often found for phenotypes that are achieved through interactive and interrelated metabolic and ontogenetic pathways (Lee 1995). It might be reduced or even disappear if input traits of a model that accounts for interrelations among relevant processes are subjected to analysis. Such possibility agrees with the awareness of geneticists that epistasis can often be removed by a physiologically based scaling of trait values (Kearsey and Pooni 1996). For example, crop yield is analysed in agronomy as the product of several yield component traits; independent QTL on various yield components must exhibit an epistatic effect on yield (Yin et al. 2002). Chapman et al. (2003) used the crop model APSIM to generate a state space of genotype performance based on 15 genes controlling 4 traits and then search this space for selection. They showed complex epistatic and G×E effects were generated for yield even though gene action at the trait level had been defined as simple additive effects. Similarly, White and Hoogenboom (1996), Messina et al. (2006), and White et al. (2008) used simple linear additive models to regress model-input parameters against several known gene loci across cultivars or genotypes, implicitly modelling the epistatic effects of these genes on the aggregated traits such as yield or days to flowering. It should be acknowledged that use of crop models to resolve epistasis in real experimental populations may be a more difficult task than to resolve G×E, and for the required accuracy crop models should evolve into crop systems biology models (see Chap. 1 of this book by Baldazzi et al.).

9.3.5 QTL-Based Modelling Can Quantify Constraints in Breeding

Model simulation could inspire breeders. However, Stam (1998) and Koornneef and Stam (2001), from a geneticist's perspective, expressed their concerns that the ignorance of the inheritance of the model-input parameters is a major constraint for breeders to adopt the results of model-based approaches. Often in ideotype design by modelling, modellers implicitly assumed that plant traits can be combined at will

into a single genotype. As stated earlier, such an unrealistic practice ignores the possible existence of constraints, feedback mechanisms and correlations among traits. By integrating crop modelling with genetics – QTL-based modelling, it is possible to evaluate constraints in breeding either due to limited genetic variation or to correlations. Gu et al. (2012b) showed trade-offs between improving photosynthesis and TE either due to tight linkage or to pleiotropic effects of QTL related to g_m and g_s. If the linkage between g_m and g_s, or co-location of QTL of g_m and g_s could be broken, the virtual ideotype could have both improved photosynthesis and TE. The quantitative importance of breaking this linkage could be used together with insights of geneticists about chances of success in guiding decisions in breeding programmes, thus strengthening the scientific basis for designing breeding activities.

9.3.6   QTL-Based Modelling Can Assist Marker-Assisted Selection

Marker-assisted selection (MAS), combined with conventional breeding approaches, has been used to effectively integrate major genes or QTL with large effect into widely grown varieties (Jena and Mackill 2008). The use of cost-effective DNA markers and a MAS strategy will provide opportunities for breeders to develop high-yielding, stress-tolerant, and better quality rice cultivars. For example, pyramiding different resistance genes using MAS provided opportunities to breeders to develop broad-spectrum resistance against diseases and insects (Huang et al. 1997). An example of the latter approach is the insertion of cassettes of up to four resistance genes from wild potato species into existing cultivars using cisgenesis techniques to make these existing cultivars resistant to late blight (Haverkort et al. 2009). By stacking several resistance genes, the resistance cannot be broken easily by the causal agent *Phytophthora infestans*, certainly not when this approach is combined with a well-designed resistance management strategy (Haverkort et al. 2009).

Gu et al. (2014b) also showed that the existing GECROS model can be a useful tool to enhance marker-assisted breeding through a model-based ideotype design. Using the principles for QTL-based modelling as defined earlier (Yin et al. 2000a, 2004, 2005), marker-based crop modelling was performed to rank the markers identified for various yield-determining physiological traits that are input parameters of GECROS (Table 9.1). It was found that the relative importance of markers differed markedly between well-watered and drought-stressed environments (the correlation coefficient in the rank between the two environments was 0.09; $P > 0.10$). Such an analysis detected markers that breeders can prioritize in their MAS programmes for specific environments. Gu et al. (2014b) showed that compared with identification of markers through multiple regression for yield per se, the model-based approach identified additional QTL and could be complementary to the analysis of yield per se.

Table 9.1 Percentage of the phenotypic variation in yield among rice introgression lines (ILs) (R^2) accounted for by different sets of simulations using the marker-based version of the GECROS model, when markers were fixed one at a time to calculate different sets of marker-based parameter inputs for GECROS. Marker positions are denoted as 'Chr_cM', that is, Chromosome_centiMorgan, as identified during QTL-analysis (Based on Gu et al. (2014b))

Fixed marker		Well-watered		Drought-stressed	
Chr_cM	Name	R^2 (%)	Rank	R^2 (%)	Rank
1_9.5	RM8068	51.6	11	42.6	13
1_25.4	RM8145	53.9	18	41.4	10
1_98.1	RM306	51.6	11	39.5	6
1_124.8	RM1152	50.9	9	44.6	18
2_92.5	RM475[a]	46.2	3	37.7	5
2_110.9	RM1367	51.7	14	45.5	19
2_139.3	RM8030[a]	34.2	1	40.9	9
3_79.1	RM251	47.9	5	46.2	20
3_108.4	RM338[b]	52.6	17	29.8	1
4_25.5	RM518	59.4	20	44.1	17
4_123.8	RM2799	51.8	16	40.2	7
5_20.6	RM7302[b]	51.6	11	33.2	2
7_43.5	RM432	50.7	8	36.9	4
7_47	RM11	51.7	14	43.8	16
7_81.05	RM3753	49.4	7	41.9	11
8_83.7	RM284[a]	45.7	2	42.6	13
9_0.8	RM5799	48.3	6	42.6	13
9_64.4	RM410[a, b]	47.3	4	35.9	3
10_87.1	RM294A	51.1	10	40.5	8
12_61.6	RM1261	53.9	19	42.1	12
Baseline simulation		51.6		42.6	

The baseline simulation gives the R^2 values for the simulation, in which no marker was fixed, i.e., IL-specific allelic values (−1 or 1) were used for all markers in calculating marker-based inputs; for other sets of simulations, markers were fixed one at a time, in which all ILs were assumed to carry an identical allele (i.e., 0) at the locus of the considered marker in calculating marker-based inputs

[a]These markers were also identified for yield per se under well-watered conditions
[b]These markers were also identified for yield per se under drought-stressed conditions

9.4 Past Experiences in Integrating Ecophysiological Modelling and Genetic Mapping

The main purpose of practising QTL-based ecophysiological modelling on the basis of using a mapping population is to predict genotypic impact on phenotypes, in contrast to conventional crop modelling which usually aims to predict the impact of environmental variables on crop productivity. The following contains summaries of current experiences.

9.4.1   Models Generally Perform Better in Simulating Phenotypic Differences Caused by Environmental Variation than by Genotypic Differences

This may not be surprising given that when individuals of mapping population are phenotyped in multiple environments, it is common that variance due to environmental differences is much more significant than the variance due to genotypic differences (Yin et al. 2000a, b; Gu et al. 2012b; Khan et al. 2014). Crop models generally perform well in assessing the impact of significant environmental variation due to changes in climatic variables and nutrient availabilities, as most existing crop models were built to predict the impact of environmental variables on crop productivity. It is a challenge to predict the impact of a subtle change in traits among relatively similar lines within a breeding population. Therefore, model's suitability in analysing genotype-to-phenotype relationships in an experimental population needs critical examinations (Parent and Tardieu 2014). The following two aspects deserve particular attentions. First, better modelling of the final spikelet or seed number of cereals under stress conditions is needed as stress sensitivity of this sink-size trait often shows larger genetic variability than that of the source-activity traits. Second, the genetic difference in response to soil environments is currently subject to huge uncertainty, due partly to the lack of sufficient site-specific information about the soil and partly to uncertainties in modelling root growth and soil processes (Gu et al. 2014b; Khan et al. 2014). There is an obvious need for robust algorithms for rooting density and depth for resource capture from the soil and their genotypic variabilities.

9.4.2   Some Model-Input Parameters Do Not Contribute to the Model in Explaining Differences among the Genotypes

The importance of model parameters in contributing to explaining yield differences among individuals of mapping population can be evaluated by fixing them once at a time at their average value (Yin et al. 2000b; Khan et al. 2014). It is expected that the model explained percentage of yield differences will drop if the parameter fixed is important for yield determination. Counter-intuitively, fixing some parameters, which seem to be physiologically important, even increased the explained percentage of phenotypic variation. Identified examples for such parameters are: specific leaf area (SLA), leaf nitrogen content, post-flowering duration for barley (Yin et al. 2000b), vegetative growth period and maximum plant height for potato (Khan et al. 2014). Similarly, when introducing genetic variation of individual biochemical parameters of leaf-photosynthesis into the photosynthesis sub-model of GECROS, the variation of yield accounted for by GECROS decreased significantly for both well-watered and drought-stressed conditions (Gu et al. 2014b). Such model based sensitivity analysis suggests whether or not the model has incorporated right

parameters in explaining yield differences among genotypes in a population. The reasons for the unimportance of those seemingly important parameters in terms of yield physiology in explaining genetic differences among genotypes remain to be elucidated.

9.4.3 Some Model-Input Parameters Are Hard to Measure for the Whole Population, and Heritability Estimates (h^2) of Model-Input Parameters Are Generally Lower than Those of Classical Plant Traits

Some model-input parameters are used in crop models in a tabular form, e.g., coefficients for assimilate partitioning among growing organs in Wageningen SUCROS-family models. Determining values of these coefficients require frequent destructive samplings during growing season, which can be implemented in classical agronomic experiments but are practically infeasible for individual lines of a mapping population. These types of parameters certainly do not allow high throughput measurements, and many of them need many steps to measure. Measurement noise accumulates over steps; some parameters require curve-fitting method to estimate, which again involves some fitting uncertainty/noise (also see Chap. 5 by Luquet et al.). So, the h^2 for measured phenotypic data of these parameters is often lower than for traits relating to classical agronomic, plant size, and architecture traits. This is in analogy to the result of Jahn et al. (2011) that physiological traits such as stomatal conductance (which involve various steps of measurements and calculations) had a lower h^2 than the classical agronomic and morphological traits. As a consequence, the percentage of phenotypic variation explained (r^2) by QTL identified for model parameters is often lower than the r^2 of QTL for classical plant traits if measured in the same experiments (Yin et al. 1999b; Gu et al. 2014b).

9.4.4 The Percentage of Phenotypic Variation of a Complex Trait Accounted for by the QTL-Based Model Is Comparable with, or Slightly Lower Than, That Obtained from the Original Parameter Values

QTL-based model parameters values derived from QTL-statistics can partly remove the noise of phenotypic values of model-input parameters. On the other hand, QTL identified by mapping analysis account for only part of genetic variance of parameters. So, a common feature of these studies is that predictability of QTL-based models is nearly comparable with that of the model using original parameter values (see Fig. 9.2 for an example), as the gain from removing random noise in original
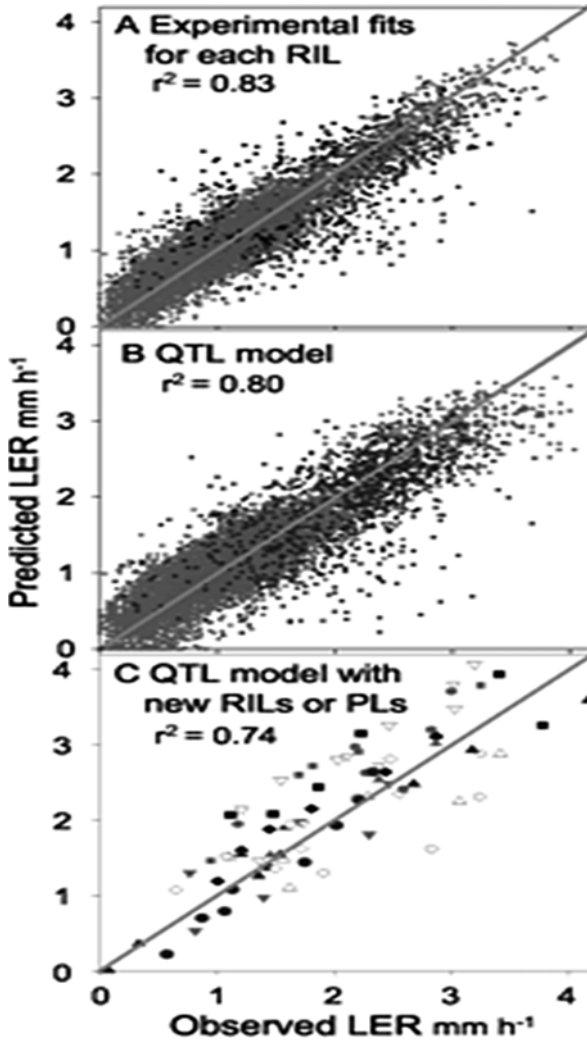
Fig. 9.2 Comparison between observed maize leaf elongation rates LER and those fitted by a simple ecophysiological model for LER (*A*), simulated by the model using QTL-based model input parameters (*B*), and QTL-based model predicted LER for those recombinant inbred lines RILs that were not included for QTL analysis (C) (Redrawn from Reymond et al. 2003)

parameters by QTL-statistics is roughly cancelled out by the loss due to the fact that the identified QTL cannot explain 100 % of the genetic variance of the parameter values (e.g., Yin et al. 2000a, 2005; Reymond et al. 2003; Gu et al. 2014b). However, QTL-based models can predict the performance of genotypes that were not phenotyped for model-input parameters (Fig. 9.2c) as long as marker data at or near QTL are available for these genotypes.

9.4.5 Despite the Outnumbering of QTL for Model-Input Parameters Relative to Those for the Complex Trait per Se, the Percentage of Phenotypic Variation of a Complex Trait Accounted for by Its QTL of Model-Input Parameters Is Lower than That Obtained from Complex-Trait QTL

Current QTL statistics can hardly find more than eight QTL for a quantitative trait to avoid false positives in QTL analysis (Kearsey and Farquhar 1998), although this may greatly depend on the population size. Therefore, rather than looking for QTL for a complex trait itself, determining QTL for underlying component traits might give more information. Indeed, using crop models to dissect a complex trait into its individual components will help to identify more QTL than analysing the complex trait per se (Gu et al. 2014b; Amelong et al. 2015). However, the percentage of phenotypic variation of a complex trait accounted for by its QTL of model-input parameters is lower than that obtained from complex-trait QTL (Prudent et al. 2011; Yin et al. 2000a; Gu et al. 2014b). Low predictability of the models could be the reason for that. However, even when yield is dissected into yield components using a simple arithmetic formula (typically: yield is equal to the product of yield component traits) and the formula perfectly predicts yield variation, the percentage of phenotypic variation of yield accounted for by QTL of its component traits is lower than that obtained from complex-trait QTL (Yin et al. 2002; Fig. 9.3). This suggests that phenotyping of model-input parameters and yield component traits may involve more random noise.

9.4.6 Number of QTL Identified for Model-Input Parameters Based on a Bi-parental Population Is Limited; Most Model-Input Parameters Are Often Affected by the Pleiotropic Effect of 1–2 Major QTL

Dominance of a major QTL is a common phenomenon, presumably due to the contrast between the parents intentionally chosen in making the bi-parental mapping population. Typically, one parent represents a modern cultivar that is currently widely cultivated whereas the other is an old traditional genotype that was probably cultivated before the Green Revolution. This means that one or two major genes are segregating in the population. Very often major genes not only affect major morphological characteristics and yield level, but also have pleiotropic effects on multiple phenological and physiological traits including model-input parameters. This has
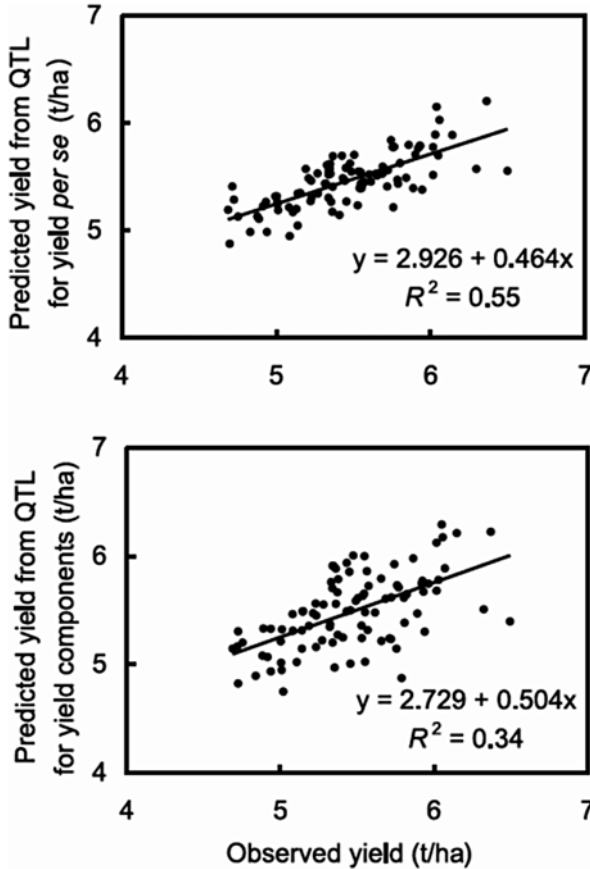
Fig. 9.3 Comparison between observed values of grain yield and those predicted from quantitative trait loci (QTL) identified for yield itself, and between observed values of grain yield and those predicted from QTL identified for its three component traits: spikes per m^2, number of kernels per spike and 1000-kernel weight (Redrawn from Yin et al. 2002)

been shown by, for example, the *denso* gene in barley (Yin et al. 1999a, b), *rht* genes in wheat (Baenziger et al. 2004; Laperche et al. 2006; White 2006), the maturity-class gene on chromosome V in potato (Khan 2012), and the RM410 locus on chromosome 9 in rice (Gu et al. 2012a). Of these, the *denso* gene in barley is particularly pleiotropic and its dominant effect is ubiquitous, not only on plant height, yield and yield components (Yin et al. 1999b), but also for flowering parameters (Yin et al. 2005) and traits like SLA (Yin et al. 1999a) and nitrogen use efficiency (Kindu et al. 2014).

9.4.7 Medium-Size Population Is the Best Option That Combines Feasibility and Robustness in Integrated Ecophysiological Modelling and Genetic Mapping Studies

From a statistical point of view, the larger is the population, the more robust is QTL mapping (Vales et al. 2005). Going for a large population size is generally not feasible because most individual input parameters in existing crop models do not allow high-throughput phenotyping. Because of the cost and/or time needed, researchers often went for selective phenotyping, and some were even pushed to phenotype only 46 individuals as a subset of a population to identify QTL for model-input parameters (Uptmoor et al. 2012), thereby, greatly sacrificing the statistical power of QTL detection. In that sense, crop models should be improved in a way that most parameters would be measurable in phenotyping facilities (Yin et al. 2004; Parent and Tardieu 2014). Before such a model becomes available, a medium-size population consisting of ca 100 individuals as a comprise of phenotyping feasibility and QTL-detection robustness, may be the best option, if model input-parameters can be measured with the currently available methods. If model input-parameters are hard to measure, it is better to use an introgression line (IL) population other than populations like recombinant inbred lines (RILs), doubled-haploid lines (DHs), as ILs differ in a lower number of loci. For example, leaf photosynthesis is commonly measured by gas exchange that does not allow high-throughput phenotyping. In a study where the entire light- and CO_2-response curves needed to be phenotyped via gas exchange, Gu et al. (2012b) chose 13 ILs (including parents) and did succeed to localize the genomic regions for seven parameters of a biochemical photosynthesis model.

9.5 Future Prospects

9.5.1 Understanding Physiological Basis of QTL and Genetic Variation

From a physiologists' point of view, a logic step following the mapping of genetic basis (i.e., QTL analysis) of a physiological trait is to elucidate the deeper-level physiological basis of the detected QTL underlying its genetic variation. Few studies have investigated the physiological basis of QTL underlying genetic variation of quantitative traits. The physiological basis of QTL may best be elucidated with physiological models that dissect complex traits into individual component traits. This was recently reported for leaf photosynthesis. QTL for light saturated leaf photosynthesis (A_{max}) and other related traits were first identified using an introgression line population (Gu et al. 2012a). To elucidate the physiological basis of these

QTL, combined gas exchange and chlorophyll fluorescence data were collected for entire CO_2 and light response curves of leaf photosynthesis (A), with which biochemical and physiological parameters of a combined conductance-biochemical photosynthesis model were estimated. Because measuring entire response curves is time consuming and does not allow high throughput, 13 lines (including the two parents) were carefully selected as representatives of the population, based on the QTL for leaf photosynthesis reported by Gu et al. (2012a). The curves were assessed at two stages (flowering and grain filling) for plants grown under moderate drought and well-watered conditions (Gu et al. 2012b). Using these curves, photosynthesis was then quantitatively dissected into three different physiologically relevant component traits: (1) stomatal conductance (g_s), (2) mesophyll conductance (g_m), and (3) biochemical efficiency and capacity. Although the effects of development stage and water supply on photosynthesis were predominant, significant genetic variation in the three mentioned component traits was found. Genomic regions of the variation of these biochemical parameters of photosynthesis were localised (Fig. 9.4). Genetic variation in A_{max} and TE (transpiration efficiency) was mainly caused by variation in g_s and g_m, which suggests more efforts should be focused on g_s and g_m in breeding programmes for improving photosynthesis and TE. Gu et al. (2012b) showed that relationships between these photosynthetic parameters and leaf nitrogen or dry matter per unit area, which were previously found across environmental treatments, were also valid for variation across genotypes. Therefore, they speculated that variation in photosynthesis due to environmental conditions and the variation in photosynthesis due to genetic variation within the same environment may share common physiological mechanisms.

Gu et al. (2012b) next used the model to evaluate the potential of utilizing the genetic variation in these components for improving photosynthesis (A) and transpiration efficiency (TE). Based on the genetic variation of physiological components underlying A and TE, ideotypes were designed by combining alleles positively influencing different components of photosynthesis. Model calculations showed that these ideotypes can potentially improve photosynthesis and TE significantly, compared with the best genotype of the 13 lines investigated. It was shown that if the tight link between g_m and g_s could be broken, both photosynthesis and TE could be improved simultaneously, despite the common negative correlation between A and TE (e.g., Condon et al. 2004). This result would be especially interesting for breeding for semi-arid environments.

The importance of mesophyll conductance in improving leaf photosynthesis has also been identified for materials of other genetic backgrounds in rice (Adachi et al. 2013). Adachi et al. (2014) further indicated that high leaf nitrogen content and high hydraulic conductivity are two additional physiological mechanisms contributing to high leaf photosynthesis of their near-isogenic lines (NILs), which differ from the recipient parent in only one or two introgression regions of previously mapped QTL and therefore best suit for elucidation of physiological basis for individual QTL. Similar results have been found for the genetic variation in leaf photosynthesis across cultivars in rice (Taylaran et al. 2011; Lauteri et al. 2014) and wheat (Jahan et al. 2014).
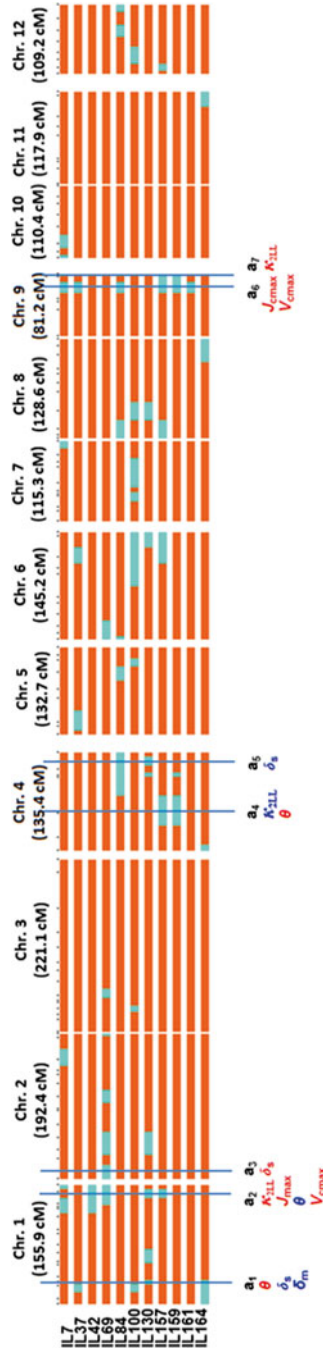
Fig. 9.4 Graphical genotypes for the 11 introgression lines which were used in the previous analysis of Gu et al. (2012b) for assigning the genome regions (QTL) to six photosynthesis parameters in a biochemical model: κ_{2LL} (efficiency of converting incident light into linear electron transport under limiting light), J_{max} (maximum electron transport rate under saturating light), θ (convexity factor of light response curve of electron transport), δ_m (parameter associated with mesophyll conductance), δ_s (parameter associated with stomatal conductance), and V_{cmax} (maximum capacity of Rubisco carboxylation). The length of each linkage group is shown in centiMorgan (cM). The *light-blue* regions indicate the introgression regions from the donor parent 'Haogelao'; the *ochre* backgrounds indicate the homozygous regions from the recurrent parent 'Shennong265'. Parameters, on which genome alleles from 'Haogelao' have positive effects and negative effects, are shown in *red* and *blue colours*, respectively (Redrawn from Gu et al. (2012b))

For more complex traits rather than leaf photosynthesis, Gu et al. (2014b) used the crop model GECROS to dissect yield into seven phenological, morphological and physiological parameters. It was found that nitrogen uptake, grain nitrogen concentration and pre- and post-flowering durations are important, whereas leaf photosynthesis was surprisingly not important, in explaining yield differences among the individual lines within a genetic population. Using rice NILs that harbour one or two spikelet-number QTL, Ohsumi et al. (2011) found that the NILs having increased spikelet number per panicle did not greatly increase grain yield because of compensation between different yield components. They also showed that the slight yield advantage of the NILs harbouring double QTL relative to other genotypes was associated with higher translocation of carbohydrates from reserves to panicle. These indicate a pleiotropic effect of the spikelet-number QTL on other physiological traits.

We call for more studies on elucidating physiological basis of QTL and pleiotropic effect of the QTL on other physiological traits/processes. Such information will facilitate to improve existing crop models that better capture physiological processes and parameters related to genetic variation of crop yield.

9.5.2   Broadening Genetic Background of the Mapping Population

While the proposed QTL-based modelling approach could potentially deal with G×E, it cannot solve all limiting factors, especially not the non-transferability of information obtained from one cross to another. The non-transferability can be largely due to the possibility that a QTL detected in one cross does simply not segregate in a second cross because the parents of the second cross carry identical alleles at that QTL – the lack of allelic diversity within a mapping population. A gene 'important' for physiologists or modellers might be useless for geneticists or breeders because if the gene is physiologically crucial, its variation will have been strongly reduced over generations of breeding (Prioul et al. 1997); so QTL will hardly be detected at such a gene locus. In this context, the approach as practised for a bi-parental cross should be extended in future although such extended studies with a broader genetic background can be most feasibly applied to simple traits that can be scored by high-throughput phenotyping.

9.5.2.1   Use of Multiple Mapping Population

With the crop model GECROS, Gu et al. (2014b) used the marker-based parameter values derived from a population of 94 introgression lines to simulate yield variation among 251 recombinant inbred lines of the same parents in rice. More directly working with multiple populations, Welcker et al. (2011) have compared the genetic

architectures of the sensitivities of maize (*Zea mays*) leaf elongation rate with evaporative demand and soil water deficit as quantified in a simple ecophysiological model. The former was measured via the response to leaf-to-air vapour pressure deficit in well-watered plants, the latter via the response to soil water potential in the absence of evaporative demand. Genetic analyses of each sensitivity were performed over 21 independent experiments with (1) three mapping populations, with temperate or tropical materials, (2) one population resulting from the introgression of a tropical drought-tolerant line in a temperate line, and (3) two introgression libraries genetically independent from mapping populations. A very large genetic variability was observed for both sensitivities. Some lines maintained leaf elongation at very high evaporative demand or water deficit, while others stopped elongation in mild conditions. A complex architecture arose from analyses of mapping populations, with 19 major meta-QTL involving strong effects and/or more than one mapping population. A total of 68 % of those QTL affected sensitivities to both evaporative demand and soil water deficit. In introgressed lines, 73 % of the tested genomic regions affected both sensitivities. They demonstrated that hydraulic processes, which drive the response to evaporative demand, also have a large contribution to the genetic variability of plant growth under water deficit in a wide range of genetic material comprising of multiple populations.

On the genetic side, geneticists are trying to improve QTL mapping resolution with several generations of intercrossing when establishing the RIL population, e.g. advanced intercross RILs. Meanwhile allelic diversity within a mapping population can be increased by intercrossing multiple genetically diverse genotypes before establishing the RILs, e.g., MAGIC – the Multi-parent Advanced Generation Inter-Cross (Huang et al. 2011).

9.5.2.2   Genome-Wide Association Study

Virk et al. (1996) empirically showed that quantitative variation of many agronomic traits in the rice germplasm is associated with allelic variation of DNA markers, indicating that marker-trait associations not only may be present in segregating populations, but can also be manifest across a germplasm collection of a crop species. Later studies more systematically demonstrated that the bi-parental analysis could also be extended by using genome-wide association study (*GWAS*) based on the linkage disequilibrium mapping, in which association between genotypes and phenotypes is scrutinized over a large germplasm collection (e.g., Remington et al. 2001). This development in association genetics may enhance opportunities for gene-based physiological modelling, especially with development of genome-wide surveys of variation using high-throughput genotyping tools (such as SNP – single-nucleotide polymorphism) across crop germplasm collections (McNally et al. 2009; Huang et al. 2010; Jahn et al. 2011). This development in association genetics may enhance opportunities for gene-based crop modelling. So-called gene-based modelling has empirically been practised by White and Hoogenboom (1996), Messina et al. (2006), White et al. (2008), Zheng et al. (2013), and Bogard et al. (2014), who predicted

flowering and yield traits of crop cultivars via regressing input parameters against binary values of a few relevant candidate genes or markers (also see Chap. 8 of this book by Boote et al.). The SNP-based GWAS can detect many genes and unravel their functions, thereby enabling model-input parameters to be potentially related to many genes. Research on GWAS-based crop modelling is now in the pipeline. However, such an analysis requires considerable attention to population structure and size to satisfy its required statistical power (Hamblin et al. 2011).

9.6   Concluding Remarks

Crop physiology research, traditionally working typically on several distinctive genotypes, has not contributed much to breeding so far. Furthermore, there has been an imbalance in gaining insight and data as geneticists and physiologists seemed to do the things separately. However, the evidence reviewed in this chapter suggests that there is now an opportunity to do a better job because we have more knowledge about QTL (or gene)-function relationships and we have better analytical tools such as QTL (or gene)-based models, dealing with relationships in a genetic population. An integrated QTL-based modelling also provides a common platform for physiologists and geneticists of working all in a synchronous and balanced way, thereby being much more effective in terms of resource use and synergy between approaches. Growing studies on functional genomics and molecular biology will increasingly enable the elucidation of the molecular genetic basis of agronomically and physiologically relevant traits for crop improvement. In the meantime, high-throughput facilities to phenotype a large population for various crop traits, sometimes with high-resolution, are increasingly becoming available. Future crop models should face this unprecedented opportunity. On one hand, model-input parameters should be designed either to be close to those traits breeders, geneticists and biologists commonly score or to be easily measured by modern high-throughput phenotyping facilities, as the optimisation procedure that current crop models often rely on to estimates their parameters (see Chap. 5 of this book by Luquet et al.) may involve another round of uncertainty. On the other hand, model structure and algorithms have to be upgraded, which has been stressed in various preceding chapters of this book and will be further elaborated in the next chapter.

References

Adachi S, Baptista L, Sueyoshi T, Murata K, Yamamoto T, Ebitani T, Ookawa T, Hirasawa T (2014) Introgression of two chromosome regions for leaf photosynthesis from an *indica* rice into background of a *japonica* rice. J Exp Bot 65:2049–2056

Adachi S, Nakae T, Uchida M, Soda K, Takai T, Oi T, Yamamoto T, Ookawa T, Miyake H, Yano M, Hirasawa T (2013) The mesophyll anatomy enhancing CO_2 diffusion is a key trait for improving rice photosynthesis. J Exp Bot 64:1061–1072

Aggarwal PK, Kropff MJ, Cassman KG, ten Berge HFM (1997) Simulating genotypic strategies for increasing rice yield potential in irrigated, tropical environments. Field Crops Res 51:5–17

Amelong A, Gambin B, Severini AD, Borrás L (2015) Predicting maize kernel number using QTL information. Field Crops Res 172:119–131

Baenziger PS, McMaster GS, Wilhelm WW, Weiss A, Hays CJ (2004) Putting genes into genetic coefficients. Field Crops Res 90:133–143

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. J Exp Bot 61:955–967

Bogard M, Ravel C, Paux E, Bordes J, Balfourier F, Chapman SC, Le Gouis J, Allard V (2014) Predictions of heading date in bread wheat (*Triticum aestivum* L.) using QTL-based parameters of an ecophysiological model. J Exp Bot 65:5849–5865

Boote KJ, Tollenaar M (1994) Modeling genetic yield potential. In: Boote KJ et al (eds) Physiology and determination of crop yield. ASA/CSSA/SSSA, Madison, pp 533–565

Boote KJ, Kropff MJ, Bindraban PS (2001) Physiology and modelling of traits in plants: implications for genetic improvement. Agric Syst 70:395–420

Buck-Sorlin GH (2002) The search for QTL in barley (*Hordeum vulgare* L.) using a new mapping population. Cell Mol Biol Lett 7:523–535

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

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

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD (2004) Breeding for high water-use efficiency. J Exp Bot 55:2447–2460

Dingkuhn M, de Vries Penning FWT, Miezan KM (1993) Improvement of rice plant type concepts: systems research enables interaction of physiology and breeding. In: de Vries Penning FWT et al (eds) Systems approaches for agricultural development. Kluwer, Dordrecht, pp 19–35

Dua AB, de Vries Penning FWT, Seshu DV (1990) Simulation to support evaluation of the production potential of rice varieties in tropical climates. Trans Am Soc Agric Eng 33:1185–1194

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

Evans LT, Dunstone RL (1970) Some physiological aspects of evolution in wheat. Aust J Biol Sci 23:725–742

Gu J, Yin X, Struik PC, Stomph TJ, Wang H (2012a) Using chromosome introgression lines to map quantitative trait loci for photosynthesis parameters in rice (*Oryza sativa* L.) leaves under drought and well watered field conditions. J Exp Bot 63:455–469

Gu J, Yin X, Stomph TJ, Wang H, Struik PC (2012b) Physiological basis of genetic variation in leaf photosynthesis among rice (*Oryza sativa* L.) introgression lines under drought and well-watered conditions. J Exp Bot 63:5137–5153

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

Gu J, Yin X, Zhang C, Wang H, Struik PC (2014b) Linking ecophysiological modelling with quantitative genetics to support marker-assisted crop design for improved rice (*Oryza sativa* L.) yields under drought stress. Ann Bot 114:499–511

Hamblin MT, Buckler ES, Jannink J-L (2011) Population genetics of genomics-based crop improvement methods. Trends Plant Sci 27:98–106

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

Haverkort AJ, Kooman PL (1997) The use of systems analysis and modeling of growth and development in potato ideotyping under conditions affecting yields. Euphytica 94:191–200

Haverkort AJ, Struik PC, Visser RGF, Jacobsen E (2009) Applied biotechnology to combat late blight in potato caused by *Phytophthora infestans*. Potato Res 52:249–264

Heuvelink E, Marcelis LFM, Bakker MJ, van der Ploeg A (2007) Use of crop growth models to evaluate physiological traits in genotypes of horticultural crops. In: Spiertz JHJ, Struik PC, van Laar HH (eds) Scale and complexity in plant systems research: gene-plant-crop relations. Springer-Verlag, Dordrecht, pp 223–233

Huang N, Angeles E, Domingo J, Magpantay G, Singh S, Zhang G, Kumaravadivel N, Bennett J, Khush G (1997) Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and PCR. Theor Appl Genet 95:313–320

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z, Buckler ES, Qian Q, Zhang QF, Li J, Han B (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42:961–967

Huang X, Paulo MJ, Boer M, Effgen S, Keizer P, Koornneef M, van Eeuwijk FA (2011) Analysis of natural allelic variation in Arabidopsis using a multiparent recombinant inbred line population. Proc Natl Acad Sci U S A 108:4488–4493

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

Jahan E, Amthor JS, Farquhar GD, Trethowan R, Barbour MM (2014) Variation in mesophyll conductance among Australian wheat genotypes. Funct Plant Biol 41:568–580

Jahn CE, Mckay JK, Mauleon R, Stephens J, McNally KL, Bush DR, Leung H, Leach JE (2011) Genetic variation in biomass traits among 20 diverse rice varieties. Plant Physiol 155:157–168

Jena K, Mackill D (2008) Molecular markers and their use in marker-assisted selection in rice. Crop Sci 48:1266–1276

Kearsey M, Farquhar AGL (1998) QTL analysis in plants; where are we now? Heredity 80:137–142

Kearsey MJ, Pooni HS (1996) The genetical analysis of quantitative traits. Chapman & Hall, London

Khan MS (2012) Assessing genetic variation in growth and development of potato. Ph.D. thesis, Wageningen University, The Netherlands

Khan MS, Yin X, van der Putten PEL, Struik PC (2014) An ecophysiological model analysis of yield differences within a set of contrasting cultivars and an F1 segregating population of potato (*Solanum tuberosum* L.) grown under diverse environments. Ecol Model 290:146–154

Kindu GA, Tang J, Yin X, Struik PC (2014) Quantitative trait locus analysis of nitrogen use efficiency in barley (*Hordeum vulgare* L.). Euphytica 199:207–221

Koornneef M, Stam P (2001) Changing paradigms in plant breeding. Plant Physiol 125:156–159

Kropff MJ, Haverkort AJ, Aggarwal PK, Kooman PL (1995) Using systems approaches to design and evaluate ideotypes for specific environments. In: Bouma J et al. (eds) Eco-regional approaches for sustainable land use and food production. Kluwer, Dordrecht, pp 417–435

Langridge P, Fleury D (2011) Making the most of 'omics' for crop breeding. Trends Plant Sci 29:33–40

Laperche A, Devienne-Baret F, Maury O, Le Gouis J, Ney B (2006) A simplified conceptual model of carbon/nitrogen functioning for QTL analysis of winter wheat adaptation to nitrogen deficiency. Theor Appl Genet 113:1131–1146

Lauteri M, Haworth M, Serraj R, Monteverdi MC, Centritto M (2014) Photosynthetic diffusional constraints affect yield in drought stressed rice cultivars during flowering. PLoS One 9(10):e109054

Lee M (1995) DNA markers and plant breeding programs. Adv Agron 55:265–344

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

Mavromatis T, Boote KJ, Jones JW, Irmak A, Shinde D, Hoogenboom G (2001) Developing genetic coefficients for crop simulation models with data from crop performance trials. Crop Sci 41:40–51

McNally KL, Childs KL, Bohnert R et al (2009) Genome wide SNP variation reveals relationships among landraces and modern varieties of rice. Proc Natl Acad Sci U S A 106:12273–12278

Messina CD, Jones JW, Boote KJ, Vallejos CE (2006) A gene-based model to simulate soybean development and yield responses to environment. Crop Sci 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. J Exp Bot 62:855–868

Nakagawa H, Yamagishi J, Miyamoto N, Motoyamo M, Yano M, Nemoto K (2005) Flowering response of rice to photoperiod and temperature: a QTL analysis using a phenological model. Theor Appl Genet 110:778–786

Ohsumi A, Takai T, Ida M, Yamamoto T, Arai-Sanoh Y, Yano M, Ando T, Kondo M (2011) Evaluation of yield performance in rice near-isogenic lines with increased spikelet number. Field Crops Res 120:68–75

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

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Penning de Vries FWT (1991) Improving yields: designing and testing VHYVs. In: FWT Penning de Vries et al. (eds) Systems simulations at IRRI. IRRI research paper 151, IRRI, Los Baños, pp 13–19

Prioul JL, Quarrie S, Causse M, de Vienne D (1997) Dissecting complex physiological functions through the use of molecular quantitative genetics. J Exp Bot 48:1151–1163

Prudent M, Lecomte A, Bouchet J-P, Bertin N, Causse M, Génard M (2011) Combining ecophysiological modelling and quantitative trait locus analysis to identify key elementary processes underlying tomato fruit sugar concentration. J Exp Bot 62:907–919

Quilot B, Génard M, Kervella J, Lescourret F (2004) Analysis of genotypic variation in fruit flesh total sugar content via an ecophysiological model applied to peach. Theor Appl Genet 109:440–449

Quilot B, Kervella J, Génard M, Lescourret F (2005) Analysing the genetic control of peach fruit quality through an ecophysiological model combined with a QTL approach. J Exp Bot 56:3083–3092

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES IV (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci U S A 98:11479–11484

Reuning GA, Bauerle WL, Mullen JL, Mckay JK (2015). Combining quantitative trait loci analysis with physiological models to predict genotype-specific transpiration rates. Plant Cell Environ 38:710–717

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 Physiol 131:664–675

Semenov MA, Halford NG (2009) Identifying target traits and molecular mechanisms for wheat breeding under a changing climate. J Exp Bot 60:2791–2804

Setter TL, Conocono EA, Egdane JA, Kropff MJ (1995) Possibility of increasing yield potential of rice by reducing panicle height in the canopy. I. Effects of panicles on light interception and canopy photosynthesis. Aust J Plant Physiol 22:441–451

Shorter R, Lawn RJ, Hammer GL (1991) Improving genotypic adaptation in crops – a role for breeders, physiologists and modelers. Exp Agric 27:155–175

Spitters CJT, Schapendonk AHCM (1990) Evaluation of breeding strategies for drought tolerance in potato by means of crop growth simulation. Plant and Soil 123:193–203

Stam P (1998) Crop physiology, QTL analysis and plant breeding. In: Lambers H et al. (eds) Inherent variation in plant growth: physiological mechanisms and ecological consequences. Backhuys Publishers, Leiden, pp 429–440

Stratton DA (1998) Reaction norm functions and QTL-environment interactions for flowering time in *Arabidopsis thaliana*. Heredity 81:144–155

Struik PC (2010) Can crop physiology help to combat late blight? Potato Res 53(4):277–287

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

Taylaran RD, Adachi S, Ookawa T, Usuda H, Hirasawa T (2011) Hydraulic conductance as well as nitrogen accumulation plays a role in the higher rate of leaf photosynthesis of the most productive variety of rice in Japan. J Exp Bot 62:4067–4077

Teng S, Qian Q, Zeng D, Kunihiro Y, Fujimoto K, Huang D, Zhu L (2004) QTL analysis of leaf photosynthetic rate and related physiological traits in rice (*Oryza sativa* L.). Euphytica 135:1–7

Tuberosa R, Salvi S (2006) Genomics-based approaches to improve drought tolerance of crops. Trends Plant Sci 11:405–411

Uptmoor R, Schrag T, Stützel 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*. Mol Breed 21:205–216

Uptmoor R, Li J, Schrag T, Stützel H (2012) Prediction of flowering time in *Brassica oleracea* using a quantitative trait loci-based phenology model. Plant Biol 14:179–189

Vales MI, Schon CC, Capettini F, Chen XM, Corey AE, Mather DE, Mundt CC, Richardson KL, Sandoval-Islas JS, Utz HF, Hayes PM (2005) Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. Theor Appl Genet 111:1260–1270

van Eeuwijk FA, Malosetti M, Yin X, Struik PC, Stam P (2005) Statistical models for genotype by environment data: from conventional ANOVA models to eco-physiological QTL models. Aust J Agric Res 56:883–894

Virk PS, Ford-lloyd BV, Jackson MT, Pooni HS, Clemeno TP, Newbury HJ (1996) Predicting quantitative variation within rice germplasm using molecular markers. Heredity 76:296–304

Visker MHPW, Keizer LCP, van Eck HJ, Jacobsen E, Colon LT, Struik PC (2003) Can the QTL for late blight resistance on potato chromosome 5 be attributed to foliage maturity type? Theor Appl Genet 106:317–325

Welcker C, Sadok W, Dignat G, Renault M, Salvi S, Charcosset A, Tardieu F (2011) A common genetic determinism for sensitivities to soil water deficit and evaporative demand: meta-analysis of quantitative trait loci and introgression lines of maize. Plant Physiol 157:718–729

White JW (2006) From genome to wheat: emerging opportunities for modelling wheat growth and development. Eur J Agron 25:79–88

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, Herndl M, Hunt LA, Payne TS, Hoogenboom G (2008) Simulation-based analysis of effects of *Vrn* and *Ppd* loci on flowering in wheat. Crop Sci 48:678–687

Xu L, Henke M, Zhu J, Kurth W, Buck-Sorlin GH (2011) A functional–structural model of rice linking quantitative genetic information with morphological development and physiological processes. Ann Bot 107:817–828

Yin X, Chasalow S, Dourleijn CJ, Stam P, Kropff MJ (2000a) Coupling estimated effects of QTLs for physiological traits to a crop growth model: predicting yield variation among recombinant inbred lines in barley. Heredity 85:539–549

Yin X, Chasalow S, Stam P, Kropff MJ, Dourleijn CJ, Bos I, Bindraban PS (2002) Use of component analysis in QTL mapping of complex crop traits: a case study on yield in barley. Plant Breed 121:314–319

Yin X, Kropff MJ, Aggarwal PK, Peng S, Horie T (1997) Optimal preflowering phenology of irrigated rice for high yield potential in three Asian environments: a simulation study. Field Crops Res 51:19–27

Yin X, Kropff MJ, Goudriaan J, Stam P (2000b) A model analysis of yield differences among recombinant inbred lines in barley. Agron J 92:114–120

Yin X, Kropff MJ, Stam P (1999a) The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. Heredity 82:415–421

Yin X, Stam P, Dourleijn CJ, Kropff MJ (1999b) AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. Theor Appl Genet 99:244–253

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

Yin X, Struik PC, Kropff MJ (2004) Role of crop physiology in predicting gene-to-phenotype relationships. Trends Plant Sci 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. J Exp Bot 56:967–976

Zhao XQ, Xu JL, Zhao M, Lafitte R, Zhu L, Fu BY, Gao YM, Li ZK (2008) QTLs affecting morph-physiological traits related to drought tolerance detected in overlapping introgression lines of rice (*Oryza sativa* L.). Plant Sci 174:618–625

Zheng B, Biddulph B, Li D, Kuchel H, Chapman S (2013) Quantification of the effects of *VRN1* and *Ppd-D1* to predict spring wheat (*Triticum aestivum*) heading time across diverse environment. J Exp Bot 64:3747–3761

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