QTL-based physiological modelling of leaf photosynthesis and crop productivity of rice (*Oryza sativa* L.) under well-watered and drought environments

Junfei Gu

QTL-based physiological modelling of leaf photosynthesis and crop productivity of rice (*Oryza sativa* L.) under well-watered and drought environments

Junfei Gu

Abstract

Junfei Gu, 2013. QTL-based physiological modelling of leaf photosynthesis and crop productivity of rice (*Oryza sativa* L.) under well-watered and drought environments. PhD thesis, Wageningen University, Wageningen, The Netherlands. 181 pp.

Improving grain yield of rice (*Oryza sativa* L.) crop for both favourable and stressful environments is the main breeding objective to ensure food security. The objective of this study was to amalgamate crop modelling and genetic analysis, to create knowledge and insight useful in view of this breeding objective.

Photosynthesis is fundamental to biomass production, but the process is very sensitive to abiotic stresses, including drought. Upland rice cv. Haogelao, lowland rice cv. Shennong265, and 94 of their introgression lines (ILs) were studied under drought and well-watered conditions to analyse the genetics of leaf photosynthesis. After correcting for microclimate fluctuations, significant genetic variation was found in this population, and 1-3 quantitative trait loci (QTLs) were detected per photosynthesis-related trait. A major QTL was mapped near marker RM410 on Chromosome 9 and was consistent for phenotyping at flowering and grain filling, under drought and well-watered conditions, and across field and greenhouse experiments. These results suggest that photosynthesis at different phenological stages and under different environmental conditions is, at least to some extent, influenced by the same genetic factors.

To understand the physiological regulation of genetic variation and resulting QTLs for photosynthesis detected in the first study, 13 ILs were carefully selected as representatives of the population, based on the QTLs for leaf photosynthesis. These 13 ILs were studied under moderate drought and well-watered conditions in the experiment where combined gas exchange and chlorophyll fluorescence data were collected to assess CO_2 and light response curves. Using these curves, seven parameters of a photosynthesis model were estimated to dissect photosynthesis into stomatal conductance (g_s), mesophyll conductance (g_m), electron transport capacity (J_{max}), and Rubisco carboxylation capacity (V_{cmax}). Genetic variation in light saturated photosynthesis and the major QTL of photosynthesis on Chromosome 9 were mainly associated with variation in g_s and g_m. Furthermore, relationships between these parameters and leaf nitrogen or dry matter per unit area were shown valid for variation across

genotypes and across water treatments. In view of these results and literature reports, it was argued that variation in photosynthesis due to environmental conditions and to genetic variation shares common physiological mechanisms.

QTL analyses were further extended to other physiological parameters of rice. Molecular marker-based estimates of these traits from estimated additive allele effects were used as input to the mechanistic crop model GECROS. This marker/QTL-based modelling approach showed the ability of predicting genetic variation of crop performance within ILs for a diverse set of field conditions. This approach also showed the potential of extrapolating to a large population of recombinant inbred lines from the same parents. Most importantly, this model approach may improve the efficiency of marker-assisted selection, as it provides a tool to rank the relative importance of the identified markers in determining final yield under specific environmental conditions.

To examine the extent to which natural genetic variation in photosynthesis can contribute to increasing biomass production and yield of rice, the GECROS crop model was used again to analyse the impact of genetic variation in photosynthesis on crop biomass production. It was shown that in contrast to other studies a genetic variation in photosynthesis of 25% can be scaled up equally to crop level, resulting in an increase in biomass of 22-29% across different locations and years. The difference with earlier studies seems related to the fact that variation in both Rubisco-limited and electron transport-limited photosynthesis were observed in our IL population.

This thesis has contributed to closing the gap between genotype and phenotype by integrating crop physiology and genetics through an innovative QTL/marker-based modelling approach. This approach can contribute to making the use of genomics much more efficient in practical plant breeding.

Contents

CHAPTER 1

General introduction

The Food and Agriculture Organization (FAO) of the United Nations estimates that by 2050, world population will reach 9 billion (United Nations, 2011). To accommodate this population growth, the world will have to nearly double its current output of food, feed, and fibre. On the other hand, crops are increasingly exposed to stresses like drought in various ways and to different extents (Bouman *et al.*, 2007; Tuberosa, 2012). At the same time, elevated greenhouse gas concentrations will result in worldwide climate changes that will increase the frequency and severity of drought stress in agriculture (IPCC, 2007). Improving yield for both favourable and stress environments of major crops, therefore, is pivotal for world food security.

Developments in rice breeding
Rice is the most important food crop, with almost 600 million Mg produced annually on over 150 million ha (Khush, 2005). Rice is also the staple crop with the highest water requirement (Tuong & Bouman, 2003). To ensure food security, crop cultivars with greatly improved agronomic traits will be required. Improving rice yield for both favourable and stressful environments is the main breeding objective (Cattivelli *et al.*, 2008; Miura *et al.*, 2011; Tuberosa, 2012).

One way to realize higher yields over the past decades was by improving the harvest index through introducing dwarf genes, as was done in other major cereals such as wheat (Austin, 1999) and barley (Hellewell *et al.*, 2000). For rice the semi-dwarfing gene (*sd-1*) was first introduced through cross breeding, in the late 1950s in China and in the early 1960s at the International Rice Research Institute (IRRI), Philippines. In China in 1956-1959, the first dwarf variety, Guang-chang-ai, was developed using the *sd-1* gene from Ai-zi-zhan (Huang, 2001). In 1962, plant breeders at IRRI made crosses to introduce dwarfing genes from the variety Dee-geo-woo-gen of Taiwan into tropical, tall landraces. In 1966, IR8, the first semi-dwarf, high-yielding modern rice variety was released, which produced record yields throughout Asia and formed the basis for the development of new, high-yielding, semi-dwarf plant types (Khush *et al.*, 2001). Since the 1960s, *sd-1* remained the predominant semi-dwarfing gene present in rice cultivars.

In China in 1976, the first hybrid rice was developed, which showed an increase in potential yield of ~15% compared with pure line varieties (Yuan *et al.*, 1994). Since then, hybrid rice has greatly contributed to the global increase of rice production.

In the late 1980s, IRRI scientists proposed the idea of a new plant type, NPT (Khush, 1995). The NPT ideotype design focused first on large panicle size and reduced tillering capacity, and then on low panicle position for improved lodging resistance and canopy photosynthesis (Setter *et al.*, 1995). Although the NPT did not yield as hoped for (Peng *et al.*, 1999), it stimulated China to establish a project on the

development of 'super' rice in 1996 (Cheng *et al.*, 2007). China's 'super' rice breeding optimized the top three leaves and panicle position within a canopy in order to meet the demand of heavy panicles for a large source supply. All these approaches focused mainly on morphological traits, and in some cases also on physiological traits, in order to improve agronomic yield-related traits. Further progress is expected from a strengthening of the input of physiology into the breeding process (Mir *et al*., 2012).

Quantitative trait loci (QTLs) studies on agronomic traits in rice
Most important agronomic traits, such as crop yield and stress tolerance are quantitatively inherited. The nature of quantitative traits is that their expression is controlled by tens, or even hundreds, of quantitative trait loci (QTLs), most of them having only a small effect on the trait (Mackay *et al.*, 2009). To support the efficient manipulation of agronomic traits via breeding, it is important to identify QTLs conferring the variation of these traits. Using markers associated with these QTLs, researchers and breeders could accelerate breeding through so-called marker-assisted selection (Mohan *et al.*, 1997; Dekkers & Hospital, 2002; Collard & Mackill, 2008).

 In 2002, the genomes of two rice subspecies, *O. sativa* ssp. *japonica* (cv. Nipponbare) and *O. sativa* ssp. *indica* 93-11 were sequenced (Goff *et al.*, 2002; Yu *et al.*, 2002), but only in a draft version. Subsequently the final sequence of the entire rice genome of Nipponbare was completed by the International Rice Genome Sequencing Project (Matsumoto *et al.*, 2005). This achievement provided a vast amount of information on the rice genome and enabled detailed genetic analysis. Using this information, researchers have now succeeded in isolating and characterizing many important QTLs/genes (Table 1), which have the potential to greatly improve rice production. Below I highlight a few examples:

-   As a complex agronomic trait, grain yield of rice is co-determined by several component traits (Xing & Zhang, 2011), including number of tillers per plant (Li *et al.*, 2003; Takeda *et al.*, 2003), number of grains per panicle (Ashikari *et al.*, 2005; Xue *et al.*, 2008; Huang *et al*., 2009; Jiao *et al.*, 2010; Miura *et al.*, 2010), and individual grain weight (Fan *et al.*, 2006; Song *et al.*, 2007; Shomura *et al.*, 2008; Weng *et al.*, 2008; Li *et al.*, 2011). Grain yield is also strongly determined by genes coding for processes like grain filling (Wang *et al.*, 2008). These genes have been identified, cloned and characterized.
-   Heading date is an important trait for the adaptation of crops to different cultivation areas and cropping seasons (Izawa, 2007). To date many QTLs contributing to heading date in rice have been cloned, for example, *Hd1*, *Hd6*, *Ehd1*, *Hd3a*, *RID1*, and *DTH8* (Yano *et al.*, 2000; Takahashi *et al.*, 2001; Doi *et al.*, 2004; Tamaki *et al.*, 2007; Wu *et al.*, 2008; Wei *et al.*, 2010).

- There are many genes (Sasaki *et al.*, 2002; Zou *et al.*, 2006) associated with a semi-dwarf growth habit that gives plants a shorter stem that is more lodging resistant, especially in high nitrogen input environments. An alternative strategy to achieve lodging resistance is increasing culm strength (e.g. Ookawa *et al.*, 2010).
- Two QTLs for broad-spectrum blast resistance, *pi21* (Fukuoka *et al.*, 2009) and *Pb1* (Hayashi *et al.*, 2010) were cloned, and these are a valuable source for disease resistance. For abiotic stresses, many QTLs have also been found. For example, for submergence tolerance, *Sub1A* (Xu *et al.*, 2006), *SK1* and *SK2* (Hattori *et al.*, 2009); for salt tolerance, *SKC1*(Ren *et al.*, 2005); and for cold tolerance, *qLTG3-1* (Fujino *et al.*, 2008).
- Cloned genes (Chen *et al.*, 2008; Ding *et al.*, 2012; Yang *et al.*, 2012) related to hybrid sterility would allow for hybridization and gene flow, thus providing a chance for cross breeding across rice species.
- When favourable genes are introduced into cultivated rice from wild relatives it is necessary to eliminate any associated negative traits by marker-assisted selection. Thus, knowledge on genes conferring important traits lost during domestication is necessary for effective breeding. For example, genes for seed shattering (Konishi *et al.*, 2006; Li *et al.*, 2006) and prostrate growth (Jin *et al.*, 2008; Tan *et al.*, 2008) have been cloned.

All these findings have provided a promising way to increase yield for both favourable and stressful environments by combining all favourable QTLs or genes into one single variety through QTL pyramiding or a transgenic approach (Takeda & Matsuoka, 2008). But in most studies, the impacts of the genes on yield or stress tolerance were evaluated on the basis of a single isolated rice plant at optimum conditions, and hence it is unclear whether the gene can result in a real improvement of yield in terms of grain yield on per area basis in the field under variable and often limiting conditions. Although successful stories have been reported on the use of such genes to tailor high yield or drought-tolerant genotypes, this approach seldom led to release of new cultivars (Tardieu & Tuberosa, 2010). In field conditions, the gene/QTL expression is highly conditional on the environment (Stratton, 1998; van Eeuwijk *et al.*, 2005). It has been suggested that to help solve real-world problems more effort should be invested in integrating functional genomics with whole-crop physiology by considering all feed-back, feed-forward, and compensation mechanisms involving crop responses to environmental perturbations (Yin & Struik, 2008).

Table 1. Genes responsible for major QTLs related to crop yield and stress tolerance in rice.

Trait	QTL/gene	Encoded protein	Chr.	Reference
Yield and yield components				
Tillering	*MOC1*	GRAS family nuclear protein	6	(Li *et al.*, 2003)
Tillering	*OsTB1*	Transcription factor with TCP domain	3	(Takeda *et al.*, 2003)
Grain number	*Gn1a*	Cytokinin oxidase	1	(Ashikari *et al.*, 2005)
Grain number, plant height and heading date	*Ghd7*	CCT domain protein	7	(Xue *et al.*, 2008)
Grain number and strong culm	*dep1*	PEBP-like domain protein	9	(Huang *et al.*, 2009)
Grain number, low tiller number and strong culm	*Ipa*	OsSPL14	8	(Jiao *et al.*, 2010)
Grain number	*WFP*	OsSPL14	8	(Miura *et al.*, 2010)
Grain size	*gs3*	Transmembrane protein	3	(Fan *et al.*, 2006)
Grain size	*GS5*	Serine carboxypeptidase	5	(Li *et al.*, 2011)
Grain size and filling	*gw2*	RING-type ubiquitin E3 ligase	2	(Song *et al.*, 2007)
Grain size	*qSW5/GW5*	Unknown	5	(Shomura *et al.*, 2008; Weng *et al.*, 2008)
Grain filling	*GIF1*	Cell wall invertase	4	(Wang *et al.*, 2008)
Duration of the basic vegetative growth and photoperiod sensitivity				
Heading date	*Hd1*	CONSTANS-like protein	6	(Yano *et al.*, 2000)
Heading date	*Hd6*	α subunit of protein kinase	3	(Takahashi *et al.*, 2001)
Heading date	*Ehd1*	B-type response regulator	10	(Doi *et al.*, 2004)
Heading date	*Hd3a*	FT-like	6	(Tamaki *et al.*, 2007)
Switch from vegetative to floral development	*RID1*	Cys-2/His-2-type zinc finger transcription factor	10	(Wu *et al.*, 2008)
Days to heading	*DTH8*	CCT domain protein	8	(Wei *et al.*, 2010)
Plant height				
Plant height, high tillering	*Htd1*	OsCCD7	4	(Zou *et al.*, 2006)
Plant height	*sd1*	Gibberellin 20 oxidase	1	(Sasaki *et al.*, 2002)
Disease resistance				
Blast resistance	*pi21*	Proline-rich protein	12	(Fukuoka *et al.*, 2009)
Blast resistance	*Pb1*	CC-NBS-LRR protein	11	(Hayashi *et al.*, 2010)
Abiotic stress resistance				
Lodging resistance	*SCM2/OsAPO1*	F-box protein	6	(Ookawa *et al.*, 2010)
Submergence tolerance	*Sub1A*	ERF-related factor	9	(Xu *et al.*, 2006)
Internode elongation under submergence condition	*SK1, SK2*	ERF-related factor	12	(Hattori *et al.*, 2009)
Salt tolerance	*SKC1*	HKT-type transporter	1	(Ren *et al.*, 2005)
Cold tolerance	*qLTG3-1*	GRP and LTP domain	3	(Fujino *et al.*, 2008)
Sterility				
Regulate photoperiod-sensitive male sterility	*pms3*	A long noncoding RNA	12	(Ding *et al.*, 2012)
Hybrid sterility	*ORF3, ORF4, ORF5*	A killer-protector system	6	(Chen *et al.*, 2008; Yang *et al.*, 2012)
Domestication				
Seed shattering	*sh4*	Myb3 transcription factor	4	(Li *et al.*, 2006)
Seed shattering	*qSH1*	BEL1-like homeobox protein	1	(Konishi *et al.*, 2006)
Prostrate growth	*PROG1*	Zinc finger transcription factor	7	(Jin *et al.*, 2008; Tan *et al.*, 2008)

Physiological perspective of crop production

From a crop physiological perspective, grain yield of rice is the outcome of three constituents: the source, sink, and flow. The source is the capacity to supply sufficient assimilates to the sink via photosynthesis; the sink is the yield potential represented by the mathematical product of yield components; and the flow is the ability of transport processes to translocate the photosynthetic products and other nutrients to feed the sink at the highest possible rate. Improved yield can only be achieved if the source, the sink and the flow are in full balance and functioning well throughout the crop cycle. Simultaneously, the increase in crop production should be reached with a decrease in availability and certainty of water resources for crop production. For rice this poses an additional challenge as the crop is highly drought susceptible.

Crop photosynthesis, being the source of energy and inorganic carbon necessary for crop growth, depends on the ability of the crop to build up and maintain a canopy for capturing incoming light, but also on the photosynthetic capacity and efficiency of individual leaves. For rice, leaf area dynamics and canopy architecture may have been effectively optimized for maximum light capture through breeding (Horton, 2000). At the leaf level, photosynthesis is not only controlled by diffusion components [stomatal conductance (g_s) and mesophyll conductance (g_m)], but also by various biochemical capacities of protein complexes. The potential activity of Rubisco (V_{cmax}) limits photosynthesis at low CO_2 concentration in the chloroplast stroma (C_c), the electron transport capacity of the chloroplast (J_{max}) limits photosynthesis at high C_c (Farquhar *et al.*, 1980). At ambient CO_2 concentration, the light-saturated photosynthesis is limited by V_{cmax} (Farquhar & Sharkey, 1982), and the V_{cmax} and J_{max} are closely related to leaf nitrogen content per unit area (N_a) (Harley *et al.*, 1992b). N_a is not only a genetic trait (Cook & Evans, 1983), but also affected by plant ontogeny and the competition for nitrogen between source and sink (Mae, 1997).

Photosynthesis is also greatly influenced by the environment during growth, particularly as the microclimate unavoidably fluctuates under natural field conditions (Flood *et al.*, 2011). Abiotic stress and leaf ontogeny have a large effect on photosynthesis (Lawlor & Cornic, 2002; Flexas *et al.*, 2004; Grassi & Magnani, 2005; Chaves *et al.*, 2009). Especially water stress will dramatically decrease photosynthesis through control of g_s and g_m (Flexas *et al.*, 2004), or a decrease of the contents of RuBP as well as the activities of the major carbon reduction cycle enzyme Rubisco (Tezara *et al.*, 1999; Lawlor & Cornic, 2002).

Not only the source (photosynthesis), but also the sink and the flow are regulated by environmental factors. For example, spikelet sterility is determined towards the end of panicle formation, successful pollination is determined during flowering, and the majority of the carbohydrates present in mature seeds is determined by assimilation

during grain filling. Low availability of assimilates from photosynthesis during periods each of the yield components are determined reduces sink size (Bindraban *et al.*, 1998; Boyer & Westgate, 2004; Ji *et al.*, 2010). On the other hand, the source activity (photosynthesis) may also be limited by metabolite transport processes (Bräutigam & Weber, 2011) and sink strength (Herold, 1980; McCormick *et al.*, 2006). Because of the complexity of regulatory networks in plant and crop systems and given the complex interactions with the environment, field crops show strong genotype $\times$ environment (G$\times$E) interactions. This complexity is especially relevant when breeding for drought tolerance.

Drought tolerance is important, but, sometimes, spectacular results obtained in one drought scenario might have limited relevance for improving yield in other scenarios as drought varies in intensity and timing (Tardieu, 2011). For example, in wheat, selection for genotypes of higher transpiration efficiency (low $\Delta^{13}C$) could improve yield by 10% in very dry seasons (Rebetzke *et al.*, 2002), but the yield advantage can disappear at moderate stress (Rebetzke *et al.*, 2002), and even hamper plant growth resulting in smaller plants with reduced transpiration, biomass and yield (Condon *et al.*, 2004; Blum, 2005). These G$\times$E interactions always result in inconsistency of morphophysiological traits, which make the selection criteria for breeding complex and unstable, especially under drought. Therefore, it is necessary to accurately model and predict G$\times$E interactions for improving breeding efficiency and MAS.

Integration of crop physiology with genetics — QTL-based modelling
Since the pioneering work on plant modelling by C.T. de Wit (1959), ecophysiological crop models have been developed extensively by integrating knowledge from different disciplines, such as crop physiology, micrometeorology, soil science, and computing technologies (Loomis *et al.*, 1979; Bouman *et al.*, 1996; McCown *et al.*, 1996). Now, crop models based on solid crop-physiological knowledge can quantify causality between relevant physiological processes and responses of these processes to environmental variables. Therefore, in principle, these crop models enable predictions beyond the environments in which the model parameters were derived and can reveal how G$\times$E interactions come about (Yin *et al.*, 2000a; 2004; Sinclair, 2011). Crop related model input parameters are also referred to as 'genetic coefficients' (White & Hoogenboom, 1996; Mavromatis *et al.*, 2001; Messina *et al.*, 2006; White *et al.*, 2008) because these model-input parameters might be (at least partly) under genetic control. Therefore, crop modelling has been used to give suggestions for ideotype breeding (Penning de Vries, 1991; Dingkuhn *et al.*, 1993; Kropff *et al.*, 1995; Haverkort & Kooman, 1997).

However, crop models often do not consider the genetic basis of model parameters that describe genotypic differences (Stam, 1998; Koornneef & Stam, 2001), nor do they consider how much genetic variation exists in the genetic materials available for breeding. Yin *et al.* (1999a, b; 2000b) first tried to combine crop modelling with QTL mapping using a SUCROS-type crop model. The QTL analysis was first applied to the model-input traits. After the QTL analysis, the identified QTLs were then coupled to the crop model by replacing the original, measured input trait values with those predicted from the QTL effects (Yin *et al.*, 2000b). This approach was first showcased for predicting differences in yield among relatively similar lines from a genetic population. This QTL-based modelling approach was later used to study crop traits such as leaf elongation rate in maize, flowering time, and fruit quality (Reymond *et al.*, 2003; Quilot *et al.*, 2004; Nakagawa *et al.*, 2005; Quilot *et al.*, 2005; Yin *et al.*, 2005b; Uptmoor *et al.*, 2008; Bertin *et al.*, 2010; Prudent *et al.*, 2011). These later studies showed that this approach was robust in predicting genetic differences in bi-parental crossing populations under different conditions (in terms of vapour pressure deficit, soil moisture content, temperature and photoperiod). The QTL-based modelling approach applied to complex traits (e.g. yield) was not very successful (Yin *et al.*, 1999a,b; 2000b), when compared with results applied to single crop traits. The reason is that yield is much more complex considering the hierarchy from leaf photosynthesis to crop yield (Yin & Struik, 2008), and further improvements in crop models were suggested by Yin & Struik (2008).

QTL-based modelling could potentially evaluate constraints in breeding due either to limited genetic variation or to correlations between the traits. QTL-based modelling could evaluate the effect of QTLs for traits at organ level on crop yield under different environments, which could be useful in breeding for specific environments. For example, Chenu *et al*. (2009) using the crop model APSIM-Maize simulated that a QTL accelerating leaf elongation could increase yield in an environment with water deficit before flowering, but could reduce yield under terminal drought stress. QTL-based modelling could also be useful in supporting marker-assisted selection.

In short, QTL-based crop modelling, combining ecophysiological modelling and genetic mapping, can dissect complex yield traits into component traits, integrate effects of QTLs of the component traits over time and space at the whole crop level, and predict yield of various allele combinations under different environmental conditions.

Objectives and approach

In this study, I tried to amalgamate physiological and genetic approaches to study a rice genetic population, using a crop model to create knowledge and insight useful for

breeding. The difficulty to phenotype a large germplasm collection for specific trait performance, has been a critical limitation in applying physiological information and crop modelling in genetic analysis for more than half a century (Yin *et al.*, 2003a, 2004; Sinclair *et al.*, 2004; Houle *et al.*, 2010; Sinclair, 2011). Most physiological studies require detailed, sophisticated and usually expensive techniques to phenotype plants, and can be only applied to a few genotypes, while a genetic analysis always involves quick and simple phenotyping of many (often >100) genotypes. The dilemma of phenotypic screens is that they are either too difficult and sophisticated, or too crude and with poor resolution (Salekdeh *et al.*, 2009; Sinclair, 2011).

In this thesis, I will use a relatively new crop model GECROS (Genotype-by-Environment interaction on CROp growth Simulator, Yin & van Laar, 2005), which requires relatively few, easily phenotyped input parameters. The model was structured from the principles of the whole-crop system dynamics to embody physiological causes of crop performance. I will test whether the model is potentially useful for designing ideotypes in support of marker-assisted selection. Furthermore, the information from combined genetic and physiological analysis of GECROS model input parameters will be scaled up to predict crop yield under diverse environmental conditions. The general methodological steps are outlined as:

First, a simple screen that allows a large genetic population to be examined is a first-tier run. This round of screening could focus on simple and easily measured traits.

Second, a more sophisticated physiological study was conducted on a smaller, but for physiologists still relatively large number of genotypes (i.e. 10~30 lines), selected from the results of the QTL analysis from the first step. Methods are available to describe mathematically traits that vary rapidly with environmental conditions (e.g. photosynthesis) (Yin *et al.*, 2009b). In this step, for each genotype, a set of genotype-specific physiological parameters was calculated.

Last and most importantly, the genetic information and physiological analysis of processes at the lower level will be scaled up, based on QTL effects and allelic information, to crop level by using the GECROS model. A QTL-based crop model will play an important role in the upscaling (Yin *et al.*, 2000b, 2004; Tardieu, 2003; Chenu *et al.*, 2009; Tardieu & Tuberosa, 2010), because these process-based crop growth models have the potential to assess a complex trait at a higher organizational level, via integrating the information about processes at lower level (Keating *et al.*, 2003; Yin & Struik, 2008; Hammer *et al.*, 2010; Zhu *et al.*, 2011).

In this thesis I will apply this methodology, with a focus on photosynthesis, which is the source of energy and inorganic carbon necessary for crop growth. The physiological process of photosynthesis and the photosynthetic responses to environment changes have been intensively studied (von Caemmerer, 2000; Bernacchi

et al., 2001; Farquhar *et al.*, 2001; Bernacchi *et al.*, 2002; Tuzet *et al.*, 2003; Flexas *et al.*, 2004; Yin *et al.*, 2009b). Extensive genetic variation in photosynthesis has been found in rice germplasm (Cook & Evans, 1983; Dingkuhn *et al.*, 1989; Sasaki & Ishii, 1992; Adachi *et al.*, 2011). Given the well-defined understanding of photosynthesis, photosynthesis is a good target for applying integrated genetic analysis and physiological analysis. More specifically, I will

1. Map QTLs for photosynthetic parameters of rice assessed by both gas exchange and chlorophyll fluorescence measurements under drought and well watered field conditions.

2. Investigate the physiological basis of genetic variation and resulting QTLs identified.

3. Examine the ability of the ecophysiological crop growth model GECROS to account for yield differences among individual genetic lines of rice.

4. Analyse the ability of the GECROS model with QTL-based estimates of physiological input parameters to predict the yield of the lines.

5. Analyze the relative importance of individual physiological traits or markers in accounting for genetic variation in yield.

6. Test the ability of the marker-based approach to predict yield differences in an independent set of genotypes of the same parents.

7. Examine the extent to which exploiting the natural genetic variation in leaf photosynthesis can contribute to variation in canopy photosynthesis and in crop productivity in rice.

I expect, such an approach, consisting of using knowledge of fundamental plant biology for elementary traits, at the same time considering their genetic variation, will improve the understanding of plant responses to environmental factors, and improve the efficiency of breeding for traits at the crop level. The approach should be considered as a contribution of genetics and crop physiology to systems biology scaled up at the whole-plant level and aimed at bridging the gap between functional genomics and crop field performance.

Outline of the thesis

To achieve the above research objective to support plant breeding, an introgression line (IL) population is preferred for this study. ILs are plant series that possess segments of the donor parent chromosome in the background of the recurrent parent. Phenotypic characterization of each line can reveal which chromosome fragment from the donor has the gene(s) associated with an interesting trait. ILs have been widely used in QTL validation, QTLs pyramiding, and map-based cloning because of the

simple genetic background (Ashikari & Matsuoka, 2006; Ando *et al.*, 2008; Takeda & Matsuoka, 2008).

In this thesis, I first evaluate the genetic variation of leaf photosynthesis parameters in a rice genetic population consisting of 94 advanced backcross ILs and two parents (**Chapter 2**). The parents were the lowland rice cv. Shennong265 (*japonica*, recurrent parent) and the upland rice cv. Haogelao (*indica-japonica* intermediate, donor parent). Haogelao is drought tolerant, but low-yielding; Shennong265 is drought susceptible, but high-yielding under irrigated conditions. After a cross between the two parents, the resultant F_1 plants were backcrossed with Shennong265 three times, and these BC_3F_1 plants were consecutively self-pollinated five times to construct the genetic population BC_3F_6 by the single-seed descent method (Gu, 2007).

Based on the QTLs related to gas exchange and chlorophyll fluorescence detected under well-watered and drought-stressed field condition in Chapter 2, 13 ILs were selected for detailed physiological study. In **Chapter 3**, the biochemical photosynthesis model of Farquhar, von Caemmerer & Berry (1980), combined with a phenomenological model for quantifying stomatal and mesophyll conductance (Yin *et al.*, 2009b), is used to study the physiological basis of genetic variation and resulting QTLs for photosynthesis in the 13 selected ILs. With this combined model, photosynthesis was dissected into stomatal conductance (g_s), mesophyll conductance (g_m), electron transport capacity (J_{max}), and Rubisco carboxylation capacity (V_{cmax}). Significant genetic variation in these parameters was found. This genetic variation in photosynthesis model parameters, together with other measured physiological input parameters related to phenological and morphological development, is used to feed the ecophysiological crop model GECROS to test the ability to predict yield differences in the IL population, and to extrapolate the prediction to an independent recombinant inbred line population (**Chapter 4**). A model-based sensitivity analysis is presented to provide breeders with more information for marker-assisted selection.

Although the crop model GECROS gave a fair prediction, Chapter 4 also shows that introducing genetic variation in leaf photosynthesis into model analysis did not improve the ability to predict crop yields. This could be caused by the fact that ILs differed genetically in many respects beyond photosynthesis. To examine to what extent observed natural genetic variation in leaf photosynthesis can potentially contribute to increasing canopy photosynthesis and rice productivity, the crop model GECROS was also used to scale up, and project the genetic variation at leaf level to crop growth during the whole growing season at different climatic scenarios and in different years (**Chapter 5**).

I hope the integration of crop physiology with genetics — QTL-based modelling of crop yield - could accelerate plant breeding. I also hope this thesis research will exemplify the concepts of 'crop systems biology' (Yin & Struik, 2008, 2010): to bring the information from functional genomics to crop level; to better understand the organization, intra- and inter-plant competition and crop responses to environmental conditions; to fill the vast middle ground between 'omics' research and relative simple crop models; and to promote communication across various biological scales. These expectations will be discussed in **Chapter 6** on the basis of the results presented in the thesis.

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

Junfei Gu [a], Xinyou Yin [a], Paul C. Struik [a], Tjeerd-Jan Stomph [a], and Huaqi Wang [b]

[a] Centre for Crop Systems Analysis, Department of Plant Science, Wageningen University, P. O. Box 430, 6700 AK Wageningen, The Netherlands
[b] Plant Breeding & Genetics, China Agricultural University, 100193 Beijing, P.R. China

ABSTRACT

Photosynthesis is fundamental to biomass production, but sensitive to drought. To understand the genetics of leaf photosynthesis, especially under drought, upland rice cv. Haogelao, lowland rice cv. Shennong265 and 94 of their introgression lines (ILs) were studied at flowering and grain filling under drought and well watered field conditions. Gas exchange and chlorophyll fluorescence measurements were conducted to evaluate eight photosynthetic traits. Since these traits are very sensitive to fluctuations in micro-climate during measurements under field conditions, we adjusted observations for micro-climatic differences through both a statistical covariant model and a physiological approach. Both approaches identified leaf-to-air vapour pressure difference as the variable influencing the traits most. Using the SSR linkage map for the IL population, we detected 1-3 QTLs per trait-stage-treatment combination, which explained between 7.0 and 30.4% of the phenotypic variance of each trait. The clustered QTLs near marker RM410 (the interval from 57.3 to 68.4 cM on chromosome 9) were consistent over both development stages and both drought and well watered conditions. This QTL consistency was verified by a greenhouse experiment under controlled environment. The alleles from the upland rice at this interval had positive effects on net photosynthetic rate, stomatal conductance, transpiration rate, quantum yield of PSII and the maximum efficiency of light adapted open PSII. However, the allele of another main QTL from upland rice was associated with increased drought sensitivity of photosynthesis. These results could potentially be used in breeding programs through marker assisted selection to simultaneously improve drought tolerance and photosynthesis.

Key words: Drought, photosynthesis, physiological model, quantitative trait locus (QTLs), gas exchange, chlorophyll fluorescence.

INTRODUCTION

Drought is considered to be the greatest threat to rice (*Oryza sativa* L.) production (Sharma & De Datta, 1994). The complex quantitative genetics nature of drought tolerance was once thought to be the main constraint for breeding for improved rice varieties under drought prone environments (Nguyen *et al.*, 1997). Yet, recent evidence has shown that progress can be made by direct selection for grain yield under managed stress trials (Venuprasad *et al.*, 2007, 2008; Bernier *et al*., 2007; Kumar *et al*., 2008). For further progress, indirect methods based on effective selection criteria and on molecular markers for component traits should be explored (Miura *et al*., 2011).

Today unprecedented efforts are being made in dissecting complex traits into their single genetic determinants – quantitative trait loci (QTLs), in order to support marker-assisted selection (MAS) and, eventually, cloning of genes. An increasing number of QTLs related to drought response have been reported and these include QTLs for root morphology, other root traits like root penetration ability (Price *et al.*, 2000, 2002; Babu *et al.*, 2003; Uga *et al.*, 2011); osmotic adjustment (Robin *et al.*, 2003); grain yield and yield components (Lanceras *et al.*, 2004; Lafitte *et al.*, 2004; Xu *et al.*, 2005); stay green (Jiang *et al.*, 2004); canopy temperature, leaf rolling and leaf drying (Yue *et al.*, 2005); and carbon isotope discrimination ($\Delta^{13}C$) (Takai *et al.*, 2009; Xu *et al.*, 2009).

Photosynthesis, being the basis of crop growth, biomass production and yield, is one of the primary physiological processes strongly affected by drought (Chaves, 1991; Lawlor, 1995). The photosynthesis response to drought is very complex. Generally, during the onset of drought, CO_2 diffusional resistances increase, especially because stomatal aperture can change rapidly (Chaves *et al.*, 2002; Cochard *et al.*, 2002; Lawlor, 2002). With the progress of drought and tissue dehydration, metabolic impairment will arise gradually, including decrease of the content and activities of the major photosynthetic carbon reduction cycle enzyme, ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco), as well as ribulose 1,5-bisphosphate (RuBP) (Reddy, 1996; Tezara *et al.*, 1999). Besides the CO_2 diffusion and CO_2 fixation pathways, Photosystem II (PSII) electron transport is very susceptible to drought (Havaux, 1992; Lu & Zhang, 1999). Chlorophyll fluorescence, emitted mainly by PSII in the range of 680-740 nm spectra region, has been widely used for the estimation of PSII electron transport rate *in vivo*. Combining gas exchange measurements for CO_2 fixation and chlorophyll fluorescence data for PSII electron transport may bring new insights into the regulation of photosynthesis in response to environment variables (von Caemmerer, 2000). The advent of portable open gas exchange systems integrated with chlorophyll fluorescence measuring devices enables researchers to not only simultaneously measure net photosynthetic rate (A), stomatal conductance for CO_2 (g_s), transpiration rate (T_r), intercellular CO_2 partial pressure (C_i), transpiration efficiency (TE), quantum yield of PSII (Φ_{PSII}), proportion of open PSII (qP), and maximum efficiency of open PSII in the light (F'_v/F'_m) in real time in the field, but also keep records of micro-climatic conditions during the observations such as leaf-to-air vapour pressure difference (VPD) and leaf temperature (T_{leaf}) (Long & Bernacchi, 2003).

Because of the primary importance of photosynthesis in determining crop growth, identifying QTLs controlling photosynthesis parameters is an important step in enhancing MAS for improved yield. This assertion is supported by growing evidence

that there is genetic variation for photosynthetic rates among available germplasm and that recent yield progress in cereals from breeding was associated with increased photosynthesis (Fischer & Edmeades, 2010). In rice, using 20 distinct varieties, Jahn *et al.* (2011) showed notable genetic variation in leaf photosynthetic rate. However, only a few QTL studies have been reported so far for photosynthetic traits (Teng *et al.*, 2004; Zhao *et al.*, 2008; Adachi *et al.*, 2011), probably partly because gas exchange measurements to phenotype these parameters under field conditions are laborious and phenotypes are greatly influenced by environments during growth and measurement, particularly when the micro-climate unavoidably fluctuates under natural field conditions (Flood *et al.*, 2011). It is very hard to expose genotypes to exactly the same environmental conditions in terms of temperature, soil water content, and VPD. Therefore, environmental noise is usually large and obscures genetic differences, resulting in large QTL $\times$ environment interactions or in irreproducible results (e.g., Simko *et al.*, 1999; Yin *et al.*, 1999a,b). Observations must therefore be corrected for differences in microclimate.

In this study, we aim at precision mapping of QTLs for photosynthetic parameters of rice assessed by both gas exchange and chlorophyll fluorescence under drought and well watered field conditions. Two strategies were applied. Firstly, we developed an advanced backcross introgression lines (IL) population, which allows to more precisely identify QTLs than the more commonly used populations such as recombinant inbred lines (RIL). Secondly, we explored both statistical and physiological approaches to correct for micro-climate variation during observations, thus enhancing the precision of observed phenotypic trait values for mapping. Our IL population was developed from a cross between a lowland rice and an upland rice variety, since upland rice relies exclusively on rainfall for water uptake and is generally thought to be more drought resistant.

MATERIAL AND METHODS

Plant materials
The mapping population consisted of 94 advanced backcross introgression lines (ILs). The parents were the lowland rice cv. Shennong265 (*Japonica*) and the upland rice cv. Haogelao (*Indica-Japonica* intermediate). The two cultivars were contrasting in term of their agronomic performance under drought condition (La, 2004; Gu, 2007). Haogelao is drought tolerant, but low yielding; whereas Shennong265 drought susceptible, but high yielding under irrigated conditions. After cross between the two parents, the resultant F_1 plants were backcrossed with paternal cultivar Shennong265

three times and these BC_3F_1 plants were consecutively self-pollinated five times to construct the mapping population BC_3F_6 by the single seed descent method.

DNA extraction and simple sequence repeats (SSR) analysis
Fresh leaves were collected from the BC_3F_6 lines and ground in liquid nitrogen. DNA was extracted from the ground tissue using the cetyltrimethyl ammonium bromide (CTAB) method (Rogers & Bendich, 1985). SSR primers were synthesized according to the sequences published by McCouch *et al.* (2002). A total volume of 25µl reaction mixture was composed of 1 ng/µl template DNA, 10 mmol Tris–HCl (pH 9.0), 50 mmol KCl, 1.5 mmol $MgCl_2$, 0.1% Triton X-100, 2 μmol of each primer, 2.5 mM each of dNTP (dATP, dCTP, dGTP, dTTP), and 1 unit Taq DNA polymerase. Amplification was performed on a program for the initial denaturation step with 94℃ for 5 min, followed by 35 cycles for 1 min at 94℃, 1 min at 55℃, 2 min at 72℃, with a final 10 min extension at 72℃. The PCR (polymerase chain reaction) products were separated on 8% polyacrylamide denaturing gels and the bands were revealed using the sliver-staining protocol described by Panaud *et al.* (1996).

Phenotypic evaluation
Plants of the introgression population and recipient and donor parents were grown at the experimental station of China Agricultural University, Beijing (39°N, 116°E), China, in 2009, following a complete randomized block design, with two replications, four rows per plot (plot size 2.5 m × 1.2 m), 7.5 cm between plants within each row and 30 cm between rows in both rainfed upland and fully irrigated lowland field conditions. The crops were managed according to standard local practice, with the following fertilizer applications: 48 kg N ha^{-1}, 120 kg P_2O_5 ha^{-1} and 100 kg K_2O ha^{-1} as the basal fertilizer, and additional 86 kg N ha^{-1} at the tillering stage and 28 kg N ha^{-1} at the booting stage. Weeds in both lowland and upland fields were controlled by a combination of chemical and manual methods, and insects were controlled chemically.

   The flowering of the population occurred between 105 and 120 days after sowing for the drought stressed environment, and between 107 and 119 days after sowing for the well watered environment. Gas exchange and chlorophyll fluorescence measurements covered both flowering stage and mid grain filling stage (around two weeks after flowering). The measurements were adjusted by considering the flowering time and the variation of flowering time in each line to make sure each genotype had three replicates per block. For the drought stressed environment, soil moisture was monitored with the time domain reflectometry method (TDR-TRIM-FM) at a soil depth of 0 ~ 30 cm. During the photosynthesis measurements, the soil water content

was around 13-16% (v/v) at flowering stage, and around 15-19% (v/v) at grain filling stage. Normally measurements were made during a clear day, between 9:00 and 11:30 am and between 13:00 and 15:00 pm, with photosynthetic photon flux density (PPFD) of natural sunlight between 700 and 1600 μmol m^{-2} s^{-1}; T_{leaf} varied from 23.3℃ to 36.0℃ (Fig. 1) and relative humidity from 17.4% to 67.8% (partly shown by VPD in Fig. 1) during the measurements.

The middle parts of three fully expanded flag leaves on the main culms of three central plants in each plot were measured using a portable open gas exchange system (Li-6400, Li-COR Inc., Lincoln, Nebraska, USA) with an integrated fluorescence chamber head (LI-6400-40, Li-COR Inc., Lincoln, Nebraska, USA) with setting of PPFD at 1000 μmol m^{-2} s^{-1} and a CO_2 concentration (C_a) at 400 μmol CO_2 (mol air)$^{-1}$ by using CO_2 cylinders. Gas exchange data for net photosynthesis rate (A); intercellular CO_2 partial pressure (C_i); stomatal conductance for CO_2 (g_s); transpiration rate (T_r) and fluorescence data for F_s (the steady-state fluorescence) were recorded after maintaining the leaf in the leaf chamber long enough for A to reach a steady state. Besides this, micro-climatic data (T_{leaf}, VPD, etc.) was automatically recorded at the same time. Then a saturating light pulse (>8500 μmol m^{-2} s^{-1} for 0.8 s) was applied to determine F'_m (the maximum fluorescence during the saturating light pulse). By the end, after turning off the actinic light, a "dark pulse" (using far-red light to preferentially excite PSI and force electrons to drain from PSII) was applied to get F'_o (the minimum fluorescence yield in the light-adapted state). From these data three chlorophyll fluorescence parameters were derived:

1. $\varPhi_{PSII} = (F'_m - F_s)/F'_m$, the apparent PSII e$^-$ transport efficiency (Genty *et al.*, 1989), which estimates the yield of PSII photochemistry;

2. $qP = (F'_m - F_s)/(F'_m - F'_o)$, which quantifies the photochemical capacity of PSII (Bradbury & Baker, 1984; Quick & Horton, 1984);

3. $F'_v/F'_m = (F'_m - F'_o)/F'_m$, which quantifies the extent to which photochemistry at PSII is limited by competition with thermal decay processes (Oxborough & Baker, 1997).

From gas exchange data, transpiration efficiency (TE) was calculated as A/T_r. To assess any genetic difference in the responsiveness to drought, the ratio of A under the drought treatment ($A_{drought}$) to that under the well-watered treatment (A_{water}) was calculated to indicate drought sensitivity (DS), for both flowering and grain-filling stages.
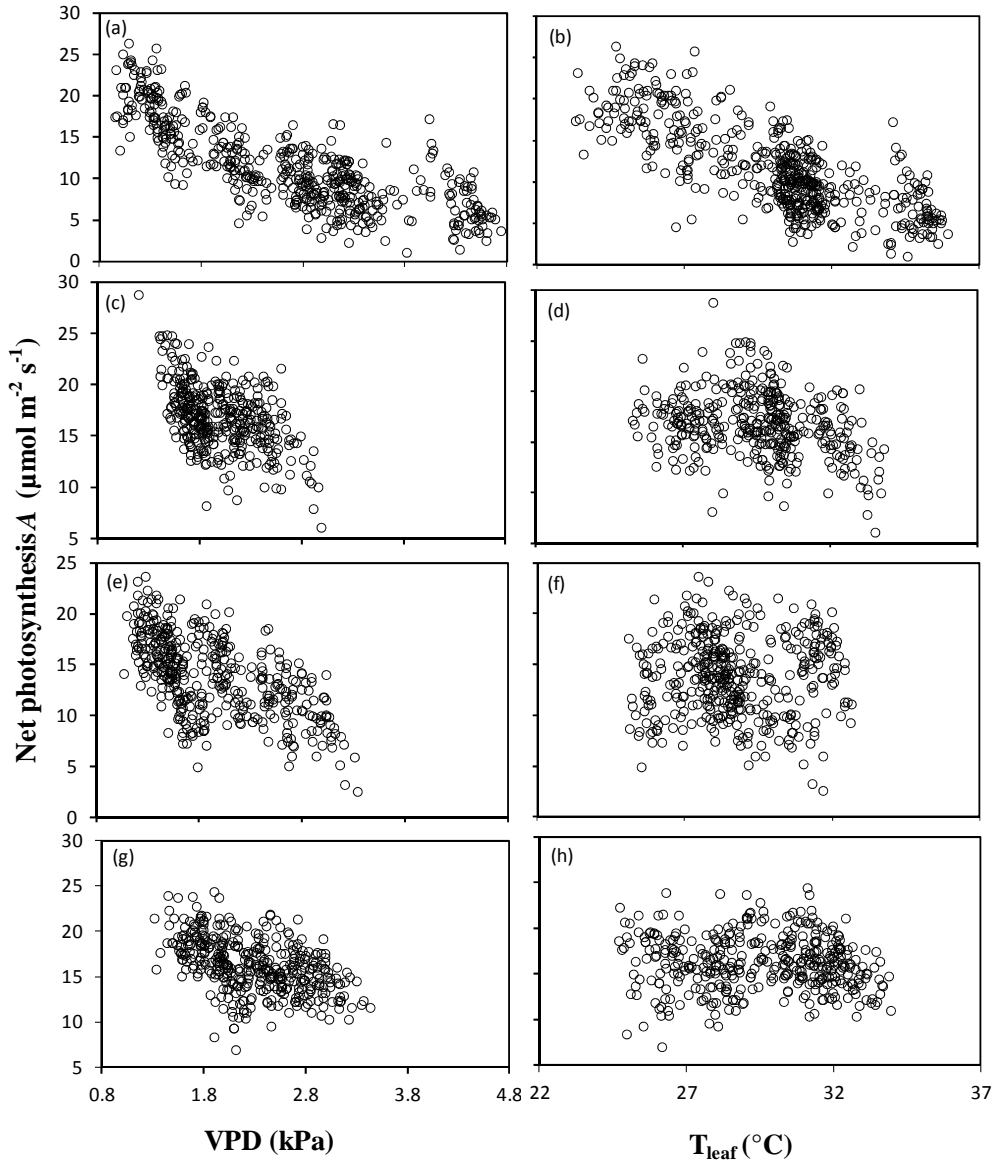
Fig. 1 Correlation between net photosynthesis A (μmol m^{-2} s^{-1}) and vapour pressure deficit VPD (the left column of plots) or T_{leaf} (the right column of plots) under different stage $\times$ treatment combinations. At flowering stage for drought stressed plants (a,b) and for well watered plants (c,d) and at mid grain filling for drought stressed plants (e,f) and for well watered plants (g,h). The minimum, mean ($\pm$ SD), maximum values are: for FS: VPD (0.96, 2.66 $\pm$ 1.01, 4.75), T_{leaf} (23.3, 30.1 $\pm$ 3.0, 36.0); for FW: VPD (1.21, 2.00 $\pm$ 0.35, 3.00), T_{leaf} (25.3, 29.5 $\pm$ 1.9, 33.9); for GS: VPD (1.07, 1.94 $\pm$ 0.56, 3.35), T_{leaf} (25.1, 28.7 $\pm$ 1.8, 32.7); for GW: VPD (1.32, 2.29 $\pm$ 0.47, 3.44), T_{leaf} (24.7, 29.5 $\pm$ 2.37, 34.0).

Adjusting for the effects of environmental fluctuations on trait values

As photosynthetic rate A or related traits (e.g. stomatal conductance) can strongly vary with environmental variables such as VPD (Cowan, 1977; Buckley & Mott, 2002), the phenotypic trait value of the i-th genetic line was expressed in a statistical covariant model using the environmental variable as a quantitative co-regressor:

$$u_{ijk} = u + G_i + E_j + (GE)_{ij} + B_k + bx_{ijk} + e_{ijk} \tag{1}$$

where, μ = general mean; G_i = genetic effect of i th genotype; E_j = treatment effect, which stands for either of the two treatments (well watered or drought stressed); $(GE)_{ij}$ = genotype × treatment interaction; B_k = the block effect; b = the effect of the environmental variable; x_{ijk}= values of the environmental variable during measurement; e_{ijk}= residual effect. This approach allows the observed trait values to be adjusted statistically to the same value (e.g. average) of the climatic variable (T_{leaf} and VPD) that had inevitably fluctuated during the field measurement conditions. The analysis identified that VPD was the most influential environmental factor (see Results).

Such a statistical approach often results in increased precision for parameter estimates and increased power for statistical tests of hypotheses (Ott & Longnecker, 2001). An alternative is to use a physiological approach (Yin *et al.*, 1999a), which helps to confirm the reliability of the statistical approach. We therefore explored the use of a physiological approach, based on the photosynthesis model of Farquhar *et al.* (1980), to correct for the effects of environmental fluctuations during measurements. Since A is Rubisco limited under our measuring conditions (i.e. light intensity of 1000 μmol m^{-2} s^{-1} in ambient CO_2 concentration), A can be expressed as a consequence of CO_2 and O_2 competing for the Rubisco binding site by carboxylation and oxygenation, respectively:

$$A = \frac{(C_i - \Gamma_*)V_{cmax}}{C_i + K_M} - R_d \tag{2}$$

where V_{cmax} is the maximum rate of Rubisco carboxylation, C_i is the intercellular CO_2 partial pressure, Γ_* is the CO_2 compensation point in the absence of day respiration (R_d), and K_M is the effective Michaelis-Menten constant. K_M is expressed as $K_{mc}(1+O/K_{mo})$, where K_{mc} and K_{mo} are the Michaelis-Menten constants for CO_2 and O_2, respectively, and O is the oxygen concentration.

In order to incorporate the effect of VPD, the model of Ball *et al.* (1987), as modified by Leuning (1990, 1995), states that

$$g_s = g_o + Af_{vpd} \tag{3}$$

where g_s is the stomatal conductance for CO_2 diffusion, g$_o$ is the residual stomatal conductance if the irradiance approaches zero, and f_{vpd} is the term for the effect of leaf-

to-air VPD. f_{vpd} is expressed as $a_1/[(C_s - \Gamma)(1+D_s/D_o)]$, where D_s is the vapour pressure deficit, C_s is CO_2 concentration at leaf surface (which was obtained from C_a and a default value for boundary-layer conductance set in the Li-Cor), and a_1 and D_o are empirical coefficients, Γ is the CO_2 compensation point, which can be derived from eqn (2) (e.g., see Azcón-Bieto *et al.*, 1981) as :

$$\Gamma = \frac{\Gamma_* + K_{mc}(1 + O/K_{mo})R_d/V_{cmax}}{1 - R_d/V_{cmax}} \tag{4}$$

Combining Eqns (2) and (3), and replacing C_i by $(C_s - A/g_s)$, and then solving for A give

$$A = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \tag{5}$$

where

$a = f_{vpd}(C_s + K_M) - 1$
$b = (C_s f_{vpd} + K_M f_{vpd} - 1)R_d + (C_s + K_M)g_o - (C_s f_{vpd} - \Gamma_* f_{vpd} - 1)V_{cmax}$
$c = [(C_s + K_M)R_d - (C_s - \Gamma_*)V_{cmax}]g_o$

The temperature response of the model parameters V_{cmax}, Γ_*, K_{mc}, K_{mo} and R_d are described, using a general Arrhenius equation:

$$parameter = \exp[c - \Delta H_a/(RT_k)] \tag{6}$$

where R is the molar gas constant and T_K is the leaf temperature in kelvin, c and ΔH_a are scaling constant and activation energy, respectively.

As constants associated with the kinetic properties of Rubisco (i.e. K_{mc}, K_{mo}, Γ_*) are generally conservative for most higher terrestrial C_3 plants (von Caemmerer, 2000; Bernacchi *et al.*, 2001), most parameters values used in the physiological model, eqns (4)-(6), were derived from literature. However, V_{cmax} and a_1 were estimated from curve fitting to our measurements for each stage × treatment combination: i.e. flowering - drought stressed environment (FS); flowering - well watered environment (FW); grain filling - drought stressed environment (GS); grain filling - well watered environment (GW). All these parameters are given in Table S1 (see Supplementary materials). Using this model, measured values for A were normalized to the mean value of observed VPD and T_{leaf} for each stage × treatment combination (Fig. S1).

Construction of marker linkage map

The initial skeleton linkage map was constructed using MAPMAKER/EXP3.0 (Lincoln *et al.*, 1993), based on a RIL population derived from the same parents (La 2004). New polymorphic SSRs were also identified for our IL population. To assign all markers (including those initially identified in the RIL population) into linkage

groups, we also took into account the ultra-dense SSR linkage map of Temnykh *et al.* (2000) and McCouch *et al.* (2002), which contains SSRs identified in our population. Their map was used as the reference to estimate marker distances, the length of chromosomes and of introgressed segments for our IL population, based on the co-linearity of markers across populations (e.g. Shen *et al.*, 2004; Wu & Huang, 2007).

QTL mapping

Significance in the difference for each trait among the ILs were tested ($P < 0.05$), and both simple and partial correlations among all the traits were estimated using SAS 9.13 to assist the analysis of any co-locations of the QTLs for various traits.

Chromosomal locations of putative QTLs for each trait were determined first by single-point analysis using the general linear model (GLM) procedure in SAS. The one-way analysis of variance (ANOVA) was used to test the significance ($P < 0.01$) of association at each locus between two genotype groups (homozygous allele from Shennong265 vs. that from Haogelao). Multivariate analysis of variance (MANOVA) with the PROC GLM in SAS was performed to calculate the total phenotypic variance explained by the identified QTLs of the same trait by using genotype data of the corresponding markers.

To improve the reliability of QTL analysis, we also used MapQTL 6 software (van Ooijen, 2009) to perform so-called composite interval mapping (or MQM in MapQTL 6) (Jansen, 1995). We followed the procedure described by Yin *et al.* (2005b). The threshold of QTL detection for each trait was based on 1000 permutation tests at the 5% level of significance in MapQTL 6. Regions with LOD score values between 2.0 and the calculated threshold were considered as suggestive QTLs (Lander & Kruglyak, 1995), once a suggestive region was approved by single point analysis.

Confirmation of an important QTL

From the field experiment, we identified a QTL around marker RM410 on Chromosome 9, that showed a consistent effect across treatments and stages for a number of the traits (see Results). Therefore, during the summer of 2010, at the research facility UNIFARM, Wageningen, plants of the introgression line (IL161) which only contains a small segment around marker RM410 from the donor parent Haogelao, as well as the recurrent parent Shennong265, were grown in the greenhouse under controlled-environment conditions, to validate the QTL expression under an independent condition. In the greenhouse, temperature was set at 26°C for the 12 h light period and 23℃ for the 12 h dark period. The CO_2 level was about 370 μmol mol^{-1}, the relative humidity was set at 65%, and extra SON-T light (providing extra PPFD ~300 μmol m^{-2} s^{-1}) was switched on when solar radiation intensity outside the

greenhouse was below 400 μmol m^{-2} s^{-1}. Sixteen plants of both genotypes were grown in hydroponic culture by using half Hoagland's solution. One week before flowering, water stress was introduced by adding 12.5% polyethylene glycol (PEG-8000) (stressed condition) or not (non-stressed condition). At flowering stage and grain filling stage, gas exchange and chlorophyll fluorescence parameters were measured on four plants (two measurements per plant) of each treatment by Li-Cor 6400. All measurements were made at a photon flux density of 1000 μmol m^{-2} s^{-1}, ambient CO_2 concentration, VPD of 1.0 to 1.6 kPa, and T$_{leaf}$ of 25℃.

RESULTS

Using statistical and physiological models to correct trait values
The environmental variables, VPD and T$_{leaf}$, fluctuated during measurements of the large set of genotypes, especially for FS (Fig. 1), VPD ranged from 0.96 to 4.75 kPa, T$_{leaf}$ ranged from 23.3 to 36.0℃. Therefore we used a statistical covariant model, eqn (1), to adjust trait values to the mean VPD and T$_{leaf}$ values for each stage × treatment combination. The model analysis showed that VPD had a stronger influence on trait values than T$_{leaf}$ did. The analysis also showed that all the trait values differed significantly among the ILs for each stage × treatment combination ($P < 0.01$).

Next, we used a physiological model, eqn (2-6), to validate the covariant model by adjusting all net photosynthesis values (*A*) to the mean VPD and T$_{leaf}$ of each stage × treatment combination. The results showed a tight correlation between the statistically corrected *A* and the physiologically corrected *A* ($R^2 = 0.92$ for FS, $R^2 = 0.92$ for FW, $R^2 = 0.99$ for GS, $R^2 = 0.99$ for GW) (Fig. S2). Further analysis using the physiological model showed that the physiologically corrected *A* using both VPD and T$_{leaf}$ closely correlated with *A* adjusted using VPD alone (Fig. S3), confirming that VPD was the more important factor as also indicated by the statistical model. This was probably because T$_{leaf}$ during measurements fluctuated only around the optimum value for photosynthesis (Fig. S1b,d,f,h), so the effect of the fluctuation on the traits, if any, was only marginal.

Phenotypic evaluations
Mean values, standard deviations, ranges, skewness, and kurtosis of all adjusted traits are shown in Table 1. All traits showed continuous distribution in the population and almost all showed a normal distribution with low levels of skewness and kurtosis. Compared with the two parents, ILs showed a larger range of variation (Table 1), indicating an obvious transgressive segregation. Among the traits, the relative range of variation (i.e. CV in Table 1) in stomatal conductance for FS was the largest, while

Table 1. Statistics of photosynthesis-related traits of two parents and the population of introgression lines after adjusting to the mean VPD for each stage × treatment combination.

	Traits	Unit	Haogelao	Shennong 265	Introgression lines Mean	CV(%)	Range	Skewness	Kurtosis
FS	A	μmol CO_2 m^{-2} s^{-1}	11.8	11.9	11.2	13.8	7.9-14.7	-0.06	-0.49
	g_s	mol m^{-2} s^{-1}	0.091	0.094	0.089	17.5	0.061-0.128	0.11	-0.85
	T_r	mmol H_2O m^{-2} s^{-1}	3.20	3.60	3.28	11.9	2.45-4.40	0.11	-0.20
	C_i	μmol CO_2 mol^{-1}	244	248	247	6.4	206-295	0.17	0.14
	TE	mmol CO_2 (mol H_2O)$^{-1}$	3.72	3.51	3.50	9.7	2.65-4.49	0.08	-0.03
	Φ_{PSII}	mol e$^-$ (mol photon)$^{-1}$	0.245	0.249	0.245	9.4	0.181-0.286	-0.39	-0.10
	qP	-	0.513	0.515	0.519	10.6	0.381-0.616	-0.43	-0.44
	$F'_\sqrt{}/F'_m$	mol e$^-$ (mol photon)$^{-1}$	0.482	0.491	0.478	5.0	0.405-0.533	-0.16	0.29
FW	A	μmol CO_2 m^{-2} s^{-1}	17.6	15.4	17.1	8.6	14.3-20.6	0.15	-0.53
	g_s	mol m^{-2} s^{-1}	0.153	0.140	0.158	8.9	0.130-0.191	0.15	-0.65
	T_r	mmol H_2O m^{-2} s^{-1}	4.81	4.42	4.97	8.6	3.83-5.97	0.02	-0.27
	C_i	μmol CO_2 mol^{-1}	263	271	269	2.9	251-290	0.11	-0.12
	TE	mmol CO_2 (mol H_2O)$^{-1}$	3.65	3.46	3.50	6.9	2.95-4.06	-0.02	-0.07
	Φ_{PSII}	mol e$^-$ (mol photon)$^{-1}$	0.306	0.281	0.289	5.6	0.246-0.324	-0.06	-0.19
	qP	-	0.593	0.528	0.541	6.3	0.450-0.614	-0.38	0.27
	$F'_\sqrt{}/F'_m$	mol e$^-$ (mol photon)$^{-1}$	0.519	0.531	0.540	3.6	0.500-0.585	0.24	-0.71
GS	A	μmol CO_2 m^{-2} s^{-1}	12.6	12.6	13.7	12.1	9.85-17.3	0.00	-0.75
	g_s	mol m^{-2} s^{-1}	0.134	0.121	0.127	15.6	0.088-0.181	0.43	-0.04
	T_r	mmol H_2O m^{-2} s^{-1}	3.82	3.39	3.68	14.8	2.63-5.37	0.36	-0.05
	C_i	μmol CO_2 mol^{-1}	284	272	268	4.1	245-293	0.03	-0.71
	TE	mmol CO_2 (mol H_2O)$^{-1}$	3.23	3.75	3.82	9.5	2.86-4.53	-0.17	-0.38
	Φ_{PSII}	mol e$^-$ (mol photon)$^{-1}$	0.246	0.247	0.267	7.7	0.219-0.313	-0.23	-0.43
	qP	-	0.466	0.455	0.531	7.6	0.424-0.615	-0.52	0.20
	$F'_\sqrt{}/F'_m$	mol e$^-$ (mol photon)$^{-1}$	0.532	0.543	0.509	5.2	0.442-0.554	-0.44	-0.35
GW	A	μmol CO_2 m^{-2} s^{-1}	16.0	18.4	16.2	9.7	11.5-20.4	0.32	0.65
	g_s	mol m^{-2} s^{-1}	0.165	0.180	0.152	10.9	0.107-0.186	0.21	-0.27
	T_r	mmol H_2O m^{-2} s^{-1}	6.03	6.30	5.44	11.0	3.99-6.73	0.22	-0.47
	C_i	μmol CO_2 mol^{-1}	285	272	270	3.1	252-287	-0.14	-0.78
	TE	mmol CO_2 (mol H_2O)$^{-1}$	2.65	2.95	3.02	8.1	2.47-3.62	0.18	-0.45
	Φ_{PSII}	mol e$^-$ (mol photon)$^{-1}$	0.292	0.281	0.278	8.0	0.212-0.343	-0.21	0.81
	qP	-	0.575	0.526	0.539	9.7	0.374-0.650	-0.53	0.15
	$F'_\sqrt{}/F'_m$	mol e$^-$ (mol photon)$^{-1}$	0.514	0.542	0.521	5.4	0.437-0.634	0.55	2.74
F	DS	-	0.672	0.773	0.653	13.4	0.451-0.826	-0.13	-0.68
G	DS	-	0.792	0.685	0.852	12.7	0.633-0.990	0.95	0.75

CV = coefficient of variation. For other definitions see the SYMBOLS AND ABREVIATIONS.

that in the intercellular CO_2 partial pressure for FW was the smallest. Between the drought environments, the ranges of variation under drought stress were relatively larger than those in the well watered environment, especially at flowering.

Simple and partial correlations for traits associated with gas exchange, chlorophyll fluorescence parameters and transpiration efficiency (TE) are given in bottom left and top right corners of Table 2, respectively, for each stage × treatment combination. In the simple correlation analysis, net photosynthesis (*A*) significantly correlated with all gas exchange and chlorophyll fluorescence parameters, except TE at grain filling, presumably reflecting the fact that photosynthesis is a complex trait associated with a number of physical and chemical reactions.

The partial correlation coefficient between g_s and *A* changed from 0.14 in the well watered environment to 0.57 in the drought stressed environment at flowering, and from 0.33 in the well watered environment to 0.58 in the drought stressed environment during grain filling. This shows the direct effect of drought stress, as the CO_2 availability decreased because of diffusional limitation through stomatal closure. The significant negative correlations between C_i and *A* in both the simple and partial correlation analyses were also supported by Fick's first law of diffusion for CO_2 transfer along the path from C_a to C_i: $C_i = C_a - A/g_s$.

There were tight correlations between the various chlorophyll fluorescence parameters (Φ_{PSII}, qP, $F'_\sqrt{}/F'_m$) in the partial correlation analysis (Table 2). These tight correlations may reflect that Φ_{PSII} is quantitatively restricted by both qP and $F'_\sqrt{}/F'_m$, i.e. $\Phi_{PSII} = qP \times F'_\sqrt{}/F'_m$.

Both the correlations between *A* and other gas exchange parameters (g_s, T_r, C_i, TE) and the correlations between *A* and chlorophyll fluorescence parameters (Φ_{PSII}, qP, $F'_\sqrt{}/F'_m$) were significant, except between *A* and qP in GS (Table 2). Compared with using only gas exchange systems, we can still get more information from combining both gas exchange and chlorophyll fluorescence data, especially on the genetic diversity in electron transport components related to photosynthesis.

Construction of genetic linkage map

To obtain SSR markers showing polymorphism between Haogelao and Shennong265, we surveyed more than 1000 SSRs and found 288 polymorphic markers. Among them 130 SSR markers were evenly distributed across the genome, and were therefore chosen to construct the linkage map. The total length of the linkage map was 1645.1 cM, with an average marker spacing of 12.65 cM (Fig. 2). A graphical representation of the 130 SSRs showed that these introgression lines covered the whole genome of donor parent Haogelao (Gu, 2007).

Table 2. Simple and partial correlation coefficients for traits associated with gas exchange, chlorophyll fluorescence parameters and water use efficiency.

		A	g_s	T_r	C_i	TE	Φ_{PSII}	qP	F'_v/F'_m
FS	A		0.57***	0.65***	-0.29**	0.38***	0.23*	-0.19	-0.12
	g_s	0.55***		-0.01	0.14	-0.35***	-0.41***	0.41***	0.40***
	T_r	0.69***	0.79***		0.05	-0.44***	0.13	-0.13	-0.07
	C_i	-0.58***	0.22*	0.08		-0.68***	0.04	-0.06	-0.01
	TE	0.48***	-0.34***	-0.21*	-0.94***		-0.06	0.06	0.11
	Φ_{PSII}	0.65***	0.13	0.24*	-0.66***	0.59***		0.98***	0.92***
	qP	0.32**	-0.11	-0.08	-0.58***	0.53***	0.88***		-0.96***
	F'_v/F'_m	0.48***	0.52***	0.60***	0.05	-0.07	-0.06	-0.52***	
FW	A		0.14	0.80***	-0.08	0.46***	0.18	-0.15	-0.11
	g_s	0.71***		0.41***	-0.21*	-0.22*	-0.12	0.08	0.06
	T_r	0.74***	0.95***		0.11	-0.38***	-0.08	0.10	0.12
	C_i	-0.30**	0.35***	0.39***		-0.82***	0.01	-0.02	-0.03
	TE	0.23*	-0.43***	-0.47***	-0.98***		-0.07	0.07	0.06
	Φ_{PSII}	0.65***	0.34***	0.42***	-0.29**	0.25*		0.99***	0.95***
	qP	0.31**	0.03	0.08	-0.30**	0.29**	0.82***		-0.98***
	F'_v/F'_m	0.44***	0.47***	0.49***	0.09	-0.13	0.05	-0.52***	
GS	A		0.58***	0.68***	-0.29**	0.49***	0.13	-0.10	-0.06
	g_s	0.70***		0.13	0.30**	-0.22*	-0.24*	0.22*	0.19
	T_r	0.78***	0.94***		-0.01	-0.52***	0.11	-0.10	-0.08
	C_i	-0.20*	0.50***	0.40***		-0.64***	0.12	-0.09	-0.09
	TE	0.10	-0.59***	-0.50***	-0.96***		0.06	-0.04	-0.05
	Φ_{PSII}	0.43***	-0.02	0.10	-0.45***	0.46***		0.99***	0.97***
	qP	-0.06	-0.38***	-0.34***	-0.43***	0.48***	0.75***		-0.97***
	F'_v/F'_m	0.67***	0.56***	0.63***	0.02	-0.10	0.26*	-0.43***	
GW	A		0.33**	0.60***	-0.45***	0.22*	0.25*	-0.20	-0.17
	g_s	0.74***		0.51***	0.01	-0.17	-0.21	0.15	0.17
	T_r	0.74***	0.98***		0.28**	-0.21*	0.04	-0.03	-0.04
	C_i	-0.21*	0.46***	0.48***		-0.74***	0.15	-0.12	-0.10
	TE	0.09	-0.57***	-0.58***	-0.98***		0.07	-0.05	-0.04
	Φ_{PSII}	0.58***	0.20*	0.25*	-0.35***	0.31**		0.99***	0.94***
	qP	0.28**	-0.08	-0.03	-0.37***	0.37***	0.87***		-0.98***
	F'_v/F'_m	0.33***	0.47***	0.44***	0.20	-0.25*	-0.12	-0.59***	

*$P<0.05$; **$P<0.01$; ***$P<0.001$.*

The simple and partial correlation coefficients are listed in the bottom left and top right corners, respectively.

Detection of QTLs

Quantitative trait locus analysis for various traits was conducted separately for the four stage × treatment combinations, by using both single point analysis and MQM. In total we detected 29 QTLs including those 'suggestive' QTLs: 8 QTLs for FS, 8 QTLs for FW, 7 QTLs for GS, 3 QTLs for GW, and 3 extra QTLs for drought sensitivity. QTLs were detected for all traits except for C_i and TE. The total fraction of the phenotypic variation explained by QTLs using genotype data of the marker at each putative QTL (single point analysis) ranged from 7.0 to 37.2%. The results are summarized in Table 3 and Fig. 2. The most significant QTLs (i.e., QTLs with LOD scores higher than the permutation calculation) were marked in bold in Table 3.

Net photosynthesis rate (A)

QTLs controlling net photosynthesis are located on chromosomes 2, 3, 7, 8 and 9. *A* values adjusted using the physiological model identified virtually the same QTLs (Fig. 2), again validating the statistical covariant analysis. The phenotypic variance explained by individual QTLs varied from 7.5 to 18.2%. The additive effect ranged from -0.92 to 1.35 μmol m^{-2} s^{-1}. On chromosome 9 near marker RM410, there was a QTL for all the four stage × treatment combinations, with a consistent positive additive effect ranging from 0.55 to 0.77 μmol m^{-2} s^{-1}.

Stomatal conductance (g_s)

Three QTLs associated with g_s were detected on chromosome 3 and 9 for FS, and on chromosome 6 for GS. The phenotypic variance explained by these three QTLs ranged from 9.5 to 13.5%. No QTLs were detected for well watered conditions (FW and GW). This difference between the well watered and stressed conditions was also shown in Table 1, which shows that the coefficient of variance of g_s changed from 8.9 to 17.5% and from 10.9 to 15.6%, when comparing well watered with stressed conditions at flowering and grain filling, respectively.

Transpiration rate (T_r)

A QTL interval was detected near marker RM410 on chromosome 9 for both FS and GS. The interval contributed to an increase of transpiration with an additive effect of 0.152 and 0.244 mmol m^{-2} s^{-1}, for FS and GS, respectively. The phenotypic variances explained were 10.8-11.4%. Given a tight correlation between T_r and g_s, the reason for no QTLs detected for g_s in well watered conditions also applied for T_r.
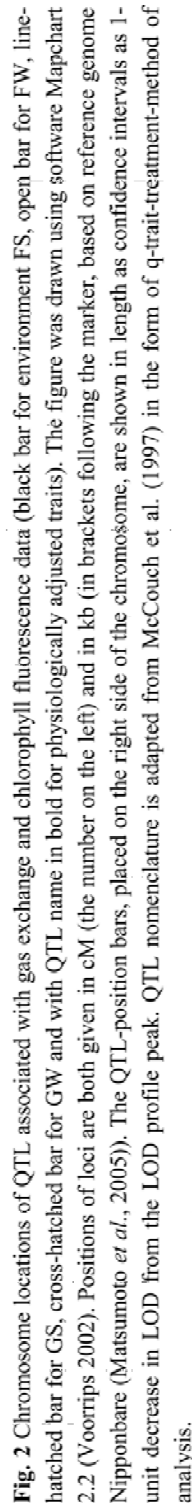
Fig. 2 Chromosome locations of QTL associated with gas exchange and chlorophyll fluorescence data (black bar for environment FS, open bar for FW, line-hatched bar for GS, cross-hatched bar for GW and with QTL name in bold for physiologically adjusted traits). The figure was drawn using software Mapchart 2.2 (Voorrips 2002). Positions of loci are both given in cM (the number on the left) and in kb (in brackets following the marker, based on reference genome Nipponbare (Matsumoto *et al.*, 2005)). The QTL-position bars, placed on the right side of the chromosome, are shown in length as confidence intervals as 1-unit decrease in LOD from the LOD profile peak. QTL nomenclature is adapted from McCouch et al. (1997) in the form of q-trait-treatment-method of analysis.

28

Table 3. QTLs identified for gas exchange and chlorophyll fluorescence parameter traits (see Table 1 for their abbreviations) in ILs from the cross Shennong 265 × Haogelao under well watered and drought stressed environments at both flowering and grain filling stages.

Traits	Stage by treatment	QTL identification	Chr	GLM/SAS marker	*P*-value	R^2 (%)	Global R^2	MQM Position (cM)	LOD	R^2 (%)	*a*
A	FS	qA_FS_MQM_1	2	RM406	0.0016	11.9	30.4	180	2.26	8.1	-0.9204
		qA_FS_MQM_2	**9**	**RM410**	**<0.0001**	**22.4**		**67.4**	**4.79**	**18.2**	**0.7725**
	FW	qA_FW_MQM_1	7	RM432	0.0096	6.9	28.5	44.5	2.45	8.3	0.8307
		qA_FW_MQM_2	9	RM5799	0.0098	3.7		0.8	2.22	7.5	-0.6558
		qA_FW_MQM_3	**9**	**RM410**	**0.0026**	**9.4**		**63.3**	**4.23**	**15.1**	**0.6408**
	GS	qA_GS_MQM_1	8	RM1235	0.0029	9.6	15.6	10.7	2.46	10.5	-0.8627
		qA_GS_MQM_2	9	RM410	0.0055	8.0		57.3	2.27	9.4	0.7092
	GW	**qA_GW_MQM_1**	**3**	**RM5178**	**0.0111**	**6.3**	**16.7**	**23.15**	**3.60**	**14.2**	**1.3485**
		qA_GW_MQM_2	**9**	**RM410**	**0.0008**	**11.5**		**64.3**	**2.85**	**12.2**	**0.5584**
g_s	FS	qGs_FS_MQM_1	3	RM338	0.0086	5.9	17.6	111.4	2.30	9.5	0.0144
		qGs_FS_MQM_2	**9**	**RM410**	**0.0009**	**11.1**		**68.4**	**3.30**	**13.8**	**0.0069**
	GS	qGs_GS_MQM_1	6	RM276	0.0055	8.4	8.4	47	2.28	10.5	-0.0102
T_r	FS	qTr_FS_MQM_1	9	RM410	0.0006	11.9	11.9	67.4	2.47	11.4	0.1515
	GS	qTr_GS_MQM_1	9	RM410	0.0040	8.6	8.6	58.3	2.34	10.8	0.2436
Φ_{PSII}	FS	qQy_FS_MQM_1	1	RM9	0.0095	7.0	7.0	94.4	2.22	10.3	0.0111
	FW	qQy_FW_MQM_1	9	RM410	0.0067	7.7	10.8	58.3	2.09	8.8	0.0066
		qQy_FW_MQM_2	11	RM1761	0.0098	7.2		0.3	2.03	8.6	0.0085
	GW	qQy_GW_MQM_1	11	RM1761	0.0084	7.5	7.5	7.3	2.35	10.9	0.0155
qP	FW	qqP_FW_MQM_1	1	RM8051	0.0097	6.3	37.2	54.4	2.26	8.1	0.0218
		qqP_FW_MQM_2	1	RM1198	0.0027	9.3		146.4	2.43	9.0	-0.0111
		qqP_FW_MQM_3	**11**	**RM1761**	**0.0047**	**8.5**		**0.3**	**3.71**	**14.2**	**0.0229**
	GS	qqP_GS_MQM_1	6	RM276	0.0003	13.8	13.8	41.3	2.29	10.6	0.0181
F'_v/F'_m	FS	**qMeo_FS_MQM_1**	**6**	**RM6836**	**0.0003**	**13.5**	**23.8**	**55.1**	**3.14**	**12.9**	**-0.0122**
		qMeo_FS_MQM_2	9	RM410	0.0016	10.2		64.4	2.33	9.3	0.0074
	GS	**qMeo_GS_MQM_1**	**4**	**RM2799**	**0.0021**	**11.8**	**21.0**	**123.8**	**3.58**	**15.1**	**-0.0119**
		qMeo_GS_MQM_2	8	RM1381	0.0097	6.2		2.9	2.05	8.3	-0.0120
DS	F	qDS_F_MQM_1	2	RM406	0.0010	7.0	7.0	174	2.62	11.5	-0.0640
	G	**qDS_G_MQM_1**	**2**	**RM6911**	**0.0003**	**13.1**	**28.1**	**39.2**	**5.44**	**20.3**	**-0.0810**
		qDS_G_MQM_2	**8**	**RM1381**	**0.0015**	**10.3**		**1.9**	**3.06**	**10.9**	**-0.0525**
	G	qDS_G_MQM_1*2[e]	2*8	RM6911_1381	0.0089	5.9					

P-value, the significance of phenotypic variation associated with markers in single-point analysis; R^2, the individual contribution of one QTL to the variation in a trait. Global R^2, the fraction of the total variation explained by the QTLs of the same trait; Position, position of maximum LOD; LOD, logarithm of odds; *a*, additive allelic value of Haogelao; [e], epistatic interaction between two markers. QTLs with LOD scores higher than the threshold set by 1000 permutation tests at 5% level of significance were marked in bold.

Quantum yield of PSII (Φ_{PSII})

QTLs, located on chromosomes 1, 9 and 11, were found for quantum yield of PSII. The locus near maker RM1761 on chromosome 11 was consistently detected for both FW and GW. The phenotypic variance explained by these QTLs ranged from 8.6 to 10.9% with a consistently positive effect.

Proportion of open PSII (qP)

Four QTLs associated with proportion of open PSII, located on chromosomes 1, 6 and 11 were detected. The direction of their effects was positive except the QTL located on chromosome 1 at 146.6 cM. Individual loci explained between 8.1 and 14.2% of the phenotypic variance, and the additive effect varied from -0.011 to 0.0218.

Maximum efficiency of open PSII in the light ($F'_{\surd}/F'_{m}$)

Four QTLs related to $F'_{\surd}/F'_{m}$ were detected on chromosomes 4, 6, 8 and 9 in the drought stressed conditions, while no QTLs were found in the well watered conditions. The total phenotypic variance explained by QTLs was 23.8% at flowering and 21.0% at grain filling. The phenotypic variance explained by individual QTLs varied from 8.3 to 15.1% with a negative effect, except for the QTL located on chromosome 9 near marker RM410.

Drought sensitivity (DS)

As indicated, DS was calculated as $A_{drought}:A_{water}$, which can characterize the relative responsiveness of each genotype to a decline in water availability. In total, three QTLs were found, one at flowering, two at grain filling. The QTLs *qDS_F_MQM_1 and qDS_G_MQM_2* coincided with QTLs of *A*: *qA_FS_MQM_1* at FS and *qA_GS_MQM_1* at GS, respectively. The coincidences were expected because these loci were expressed only at one of the treatments. However, we found a new QTL with a relatively large effect for grain filling stage (not detected for *A* at either treatments at this stage), on chromosome 2 with an additive effect of -0.081 on DS.

Verification of a QTL on Chromosome 9 in a controlled greenhouse environment

The above QTL analysis showed that the QTL near RM410 on chromosome 9 had a significant multiple effect on *A*, g_s, T_r, Φ_{PSII} and $F'_{\surd}/F'_{m}$ across development stages and treatments. In order to assess whether the effect shown by chromosome 9 is independent and whether there is any epistasis between identified QTLs, ANOVA by PROC GLM was used to evaluate epistatic interactions between pairs of QTLs, as represented by the nearest marker loci (Lin *et al*., 2000). There was no significant epistatic interaction found, except for DS at grain filling stage (Table 3). In our IL

population, we found an introgression line, IL161, which had the background of recurrent parent Shennong265 but for a small introgression segment containing marker RM410 (Fig. 3a); so, IL161 could serve as near-isogenic line of Shennong265. To further validate the QTL near RM410, we measured photosynthetically associated traits for Shennong265 and IL161 in a greenhouse experiment. The *A* of IL161 was consistently higher than that of Shennong265 (Fig. 3b), confirming the positive effect of the allele from Haogelao at the locus on *A*. This difference was significant ($P <$ 0.05) at FS and FW, which was also supported by the high LOD scores of locus *qA_FS_MQM_2* and *qA_FW_MQM_2*, respectively. The difference was insignificant ($P > 0.05$) at GS and GW, respectively, partly in line with the comparatively small additive effects and low LOD scores of locus *qA_GS_MQM_2* and *qA_GW_MQM_2*. For g_s, T_r, Φ_{PSII} and $F'_\sqrt/F'_m$ the corresponding QTLs *qGS_FS_MQM_2*, *qTr_FS_MQM_1*, *qQy_FW_MQM_1* and *qMeo_FS_MQM_2* were validated by the significant difference between IL161 and Shennong265 except for T_r at GS (Fig. 3c-f).

DISCUSSION

In this study, we aimed to identify QTLs for photosynthetic parameters of rice under drought and well watered field conditions during flowering and mid grain filling stage. Because of the limited range of genetic variation and high sensitivity to environmental perturbations, photosynthetic traits were known so far not to be amenable to QTL analysis. We, therefore, used two strategies to enhance QTL mapping precision: using both a statistical and a physiological approach to adjust phenotypic trait values for micro-climatic differences during measurements in the field, and using an advanced backcross IL population. The identified QTLs tended not only to cluster in the rice genome, but also to consistently be expressed over both development stages and both drought stressed and well watered conditions (Fig. 2).

Complexity of photosynthetic traits

Photosynthesis as a dynamic process, continuously interacts with environment. Because of micro-climate fluctuations, it is difficult to phenotype photosynthesis in the field for a large set of genotypes (Flood *et al.*, 2011). We used a covariant model which normalized all measurements to the mean VPD, because VPD has a dominant effect on g_s and photosynthesis (Ball *et al.*, 1987; Leuning, 1990, 1995). This dominant effect of VPD, relative to T_{leaf}, was confirmed by our statistical analysis. Bernacchi *et al.* (2001), however, demonstrated that T_{leaf} influenced many aspects of the biochemical and biophysical reactions which determine the rate of photosynthesis. Using a physiological model, we could separate the mixed effects of VPD and T_{leaf} (Fig. S1). A sensitivity analysis with and without considering temperature effect
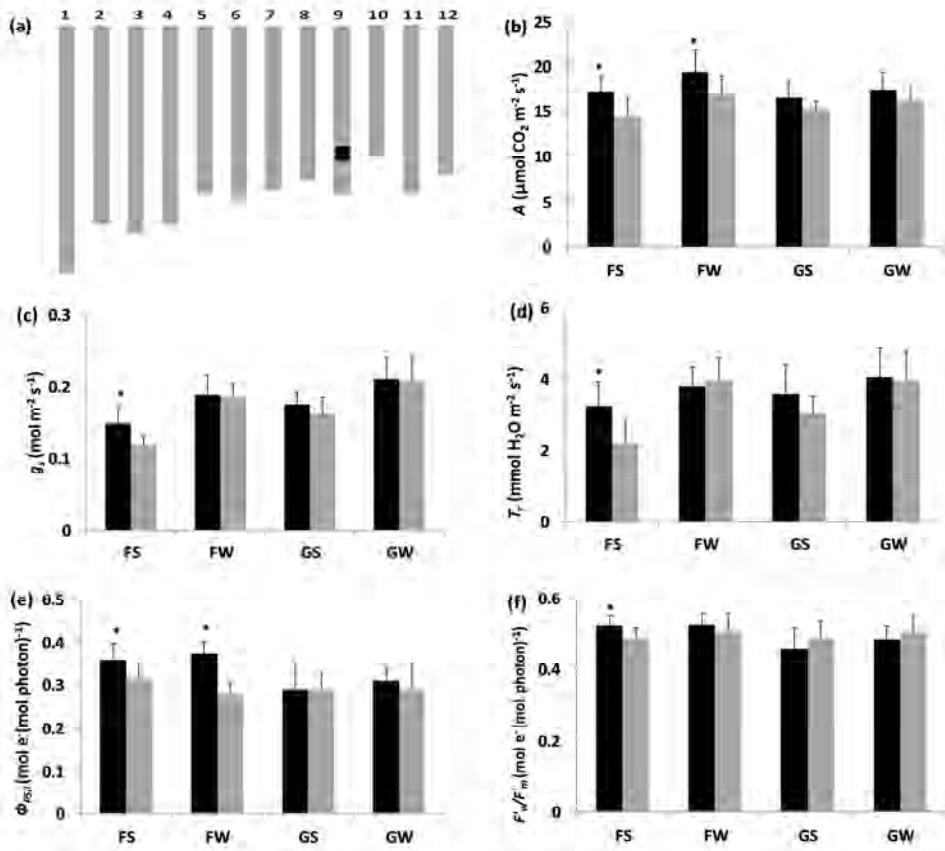
Fig. 3 Confirmation of a QTL on Chromosome 9. **(a)** Graphical representation of genotypes of IL161. *Gray blocks*, chromosome regions homozygous for Shenong265; *Black block*, chromosome region introgressed from Haogelao. The graphical genotypes shown here are based on the physical map by Matsumoto *et al.* (2005). (b)-(f) Comparisons of photosynthetic traits between Shennong265 (gray column) and IL161 (black column) in 2010: **(b)** net photosynthesis A; **(c)** stomatal conductance for CO_2 g_s; **(d)** transpiration rate T_r; **(e)** quantum yield of PSII Φ_{PSII}; **(f)** maximum efficiency of open PSII in the light F'_v/F'_m. * indicates significant differences at $P<0.05$ between IL161 and Shennong265.

showed that A was little affected by T_{leaf} but strongly by VPD (Fig. S3). Part of reason may be that temperature during measurements varied around the optimum temperature of photosynthesis (~30℃), where the temperature response is less prominent (Fig. S1). Partly because VPD = $e(T_{leaf})$ - e_a, where $e(T_{leaf})$ is the saturation vapor pressure based on T_{leaf}, e_a is the vapor pressure in the ambient air. The equation indicates that T_{leaf} may influence photosynthesis through VPD. Therefore, the physiological model confirmed the covariant model analysis in separating the mixed effects of VPD and T_{leaf} based on solid physiological principles.

Our QTL study using corrected trait values showed that fewer than 35% of the QTLs for chlorophyll fluorescence parameters coincided with those for gas exchange parameters (Table 3). Chlorophyll fluorescence parameters indicate electron transport capacity of photosynthesis (Genty *et al.*, 1989). Different QTLs identified for chlorophyll fluorescence and gas exchange parameters suggest that photosynthetic electron transport and CO_2 fixation are not entirely coupled. The partial uncoupling between the two sets of parameters could be due to the fact that *A* was limited by the Rubisco activity during our measurement conditions (FFPD = 1000 μmol m^{-2} s^{-1}), at this state, part of the electrons were used for processes other than CO_2 fixation (Yin *et al.*, 2006, 2009b). The alternative use of electrons could be especially the case when plants are facing drought stress (Chaves, 1991). During onset of water stress, stomatal aperture will first decrease to reduce the water loss; this sensitivity of stomatal conductance was also shown in our data (Table 1), with more variance in drought conditions than in well watered conditions. Net photosynthesis will be reduced after the stomatal response as a consequence of the reduced C_i. A further complication under drought is the associated increase in T_{leaf}. High temperature will have feedback effects, firstly, by increasing transpiration as a result of an increased VPD at the leaf surface. Secondly, high T_{leaf} may alter the biochemical activity of photosynthetic enzymes (e.g. V_{cmax}). Thirdly, the higher canopy temperature may accelerate ageing of the leaf, thus shortening the growing period. The complex and conflicting responses of gas exchange and electron transport to drought stress mean that physiological knowledge should be incorporated into genetic analysis of photosynthesis.

Merits of IL population
Since Eshed & Zamir (1994) constructed the first complete set of ILs in tomato carrying single *Lycopersicon pennellii* chromosomal segments into a homogeneous background of *Lycopersicon esculentum*, representing the entire wild tomato genome, the ILs also became popular in other crop species such as rice, potato and barley and in the model plant *Arabidopsis*. The ILs are plant series that possess segments of the donor parent chromosome in the background of the recurrent parent. These ILs can be considered similar to a genomic library with genome inserts. The ability to statistically identify small phenotypic effects is increased by the removal of background noise. Also the homozygous lines are immortal, and phenotypic data can be obtained from different environments (e.g., across various years).

In our IL population there was an line, IL161, with only a single desirable segment introgressed, the interval on chromosome 9, with co-location QTLs of *A*, g_s, T_r, Φ_{PSII} and F'_v/F'_m. Because all phenotypic variance between the IL and the recurrent parent (cv. Shennong265) is due to the introgressed segment, we validated the detection of

this QTL by comparing the difference in photosynthetic traits between IL161 and Shennong265 in an independent environment (Fig. 3). Because the greenhouse micro environment variables (temperature, humidity, light intensity, etc.) were controlled at relatively constant levels, the results of the greenhouse experiment for QTL verification implicitly proved the efficacy of our covariant and physiological models in adjusting phenotypic trait values for field micro-climatic differences to have a more accurate QTL analysis.

Cluster of QTLs

The phenomenon of QTL clusters has been observed in different crops, including rice (Xiao *et al.*, 1996), barley (*Hordeum vulgare* L.) (Yin *et al.*, 1999b), wheat (*Triticum aestivum* L.) (Quarrie *et al.*, 2006), cotton (*Gossypium hirsutum* L.) (Shappley *et al.*, 1998), soybean (*Glycine max* L.) (Xu *et al.*, 2011), sorghum (*Sorghum bicolor* (L.) Moench) (Lin *et al.*, 1995) and peach (*Prunus persica* L.) (Quilot *et al.*, 2004). This clustering may be due to the tight linkage of genes or to the pleiotropic effects of a single locus. By using substitution mapping, Monforte & Tanksley (2000) demonstrated in tomato (*Solanum lycopersicum* L.) that a region affecting several agronomically important traits actually resulted from the linkage of multiple QTLs. However, Xue *et al.* (2008) showed that the tight correlation between number of grains per panicle, plant height and heading date was due to the pleiotropic effect of a single QTL *Ghd7*.

In our study, four intervals located on chromosomes 6, 8, 9 and 11 were found to control two or more photosynthetic traits each. Especially in the interval from 57.3 cM to 68.4 cM of chromosome 9 (~2500 kb), QTLs related to A, g_s, T_r, Φ_{PSII} and F'_{ν}/F'_m were clustered and showed the same positive effect from the allele of upland rice Haogelao. The knowledge of the photosynthetic processes indicates chloroplast electron transport rates and carbon metabolism are coupled (at least to some extent), suggesting that pleiotropic effects are likely. A conclusion about whether the clustering is caused by pleiotropy or by gene linkage within these QTL regions cannot be made at this stage. For better characterizing these loci, it is necessary to reduce the extent of introgression and develop near-isogenic lines carrying fine-mapped QTLs.

The clustering of QTLs also indicates the difficulties of manipulating correlated traits simultaneously. For example, TE is an important target for breeding (Xu *et al.*, 2009). From theoretical perspective, Condon *et al.* (2004) indicated that under certain environment conditions, leaf-level TE could be improved by higher photosynthetic potential, lower stomatal conductance, or a combination of these two. However, in our experiment, the clustering of QTLs for A and g_s on chromosome 9 shows that the photosynthesis was improved by keeping stomata more open, resulting in higher C_i,

and a higher photosynthesis (Fig. 3). This association means a higher loss of water at the same time, thereby keeping TE virtually invariant. This could be the reason why there was no QTL found for TE near marker RM410. Again, further analysis based on finer substitution lines might answer the question whether the association among A, g_s and T_r could be broken towards a significantly improved TE.

Marker-assisted selection

The above interesting QTL clusters in the IL population for a number of traits could be explored for further MAS for an improved photosynthetic performance. Especially the QTLs at the RM410 locus was independently confirmed, showing a positive allele from upland rice Haogelao. The upland rice cultivar generally performed better under drought for a number of agronomic traits (La, 2004; Gu, 2007). Our result indicates the possibility of simultaneous improvement of drought tolerance and photosynthetic traits. For breeders, it is interesting to identify co-locations of QTLs, especially when their effects have the same positive direction. This co-location could potentially be used in a breeding program through MAS to combine multiple benefits without negative effects.

The analysis with DS (expressed as $A_{drought}$:A_{water}) identified additional QTLs (Table 3), especially *qDS_G_MQM_1* which has the highest LOD score (5.44) in our study. For breeding one would select for genotypes which have not only high photosynthetic rates but also low photosynthetic sensitivity to drought. QTLs for DS all had negative additive effects (Table 3), indicating alleles from Haogelao were surprisingly associated with high sensitivity to drought. Nevertheless, our ILs are good ready breeding materials which are most alike to the recurrent parents but are further improved by the introgression of desired traits from the donor plant. Further rounds of selection on the basis of these ILs, using the markers associated with the QTLs, could combine favourable alleles of multiple loci into a single genotype.

CONCLUDING REMARKS

This is the first paper using simultaneously measured gas exchange and chlorophyll fluorescence data to intensively study the genetic differences of photosynthesis under field conditions. We also introduced a physiological model to support our covariant model to remove the micro-environment variation noise. Through these approaches we obtained consistent results across environments and growth stages, and observed co-location of physiologically tightly related QTLs. We then successfully confirmed a QTL controlling multiple photosynthetic traits identified under field conditions. In view of climate change (CO_2 enrichment, higher temperatures, and more severe drought stress), photosynthesis as a source of crop production is directly influenced by

these factors, and has also been considered as the only remaining major trait available to further increase crop yield potential (Long *et al.*, 2006; Murchie *et al.*, 2009; Zhu *et al.*, 2010). Fischer & Edmeades (2010) have shown that recent yield progress in cereals from breeding was associated with increased photosynthesis. We expect that photosynthesis will receive an increasing attention in genetic studies and future breeding programmes (e.g. Adachi *et al.*, 2011). Great challenge for drought-prone environments is to increase photosynthetic rate and transpiration efficiency simultaneously. To that end, rich physiological knowledge should be explored to enhance the genetic analysis of the traits of photosynthesis and water use, as already illustrated for other traits (e.g. Yin *et al.*, 1999a; Bertin et al. 2010). Our results highlight that combined physiological and genetic tools can be helpful to improve screening and selection strategies in rice breeding for increased photosynthesis under field conditions.

SYMBOLS AND ABBREVIATIONS

A	Net photosynthesis rate
C_i	Intercellular CO_2 concentration
$F'_\sqrt{}/F'_m$	Maximum efficiency of open photosystem (PS) II in the light
g_s	Stomatal conductance for CO_2
qP	Proportion of open PSII
T_{leaf}	Leaf temperature
T_r	Transpiration rate
Φ_{PSII}	Quantum efficiency of PSII electron transport
DS	Drought sensitivity
TE	Transpiration efficiency
F	Flowering stage
G	Grain filling stage
FS	Flowering stage - drought stressed environment
FW	Flowering stage - well watered environment
GS	Grain filling stage - drought stressed environment
GW	Grain filling stage - well watered environment
PPFD	Photosynthetic photon flux density
QTL	Quantitative trait locus

Supplementary Materials in Chapter 2

Table S1. The values of photosynthetic parameters used in the physiological adjustment. The values of c and ΔH_a describe the temperature responses of CO_2 uptake.

Parameter		Value	c (dimensionless)	ΔH_a (kJ mol^{-1})	Reference
R_d (μmol m^{-2} s^{-1})		1[a]	18.72	46.39	Bernacchi *et al.* (2001)
V_{cmax} (μmol m^{-2} s^{-1})		—	26.35	65.33	Bernacchi *et al.* (2001)
Γ_* (μmol mol^{-1})		42.75[a]	19.02	37.83	Bernacchi *et al.* (2001)
K_{mc} (μmol mol^{-1})		404.9[a]	38.05	79.43	Bernacchi *et al.* (2001)
K_{mo} (mmol mol^{-1})		278.4[a]	20.30	36.38	Bernacchi *et al.* (2001)
g_o (mol m^{-2} s^{-1})		0.01	—	—	Leuning (1995)
D_o (kPa)		0.35	—	—	Leuning (1995)
V_{cmax} (μmol m^{-2} s^{-1})	FS	57	—	—	† (0.86)
	FW	71	—	—	† (0.81)
	GS	60	—	—	† (0.82)
	GW	68	—	—	† (0.73)
a_1	FS	16.41	—	—	† (0.86)
	FW	18.53	—	—	† (0.81)
	GS	17.09	—	—	† (0.82)
	GW	21.06	—	—	† (0.73)

[a], the values of R_d, Γ_*, K_{mc}, K_{mo} were estimated at 25℃;

†, calculated from each stage-treatment combination by fitting Leuning's stomatal-photosynthesis model (Leuning 1990, 1995) with r^2 in brackets.
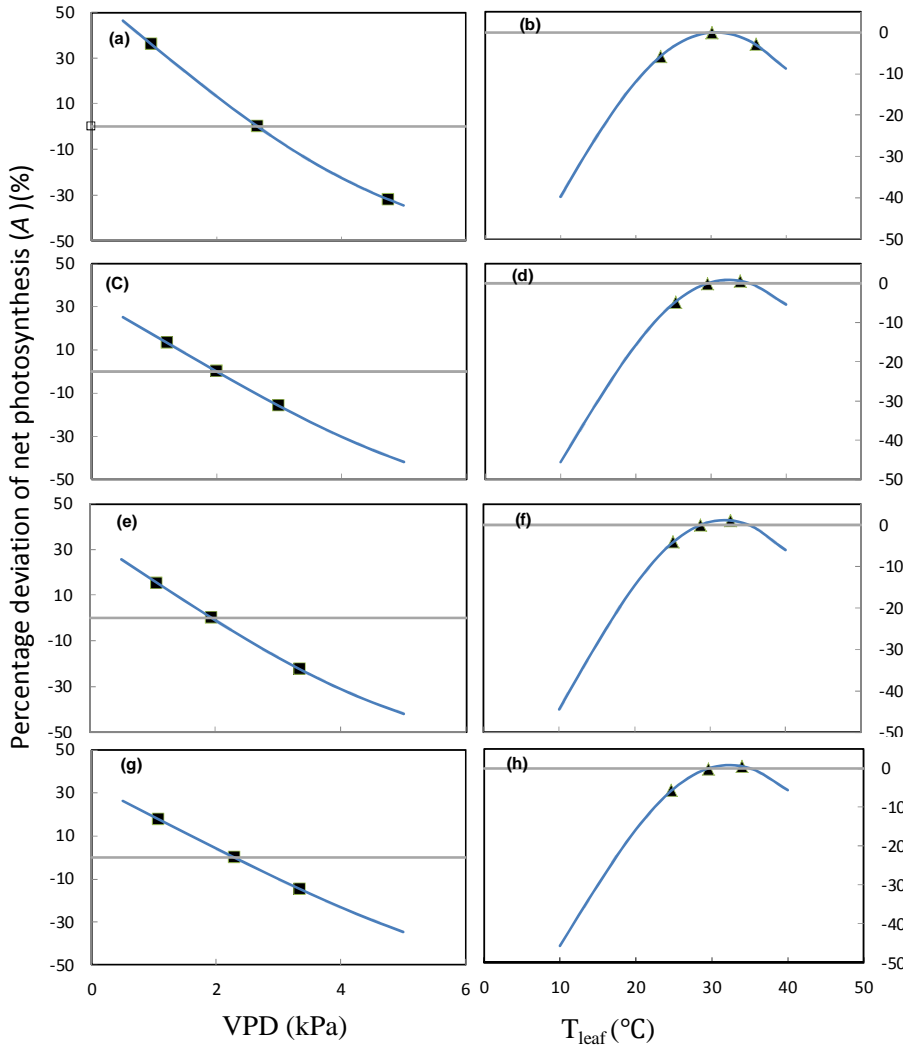
Fig. S1 Percentage deviation of net photosynthesis rate (A) to VPD and T_{leaf} predicted by the physiological model for four different stage-treatment combinations (panel (a), (b) for FS, panel (c), (d) for FW, panel (e), (f) for GS, (g), (h) for GW). Each set of functions to VPD (panel (a), (c), (e), (g)) and temperature (panel (b), (d), (f), (h)) was normalized to unity at average VPD and temperature, respectively. The minimum, mean and maximum value of VPD and temperature in each stage-treatment combination is shown by squares and triangles, respectively. The minimum, mean ($\pm$ SD), maximum value is: FS (VPD (0.96, 2.66 $\pm$ 1.01, 4.75), T_{leaf} (23.3, 30.1 $\pm$ 3.0, 36.0)); FW (VPD (1.21, 2.00 $\pm$ 0.35, 3.00), T_{leaf} (25.3, 29.5 $\pm$ 1.9, 33.9)); GS (VPD (1.07, 1.94 $\pm$ 0.56, 3.35), T_{leaf} (25.1, 28.7 $\pm$ 1.8, 32.7); GW (VPD (1.32, 2.29 $\pm$ 0.47, 3.44), T_{leaf} (24.7, 29.5 $\pm$ 2.37, 34.0).
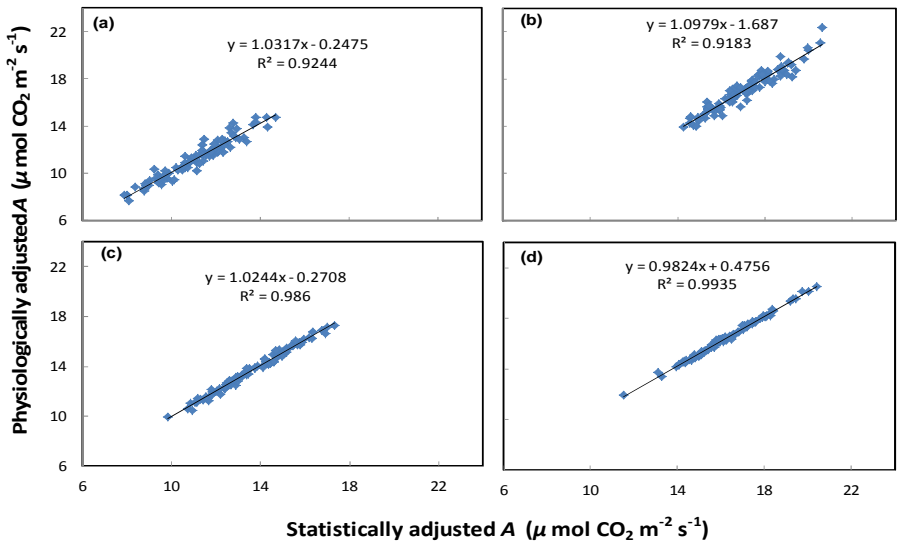
Fig. S2. Comparison between physiologically adjusted net photosynthesis rate (*A*) and the statistically adjusted *A* (panel (a) for FS, panel (b) for FW, panel (c) for GS, panel (d) for GW).
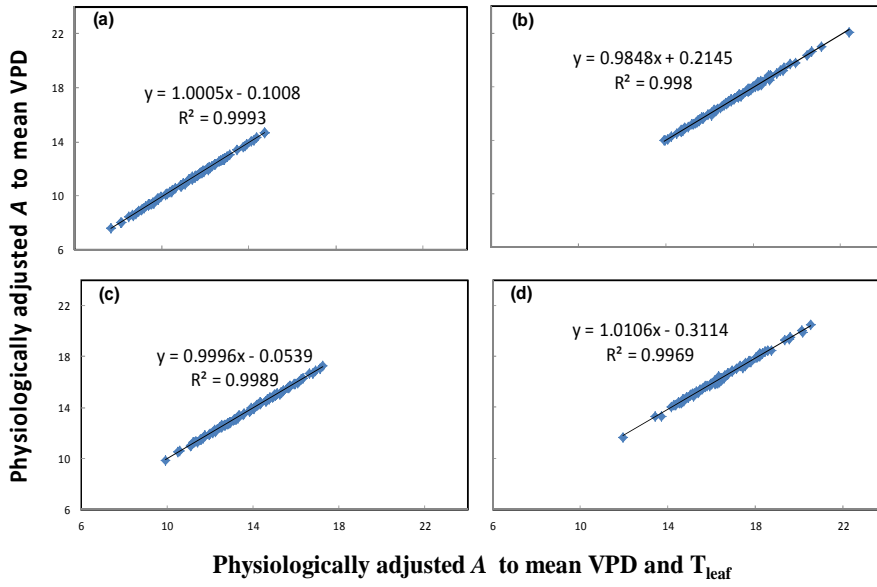
Fig. S3. Comparison between physiologically adjusted net photosynthesis rate (A, μmol CO_2 m^{-2} s^{-1}) to mean VPD and physiologically adjusted A to both VPD and T_{leaf} (panel (a) for FS, panel (b) for FW, panel (c) for GS, and panel (d) for GW).

CHAPTER 3

Physiological basis of genetic variation in leaf photosynthesis among rice (*Oryza sativa* L.) introgression lines under drought and well-watered conditions

Junfei Gu [a], Xinyou Yin [a], Tjeerd-Jan Stomph [a], Huaqi Wang [b], Paul C. Struik [a]

[a] Centre for Crop Systems Analysis, Department of Plant Science, Wageningen University, P. O. Box 430, 6700 AK Wageningen, The Netherlands
[b] Plant Breeding & Genetics, China Agricultural University, 100193 Beijing, P.R. China

ABSTRACT

To understand the physiological basis of genetic variation and resulting QTLs for photosynthesis in a rice (*Oryza sativa* L.) introgression line population, we studied 13 lines under drought and well-watered conditions, at flowering and grain filling. Simultaneous gas exchange and chlorophyll fluorescence measurements were conducted at various levels of incident irradiance and ambient CO_2 to estimate parameters of a model which dissects photosynthesis into stomatal conductance (g_s), mesophyll conductance (g_m), electron transport capacity (J_{max}), and Rubisco carboxylation capacity (V_{cmax}). Significant genetic variation in these parameters was found, although drought and leaf age accounted for larger proportions of the total variation. Genetic variation in light saturated photosynthesis and transpiration efficiency (TE) were mainly associated with variation in g_s and g_m. One previously mapped major QTL of photosynthesis was associated with variation in g_s and g_m, but also in J_{max} and V_{cmax} at flowering. So, g_s and g_m, which were demonstrated in the literature to be responsible for environmental variation in photosynthesis, were found also to be associated with genetic variation in photosynthesis. Furthermore, relationships between these parameters and leaf nitrogen or dry matter per unit area, which were previously found across environmental treatments, were shown valid for variation across genotypes. Finally, we evaluated the extent to which photosynthesis rate and TE can be improved. Virtual ideotypes were estimated to have 17.0% higher photosynthesis and 25.1% higher TE, compared with the best genotype investigated. Our analysis using introgression lines highlights possibilities to improve both photosynthesis and TE within the same genetic background.

Key words: Drought, genetic variation, mesophyll conductance, modelling, *Oryza sativa* L., photosynthesis, rice, stomatal conductance.

INTRODUCTION

The response of leaf photosynthesis to drought involves interactions between physical and metabolic mechanisms (Kramer & Boyer, 1995; Pinheiro & Chaves, 2011). The understanding of these physiological mechanisms is necessary to improve physiological dissection of the complexity of leaf photosynthesis in response to drought (Serraj *et al.*, 2008).

In general, the relationships among leaf photosynthesis (*A*), stomatal conductance (g_s), and transpiration are well understood, as g_s has been studied most when investigating photosynthetic responses to drought (reviewed by Israelsson *et al.*, 2006; Casson & Hetherington, 2010; Lawson *et al.*, 2011). However, g_s is not the only component of CO_2 diffusion in leaves. Mesophyll conductance (g_m), the conductance

from substomatal cavities to the site of carboxylation limits photosynthesis significantly as well, meaning that the CO_2 concentration in chloroplast (C_c) is lower than in intercellular space (C_i) (Lloyd *et al.*, 1992; Warren *et al.*, 2003; Warren, 2004, 2008; Flexas *et al.*, 2008). Ignoring g_m would erroneously attribute the decreased photosynthesis under drought to metabolic impairment (Delfine *et al.*, 1998; Flexas *et al.*, 2004; Centritto *et al.*, 2009).

The value of g_m is influenced by leaf traits such as leaf dry matter per unit area (LMA, Flexas *et al.*, 2008; Galmés *et al.*, 2011), but also by environmental variables, including water status (Delfine *et al.*, 1998; Galmés *et al.*, 2007; Niinemets *et al.*, 2009), temperature (Bernacchi *et al.*, 2002; Scafaro *et al.*, 2011), and nutrient supply (Warren, 2004). There is increasing evidence that g_m and g_s are tightly correlated (e.g. Evans, 1999; Flexas *et al.*, 2007a; Warren, 2008; Yin *et al*., 2009b; Barbour *et al*., 2010; Douthe *et al*., 2011) and follow the same pattern of variation: declining in response to short-term increases of CO_2 partial pressure and increasing with increases in irradiance. So the relationship between g_m and g_s is worthy to be further explored, when assessing genetic variation in leaf photosynthesis. Genetic variation in g_m/g_s ratio will allow breeding for high transpiration efficiency (TE) (Galmés *et al.*, 2011).

Photosynthesis is affected not only by diffusion components (g_s and g_m), but also by various biochemical capacities of protein complexes. The potential activity of Rubisco (V_{cmax}) limits photosynthesis at low C_c. As C_c increases, the chloroplastic electron transport capacity (J_{max}) can limit photosynthesis (Farquhar *et al.*, 1980). Both V_{cmax} and J_{max} are closely related to the amount of leaf nitrogen per unit area (N_a) (Makino *et al.*, 1984, 1985; Evans, 1989; Harley *et al.*, 1992b).

Whilst most studies have focused on photosynthetic responses to environmental factors, significant genotypic variation of *A* has long been reported among species of *Oryza*, and among progeny plants derived from crosses between varieties. For example, variation was observed among varieties of *japonica* rice (Sasaki & Ishii, 1992; Ishii, 1995), and among varieties including *indica* and *japonica* rice and wild rice species (Cook & Evans, 1983; Dingkuhn *et al.*, 1989; Yeo *et al.*, 1994; Peng *et al.*, 1998; Masumoto *et al.*, 2004; Teng *et al.*, 2004). Moreover, quantitative trait loci (QTLs) responsible for the different photosynthetic parameters have been successfully mapped (Zhao *et al.*, 2008; Takai *et al.*, 2009; Xu *et al.*, 2009; Adachi *et al.*, 2011). Recently, we (Chapter 2), using a population of introgression line (ILs) from a cross between upland rice and lowland rice, identified QTLs for light saturated gas exchange and chlorophyll fluorescence parameters under both well-watered and drought conditions in the field. QTLs affecting these parameters tended to cluster in the same genomic regions, suggesting a common genetic basis and inherent physiological connections of photosynthesis parameters.

Few studies have investigated the physiological basis for these reported genetic variations and QTLs of leaf photosynthesis. Taylaran *et al.* (2011) showed that the higher N_a and higher g_s in *indica* cultivars could be the reason for higher *A* than observed in a *japonica* variety. Similarly, Adachi *et al.* (2011) reported that two mapped QTLs of net photosynthesis actually arose from an increased N_a and g_s. Scafaro *et al.* (2011) compared a cultivar of *Oryza sativa* with two wild *Oryza* relatives, and found that the difference in mesophyll cell wall thickness was responsible for differences in g_m, which resulted in substantial variation in *A* between the cultivated and the wild rice. Masle *et al.* (2005) isolated a TE-regulating gene *ERECTA* from a population of *Arabidopsis* and found that V_{cmax}, J_{max}, stomatal density, and mesophyll development caused the genetic variation in TE and *A*.

As a follow-up of our QTL-mapping study (Chapter 2), the present paper aims to investigate the physiological basis of genetic variation and resulting QTLs identified for our IL population. Therefore, we used a model to analyse experimental data for complete curves of photosynthetic responses to CO_2 and to light measured on leaves in a representative subset of the ILs. Such a model analysis allows: (1) to identify the genetic variation in each biophysical and biochemical component, (2) to analyse the physiological basis for the genetic variation in photosynthesis, and (3) to evaluate the potential of utilising the genetic variation in these components for improving *A* and TE under contrasting drought stress. The information obtained could have an important implication for developing drought tolerant varieties.

MATERIAL AND METHODS

Plant growth conditions, treatments and experimental design

A greenhouse experiment was conducted, at the research facility UNIFARM, Wageningen, the Netherlands. Physiological dissection of photosynthesis requires complete curves of responses to various CO_2 and light levels, and it is practically infeasible to experimentally obtain these curves for all individual genotypes of the IL population described in Chapter 2. Eleven lines (IL7, IL37, IL42, IL69, IL84, IL100, IL130, IL157, IL159, IL161, and IL164) and two parents [Shennong265, *japonica*; Haogelao, *indica-japonica* intermediate] were therefore selected. The selection was based on two criteria: (i) the ILs should carry many QTLs to reflect as much as possible the genetic variation of the population; and (ii) the ILs should contain as few chromosome segments from the donor parent as possible, to remove the background noises (see also Eshed & Zamir, 1995). These eleven ILs had on average 6.5% of genome introgressed. Their graphical genotypes are shown in Fig. 1.
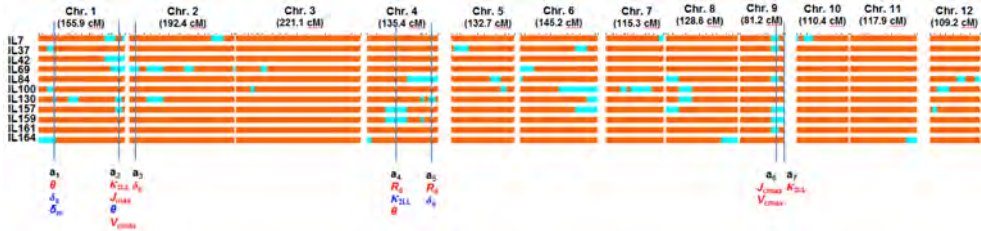
Fig. 1. Graphical genotypes of the eleven introgression lines in this study. The length of each linkage group was shown in centiMorgan (cM). The light-blue region indicate the introgression regions from the donor parent 'Haogelao'; the red regions indicate the homozygous regions from the recurrent parent 'Shennong265'. The figure was drawn using software GGT 2.0 (van Berloo, 2008). The seven regions responsible for variation of photosynthesis parameters (R_d, k_{2LL}, J_{max}, θ, δ_m, V_{cmax} and δ_s, see Table in Appendix for their definition), identified from regression analysis using Eqn (14) with additive effects (i.e. a_1, a_2, a_3, a_4, a_5, a_6 and a_7), are indicated (vertical lines). Parameters, on which genome alleles from 'Haogelao' have positive effects and negative effects, were shown in red and blue colours, respectively.

   Temperature in the greenhouse was set at 26°C for the 12 h light period and at 23°C for the 12 h dark period. The CO_2 level was about 380 μmol mol^{-1}, the relative humidity was set at 65%, and extra SON-T light was switched on when global solar radiation intensity outside the greenhouse was < 400 W m^{-2} and then switched off once it exceeded 500 W m^{-2}. Pre-germinated seeds of the 13 genotypes were sown on sand-beds twice (on the 8th and the 15th, respectively, in June 2010), to extend flowering and grain-filling periods for an enough time window of measurement. Seedlings were then transferred to containers (40 cm long, 30 cm wide, and 20 cm high) in hydroponic culture by using half-strength Hoagland's solution, according to a completely randomized block design. Sixteen plants of each genotype were grown with 7.5×7.5 cm^2 spaces between plants. One week before flowering, eight plants per genotype were exposed to a moderate water stress (comparable to the stress level as in the field experiment in Chapter 2) induced by adding 12.5% polyethylene glycol (PEG-8000) to the growth solution (Money, 1989). The stress was imposed continuously on plants till all measurements were completed. The remaining eight plants per genotype were maintained under non-stressed condition. Flowering period of 13 genotypes of the two sowings lasted from the 15th August till the 3rd September. Measurements were conducted at flowering and at grain filling (ca 14 days after flowering). Therefore, there were four stage × treatment combinations, namely flowering-drought-stressed treatment (FS), flowering-well-watered treatment (FW), grain filling-drought-stressed treatment (GS), and grain filling-well-watered treatment (GW), for the measurements as described below.

Gas exchange and chlorophyll fluorescence measurements

The flag leaves on the main stems of four representative plants (out of eight) per treatment of each genotype were used for measurements (except for IL42 at FS, because of a labour peak as flowering of the late IL42 coincided with grain filling of some of the earlier genotypes). We used an open gas exchange system (Li-Cor 6400; Li-Cor Inc., Lincoln, NE, USA) and an integrated fluorescence chamber head (Li-Cor 6400-40; Li-Cor Inc., Lincoln, NE, USA) to simultaneously measure gas exchange and chlorophyll fluorescence parameters at 21% O_2. All measurements were made at a leaf temperature of 25°C and a leaf-to-air vapour pressure difference (VPD) of 1.0-1.6 kPa. For C_i response curves, C_a was increased stepwise: 50, 60, 70, 80, 100, 150, 250, 380, 650, 1000, and 1500 μmol mol^{-1}, while keeping light intensity (I_{inc}) at 1000 μmol m^{-2} s^{-1}. For the I_{inc} response curves, photon flux densities were in an increasing series: 10, 30, 50, 70, 100, 170, 500, 1000, 1500, and 2000 μmol m^{-2} s^{-1}, while keeping C_a at 380 μmol mol^{-1}.

To properly estimate photosynthetic parameters, we also conducted measurements using a 2% O_2 gas mixture: a gas cylinder containing a mixture of 2% O_2 and 98% N_2 was used to blend with pure CO_2 to produce 2% O_2 in the leaf chamber. Under this condition, only the first half of the light or CO_2 response curves were measured: For C_i response curves, I_{inc} was kept at 1000 μmol m^{-2} s^{-1}, and C_a was increased stepwise: 50, 60, 70, 80, 100, and 150 μmol mol^{-1}; for I_{inc} response curves, I_{inc} was increased in the order of 10, 30, 50, 70, 100 and 170 μmol m^{-2} s^{-1}, and this half curve of the light response was obtained using 2% O_2 combined with 1000 μmol mol^{-1} C_a to ensure a non-photorespiratory condition. Light and CO_2 responses for the two O_2 levels were measured on the same leaves.

Leaf respiration in darkness (R_{dk}) was measured ~15 min after leaves had been placed in darkness. For measurements at each irradiance or CO_2 step, A was allowed to reach steady state, after which F_s (the steady-state fluorescence) was recorded. Then a saturating light pulse (>8500 μmol m^{-2} s^{-1} for 0.8 s) was applied to determine F'_m (the maximum fluorescence during the saturating light pulse). The apparent PSII e$^-$ transport efficiency was obtained as: $\Phi_2 = (F'_m - F_s)/F'_m$ (Genty *et al.*, 1989) for each irradiance or CO_2 step.

Leakage of CO_2 into and out of the leaf cuvette was corrected for all gas exchange data, using heat-killed leaves according to Flexas *et al.* (2007c).

Leaf N content measurements

Photosynthesis measurements at the two stages were made on the same leaf positions. After measurements at grain filling, the portion of the flag leaves used for the above described measurements was cut. The leaf material was weighed after drying at 70°C

to constant weight, and LMA (g dry matter m^{-2} leaf) was determined. Fraction of total-N in leaves was analysed using an element analyser based on the micro-Dumas combustion method. From these data, N_a (g N m^{-2} leaf) was calculated.

Model analysis

We use the model of Farquhar, von Caemmerer & Berry (1980) (FvCB model). The net CO_2 assimilation (A) is expressed as the minimum of the Rubisco limited rate (A_c) and the electron transport limited rate (A_j):

$$A = \min{(A_c, A_j)} \tag{1}$$

A_c is described, following the Michaelis-Menten kinetics:

$$A_c = \frac{(C_c - \Gamma_*)V_{cmax}}{C_c + K_{mc}(1 + O/K_{mo})} - R_d \tag{2}$$

where C_c and O are the CO_2 and O_2 levels at the carboxylation sites of Rubisco, V_{cmax} is the maximum rate of carboxylation, K_{mc} and K_{mo} are Michaelis-Menten constants of Rubisco for CO_2 and O_2, respectively, and Γ_* is the CO_2 compensation point in the absence of day respiration (R_d). In the model, $\Gamma_* = 0.5O/S_{c/o}$. As constants K_{mc} and K_{mo} are generally conservative for C_3 plants (von Caemmerer, 2000), their values were taken from Bernacchi *et al.* (2002).

A_j is described by:

$$A_j = \frac{(C_c - \Gamma_*)J}{4C_c + 8\Gamma_*} - R_d \tag{3}$$

where J is the potential PSII e$^-$ transport rate that is used for CO_2 fixation and photorespiration, and can be described by (Ögren & Evans, 1993; von Caemmerer, 2000; Yin *et al.*, 2009b):

$$J = \left(\kappa_{2LL}I_{inc} + J_{max} - \sqrt{(\kappa_{2LL}I_{inc} + J_{max})^2 - 4\theta J_{max}\kappa_{2LL}I_{inc}}\right)/(2\theta) \tag{4}$$

where κ_{2LL} is the conversion efficiency of incident light into J at strictly limiting light, J_{max} is the maximum value of J under saturated light, and θ is the convexity factor.

Model parameters were estimated according to the procedure described by Yin *et al.* (2009b). Specifically, using data of the e$^-$ transport-limited range under non-photorespiratory conditions (i.e. the irradiance response curve at 2% O_2 combined with 1000 μmol mol^{-1} C_a), a simple linear regression can be performed for the observed A against ($I_{inc}\Phi_2/4$). The slope of the regression yields the estimate of a lumped parameter s, and the intercept gives an estimate of R_d (Yin *et al.*, 2009b; 2011). This allowed the actual rate of linear electron transport to be calculated:

$$J = sI_{inc}\Phi_2 \tag{5}$$

Parameters J_{max}, κ_{2LL}, and θ were estimated by fitting Eqn (4) to the calculated J.

$S_{c/o}$ was calculated by following the procedure described by Yin *et al.* (2009b; see their Eqn (10)). Once all above parameters were estimated, their values were used as input to the model described below, upon which V_{cmax} and coefficients related to diffusional conductances were estimated.

Modelling of g_m and g_s

To examine any variation in mesophyll conductance (g_m) in response to C_i and irradiance (at 21% O_2), the variable J method (Harley *et al.*, 1992a) was first applied:

$$g_m = \frac{A}{C_i - \dfrac{\Gamma_*[J + 8(A + R_d)]}{J - 4(A + R_d)}} \tag{6}$$

where A and C_i were taken from gas exchange measurements and J was calculated by Eqn (5). This first analysis showed that g_m was variable (see Results). We therefore used a phenomenological equation of Yin *et al.* (2009b) to model g_m:

$$g_m = g_{mo} + \delta_m(A + R_d)/(C_c - \Gamma_*) \tag{7}$$

where g_{mo} is the minimum mesophyll conductance if the irradiance approaches zero, δ_m is the coefficient which defines the C_c/C_i relationship under saturating light as: $(C_c - \Gamma_*)/(C_i - \Gamma_*) = 1/(1+1/\delta_m)$ (Yin *et al.,* 2009b).

Combining Eqn (7) with Eqns (2) and (3) and replacing C_c with ($C_i - A/g_m$) yields (Yin *et al.*, 2009b):

$$A_c \text{ or } A_j = \left(-b - \sqrt{b^2 - 4ac}\right)/2a \tag{8}$$

where

$$a = x_2 + \Gamma_* + \delta_m(C_i + x_2)$$

$$b = -\{(x_2 + \Gamma_*)(x_1 - R_d) + (C_i + x_2)[g_{mo}(x_2 + \Gamma_*) + \delta_m(x_1 - R_d)] + \delta_m[x_1(C_i - \Gamma_*) - R_d(C_i + x_2)]\}$$

$$c = [g_{mo}(x_2 + \Gamma_*) + \delta_m(x_1 - R_d)][x_1(C_i - \Gamma_*) - R_d(C_i + x_2)]$$

with $x_1 = \begin{cases} V_{cmax} & \text{for } A_c \\ J/4 & \text{for } A_j \end{cases}$

$x_2 = \begin{cases} K_{mc}(1 + O/K_{mo}) & \text{for } A_c \\ 2\Gamma_* & \text{for } A_j \end{cases}$

The above model analysis showed that the pattern of variation of g_m is similar to that of g_s in response to CO_2 and irradiance levels (see Results). Therefore, Eqn (8) is

also valid to model the dynamics of an overall conductance (g_t) if places for C_i are replaced with C_a and those for δ_m are replaced with δ_t:

$$g_t = g_{to} + \delta_t(A + R_d)/(C_c - \Gamma_*) \tag{9}$$

where g_{to} is the minimum overall conductance, and δ_t is the coefficient defining the C_c/C_a relationship under saturating light as: $(C_c - \Gamma_*)/(C_a - \Gamma_*) = 1/(1+1/\delta_t)$. Note that $g_t = 1/(1/g_s+1/g_m)$, and that the value of C_a for our model analysis was adjusted to the CO_2 level at the leaf surface according to the boundary-layer conductance.

Assuming both g_{mo} and $g_{to} = 0$ (which is generally the case, see Results), and dividing Eqn (7) by Eqn (9), the following expression is obtained:

$$g_m/g_s = \delta_m/\delta_t - 1 \tag{10}$$

Eqn (10) quantitatively indicates an overall relative limitation of g_m vs g_s to photosynthesis. From Eqn (10), an equation for g_s is derived here:

$$g_s = \delta_s(A + R_d)/(C_c - \Gamma_*) \tag{11}$$

where $\delta_s = \delta_m\delta_t/(\delta_m - \delta_t)$.

Once A is calculated from Eqn (8), g_m can be calculated using the equation obtained by replacing C_c in Eqn (7) with $(C_i - A/g_m)$, and then by solving the equations for g_m (Yin *et al.,* 2009b):

$$g_m = [A + \delta_m(A + R_d)]/(C_i - \Gamma_*) \tag{12}$$

Similarly, g_t can be calculated using the equation obtained by replacing C_c in Eqn (9) with $(C_a - A/g_t)$ and then solving the equations for g_t

$$g_t = [A + \delta_t(A + R_d)]/(C_a - \Gamma_*) \tag{13}$$

To allow comparisons across genotype × treatment × stage combinations, we also estimated the value of g_m as constant, using the so-called NRH-A method ($g_{m(NRH-A)}$), based on the data obtained from high C_i of CO_2 response curves and low I_{inc} levels of light response curves at 21% O_2. The rationale for this method and the choice of data was fully discussed by Yin & Struik (2009b). This estimate using the NRH-A method should represent the average value of g_m within its lower range of the variation.

Ideotype design

The 13 lines used in this study were selected based on the QTLs detected by single-point analysis in Chapter 2. In order to quantify the additive effect of the QTLs on each parameter in our model, a statistical covariant model was used, in which the value of a parameter X of introgression line k, containing N QTLs (as represented by the nearest marker loci), for a specific stage (S_i) × treatment (T_j) combination was presented as:

$$X_{ijk} = \mu + S_i + T_j + \sum_{n=1}^{N} a_n \times M_{k,n} + e_{ijk} \tag{14}$$

where μ = the intercept; S_i = growth stage effect, which stands for either of the two stages (flowering or grain filling); T_j = treatment effect, which stands for either well watered or drought stressed; a_n = the additive effect of the n-th QTL; $M_{k,n}$ = genetic QTL scores of the individual introgression line k that take the value either -1 (allele coming from 'Shennong265') or 1 ('Haogelao' allele present), e_{ijk} is an error term.

For the ideotype design, only QTLs with significant enhancing additive effects ($P < 0.05$) were kept in Eqn (14). For example, an ideotype for improved photosynthesis was the virtual genotype of which parameter values were estimated as the sum of the allele effects that enhanced A for all QTLs of each FvCB-model component. To construct the A response of ideotype to irradiance, estimated parameters were used as inputs in Eqn (8). To calculate the TE response to irradiance, the following equation (Farquhar & Richards 1984) was used:

$$TE = \frac{C_a - C_i}{1.6\,(e_i - e_a)} \tag{15}$$

where $(e_i - e_a)$ is leaf-to-air VPD, and C_i is calculated from our model using values of A and estimates of the parameters δ_m and δ_s.

Statistics and curve fitting
A three-way analysis of variance of genotype $\times$ treatment $\times$ growth stage for the photosynthesis parameters was calculated. Non-linear fitting was carried out using the GAUSS method in PROC NLIN, multiple linear regression fitting for eqn (14) was performed using the PROC GLM, of SAS (SAS Institute Inc, Cary, NC, USA).

RESULTS

Estimates of photosynthesis parameters
The estimated values for $S_{c/o}$ did not differ among genotypes, nor among treatment $\times$ stage combinations; so a single value for $S_{c/o}$ was obtained from the pooled data (= 3.02 ± 0.03 mbar μbar^{-1}). As reported by Yin *et al.* (2009b), the estimated values for R_d did not differ between 21% and 2% O_2 levels, and a common R_d across the O_2 levels was obtained. However, the estimated values for R_d and s were genotype-, treatment- and stage-specific (See Supplementary material Table S1); and the values of R_d were generally lower than those of R_{dk} (Table S1). After parameter s was
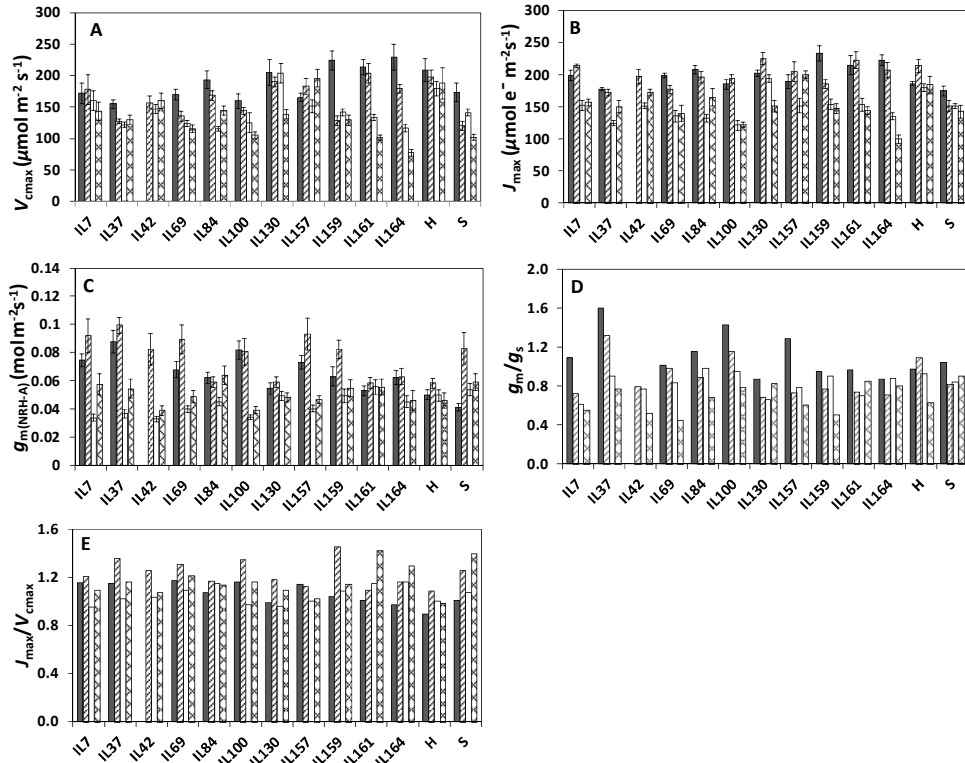
Fig. 2. Values of photosynthesis parameters estimated for flag leaves of introgression lines [including two parents: Haogelao (H) and Shennong265 (S)] at four stage × treatment combinations: flowering-drought-stressed treatment (FS,■); flowering-well-watered environment (FW,▨); grain filling-drought-stressed environment (GS,▭); grain filling-well watered environment (GW,▨). (A) maximum rate of Rubisco activity-limited carboxylation (V_{cmax}); (B) maximum value of electron transport rate used for $NADP^+$ reduction (J_{max}); (C) mesophyll conductance ($g_{m(NRH-A)}$), calculated by non-rectangular hyperbolic method (NRH-A) of Yin & Struik (2009b); (D) mesophyll conductance : stomatal conductance ratio (g_m/g_s), calculated by Eqn (10); (E) the ratio of J_{max} : V_{cmax}.

estimated, J was obtained (Eqn (5)) and J_{max}, κ_{2LL}, and θ were then estimated by fitting Eqn (4) (see Table S1).

After values of J, $S_{c/o}$ and R_d were known, we used Eqn (6) to evaluate the effects of variations of CO_2 concentration and light intensity on g_m for four stage × treatment combinations (FS, FW, GS, GW) of each genotype. In general, g_m strongly declined with an increase in C_i, and increased with an increase in light intensity, following the same response as g_s to CO_2 concentration and light intensity; so, a proportional relationship between g_m and g_s was obtained (Fig. S1). The slope of the proportional relationship, indicating the average g_m/g_s ratio, differed among the stage-treatment

combinations, and was higher for drought-stressed plants than for well-watered plants, and for flowering than for grain filling.

The variation of g_m across I_{inc} and across C_i levels was confirmed by the curve-fitting based on Eqn (8), as the value of parameter g_{mo} in the equation was found to be close to zero, whereas δ_m was found to vary from 0.452 to 1.571 (a zero g_{mo} combined with a non-zero δ_m would mean that g_m varies with C_i and I_{inc}; see Yin *et al.*, 2009b). This method allowed solving δ_m and V_{cmax} simultaneously (Table S1; Fig. 2), when using the earlier estimated $S_{c/o}$, R_d, J_{max}, θ and κ_{2LL} as inputs. In this method, a universal parameter δ_m (rather than specific g_m values) across whole photosynthesis light- and CO_2-response curves was estimated. A further analysis based on Eqn (9) also showed that parameter g_{to} did not differ significantly from zero ($P > 0.05$). Therefore, an overall g_m/g_s ratio (Eqn (10)) was calculated for each introgression line at each stage × treatment combination (Fig. 2D). The overall average g_m/g_s ratio obtained from this method for most of the stage × treatment combinations (Table S1) was slightly higher than those values shown in Fig. S1, probably because the variable J method assumes no alternative e⁻ transport, whereas the curve-fitting method does account for any alternative e⁻ transport (Yin *et al.*, 2009b).

Components of variation in and correlations among photosynthetic parameters

The variation in each estimated photosynthetic parameter can be statistically partitioned into genetic, environmental (stress vs. non-stress), and developmental (i.e. flowering vs. grain filling) components, and their two-way interactions. However, as most interactions were not significant ($P > 0.05$; results not shown), we omitted all interaction terms (Table 1). Significant genetic differences were found for R_d, κ_{2LL}, J_{max}, θ, and V_{cmax} as well as for δ_s and g_m/g_s ratio ($P < 0.05$, Table 1, Fig. 2), although environmental and developmental components were contributing most to the variation in most parameters (Table 1).

The parameters of photosynthesis model were partly correlated (Table S2). In particular, the correlations between J_{max} and V_{cmax}, δ_m and V_{cmax}, δ_m and δ_s were significant in each stage × treatment combination ($P < 0.05$). These correlations may suggest that these traits are, at least partly, under common genetic control.

Physiological basis of the genetic variation

Significant genetic differences of some model parameters ($P < 0.05$, Table 1, Fig. 2) hint a physiological basis for genetic variation in A found earlier by us in Chapter 2, in which we identified QTLs for light-saturated A (A_{max}) under field conditions.

Using our model approach, A_{max} can be dissected into four physiological components: g_s, g_m, electron transport, and Rubisco activity. To quantitatively analyse

Table 1. A three-way ANOVA of genetic effect vs. treatment vs. growth stage for the estimated photosynthesis parameters. (F and P values significant at a level of $P < 0.05$ are in bold. For definitions see Appendix.)

Parameters		*F*-value (probability of significance)		
		Genetic effect	Treatment	Stage (ontogeny)
Primary parameters of the model	R_d	**4.82 (<0.0001)**	3.06 (0.0721)	**15.45 (0.0004)**
	κ_{2LL}	**6.14 (<0.0001)**	**9.34 (0.0042)**	**38.93 (<0.0001)**
	J_{max}	**2.00 (0.0494)**	0.00 (0.9850)	**80.19 (<0.0001)**
	θ	**6.09 (<0.0001)**	**9.78 (0.0035)**	**100.85 (<0.0001)**
	δ_m	1.69 (0.1110)	**5.55 (0.0241)**	1.99 (0.1671)
	V_{cmax}	**2.44 (0.0191)**	**6.96 (0.0122)**	**29.21 (<0.0001)**
	δ_t	1.70 (0.1090)	**20.65 (<0.0001)**	0.10 (0.7587)
Other parameters	δ_s	**2.39 (0.0218)**	**53.46 (<0.0001)**	**9.46 (0.0040)**
	$g_{m(NRH-A)}$	0.82 (0.6314)	**9.39 (0.0041)**	**50.24 (<0.0001)**
	g_m/g_s	**2.79 (0.0085)**	**19.27 (<0.0001)**	**31.38 (<0.0001)**
	J_{max}/V_{cmax}	1.49 (0.1721)	**25.85 (<0.0001)**	1.88 (0.1787)

the effects of each component, A_{max} (at 380 μmol mol^{-1} CO_2, 1500 μmol m^{-2} s^{-1} irradiance, 25°C, and 1.5 kPa VPD) was first plotted against each component here. Within each stage × treatment combination, the correlation between A_{max} and each component (g_s, g_m, J_{max} and V_{cmax}) can be observed (Fig. S2), providing the evidence about where genetic differences in A_{max} possibly came about. In order to quantify the main sources of genetic variation in A_{max}, a multiple regression analysis was carried out (Table 2). For each stage × treatment combination, the genetic variation in g_s and g_m had the largest impact on the genetic variation in A_{max}. Under well-watered treatment, g_m caused more genetic variation in A_{max} than g_s did, while under drought-stressed treatment, g_s accounted for more genetic variation.

We also analysed TE under the same measurement conditions (Fig. S3). When we inspected the relationship within each stage × treatment combination, the correlation appeared very weak, except for g_s (Fig. S3A) and g_m/g_s (Fig. 3). Multiple regression analysis (Table 2) also showed that genetic variation in g_s and g_m, relative to that in V_{cmax} and J_{max}, contributed more to TE in this genetic background, and not surprisingly, g_m and g_s affected TE in an opposite direction.

Physiological basis of a major photosynthesis QTL

Of the ILs used, IL161 is unique in that it has the background of the recurrent parent Shennong265 except for a single introgression segment on chromosome 9 from the donor parent (Fig. 1). Compared with the recurrent parent, IL161 significantly increased A_{max} across stages and treatments; thus, a major QTL was consistently

Table 2. Multiple linear regression analysis of light-saturated photosynthesis (A_{max}) or transpiration efficiency (TE) as a function of g_s, g_m, J_{max} and V_{cmax} (i.e. A_{max} or TE $= b_0 + b_1g_s + b_2g_m + b_3J_{max} + b_4V_{cmax}$), based on data of 11 introgression lines and their parents, for each stage × treatment combination

Trait	Stage × treatment	Intercept (b_0)	Regression coefficient (probability of significance)			
			b_1	b_2	b_3	b_4
A_{max}	FS	**1.21**	**46.99 (4.4 × 10⁻⁵)**[1]	**27.24 (3.0 × 10⁻⁴)**[2]	**0.04 (0.0058)**[3]	0.00 (0.6866)[4]
	FW	**1.26**	**31.19 (1.3 × 10⁻⁵)**[2]	**32.82 (4.9 × 10⁻⁶)**[1]	0.02 (0.0857)[4]	**0.02 (0.0260)**[3]
	GS	**0.63**	**30.85 (7.0 × 10⁻⁸)**[1]	**46.59 (1.5 × 10⁻⁷)**[2]	**0.04 (0.0001)**[3]	0.00 (0.9929)[4]
	GW	**1.39**	**22.45 (7.7 × 10⁻⁵)**[2]	**53.15 (2.8 × 10⁻⁵)**[1]	0.00 (0.8682)[4]	**0.04 (0.0147)**[3]
TE	FS	**6.02**	**-20.70 (6.7 × 10⁻⁷)**[1]	**6.33 (0.0003)**[2]	**0.01 (0.0024)**[3]	0.00 (0.1739)[4]
	FW	**4.29**	**-15.77 (1.1 × 10⁻⁶)**[1]	**8.11 (8.0 × 10⁻⁵)**[2]	0.00 (0.3442)[4]	0.01 (0.0523)[3]
	GS	**4.84**	**-21.31 (1.0 × 10⁻⁷)**[1]	**12.89 (0.0002)**[2]	0.01 (0.0595)[3]	0.00 (0.7278)[4]
	GW	**4.47**	**-13.21 (3.2 × 10⁻⁷)**[1]	**10.00 (0.0005)**[2]	0.00 (0.4847)[4]	**0.01 (0.0394)**[3]

Coefficient values significant at a level of $P < 0.05$ are in bold. [1, 2, 3, 4] the comparative importance of each parameter, determined from the level of significance. For definitions see Appendix.
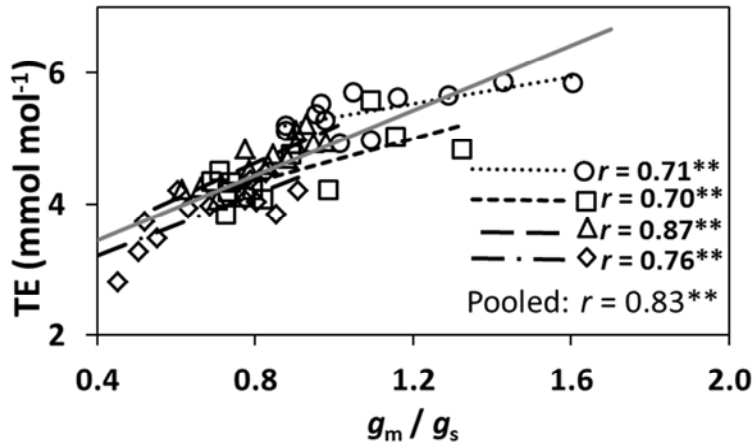
Fig. 3. Relationship between transpiration efficiency (TE) (380 μmol mol⁻¹ CO_2, 1500 μmol m⁻² s⁻¹ light intensity, 25°C, and 1.5 kPa VPD) and ratio of mesophyll conductance and stomatal conductance (g_m/g_s). Linear regressions were fitted for overall data (grey solid lines) and each stage × treatment combination: flowering-drought-stressed treatment (FS, ○, ·········), flowering-well-watered environment (FW, □, -----), grain filling-drought-stressed environment (GS, △, — —), and grain filling-well watered environment (GW, ◇, — · —). The significance of each correlation was shown as: **, $P < 0.01$.
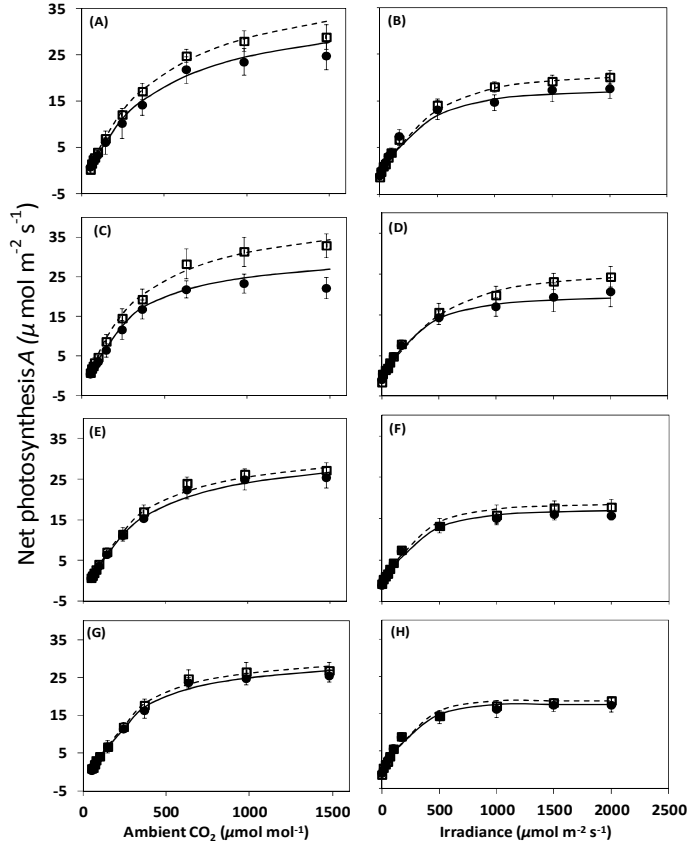
Fig. 4. Photosynthesis response curves of IL161 (empty square) and Shennong265 (filled circle) under 21% O_2 at four stage × treatment combinations: flowering-drought-stressed treatment (A, B); flowering-well-watered environment (C, D); grain filling-drought-stressed environment (E, F); grain filling-well-watered environment (G, H). The curves are drawn from the model using fitted parameter values: for IL161, dashed lines; for Shennong265, full line. Left panels (A, C, E, G) show the response of net photosynthesis A to ambient CO_2 (C_a) under light intensity of 1000 μmol m^{-2} s^{-1}. Right panels (B, D, F, H) show the response of photosynthesis A to light intensity under 380 μmol mol^{-1} CO_2. Values are means ± SD (n=4).

detected for A_{max} on chromosome 9 (Chapter 2). CO_2 and light response curves measured in the present study indicated that the QTL contributed to a higher photosynthesis rate across all irradiance and CO_2 levels (Fig. 4). Through our analysis, seven parameters of both IL161 and Shengnong265 were estimated for each stage × treatment combination (Table 3). There was no significant difference between them for R_d, κ_{2LL}, and θ ($P > 0.05$). At flowering, IL161 showed significantly higher g_m, g_s, V_{cmax} and J_{max} than Shennong265 across the two treatments. At grain filling, however,

Table 3. Parameter values (± standard error of the estimate) of the photosynthesis model, estimated for IL161 and Shennong265 (S) that differ in a single introgression region on chromosome 9 (see Gu *et al.* 2012), for four stage-treatment combinations

	Genotype	Model parameters							Other derived parameters	
		R_d	κ_{2LL}	J_{max}	θ	V_{cmax}	δ_m	δ_i	g_s^a	g_m^b
FS	IL161	0.669 ± 0.193	0.344 ± 0.044	214.8 ± 14.9*	0.790 ± 0.151	213.3 ± 12.0*	0.524 ± 0.019*	0.267 ± 0.008	0.145	0.140
	S	0.539 ± 0.177	0.321 ± 0.029	174.6 ± 7.6	0.806 ± 0.098	173.3 ± 14.5	0.572 ± 0.031	0.279 ± 0.013	0.120	0.125
FW	IL161	0.474 ± 0.175	0.332 ± 0.033	222.7 ± 13.5*	0.744 ± 0.136	204.0 ± 15.4*	0.697 ± 0.036*	0.401 ± 0.016*	0.221	0.163
	S	0.407 ± 0.139	0.326 ± 0.038	151.1 ± 8.1	0.764 ± 0.143	120.7 ± 7.2	1.044 ± 0.103	0.574 ± 0.032	0.190	0.156
GS	IL161	0.317 ± 0.147	0.322 ± 0.053	153.2 ± 10.4	0.863 ± 0.137	133.4 ± 4.6*	0.767 ± 0.039*	0.450 ± 0.016*	0.186	0.130
	S	0.365 ± 0.124	0.316 ± 0.020	151.4 ± 3.9	0.852 ± 0.054	141.1 ± 5.5	0.694 ± 0.035	0.376 ± 0.012	0.146	0.123
GW	IL161	0.059 ± 0.235	0.297 ± 0.031	143.9 ± 6.0	0.884 ± 0.078	101.5 ± 4.7	1.340 ± 0.158	0.723 ± 0.026*	0.200	0.171
	S	-0.147 ± 0.172	0.312 ± 0.050	142.2 ± 9.7	0.823 ± 0.160	102.0 ± 4.9	1.211 ± 0.139	0.635 ± 0.021	0.173	0.158

[a] g_s and [b] g_m were derived from the fitted model at saturated light of 1500 μmol m^{-2} s^{-1}, CO_2 concentration of 380 μmol mol^{-1}.

*, statistically significantly different between Introgression line 161 (IL161) and recurrent parent Shennong265 (S) ($P < 0.05$).

For parameter definitions see Appendix. FS, flowering stage – drought stressed environment; FW, flowering stage – well watered environment; GS, grain filling stage – drought stressed environment; GW, grain filling stage – well watered environment.
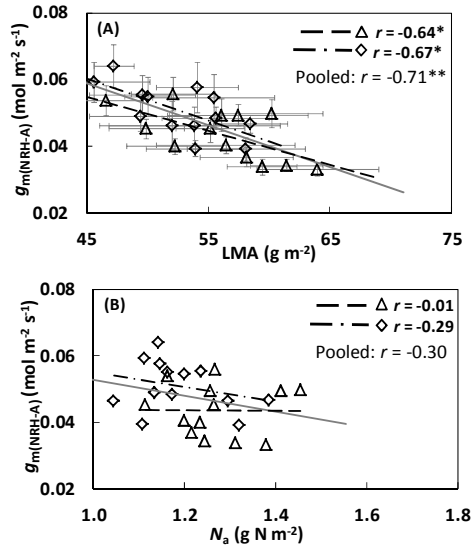
Fig. 5. Relationship between (A) mesophyll conductance ($g_{m(NRH-A)}$) calculated by the non-rectangular hyperbolic method (Yin & Struik, 2009b) and leaf mass per area (LMA); (B) $g_{m(NRH-A)}$ and leaf nitrogen per unit area (N_a). Values are means ± standard deviations of four replicates. Linear regressions were fitted for overall data (grey solid lines) and each stage × treatment combination: grain filling-drought-stressed environment (GS, Δ, − − −), and grain filling-well-watered environment (GW, ◇, — · —). The significance of each correlation was shown as: *, $P < 0.05$; **, $P < 0.01$.

only higher diffusional conductance (larger g_m and g_s) could be the reason for higher A, as V_{cmax} was even lower in IL161 than in Shennong265 for the stress treatment (Table 3). Therefore, there was greater difference between IL161 and Shenong265 at flowering than at grain filling (Fig. 4). Our whole-curve measurements are consistent with the results in Chapter 2 that larger additive effects of the QTL on A_{max} were obtained at flowering than at grain filling.

Relationships between photosynthesis parameters and leaf morpho-physiological characteristics

The variation in $g_{m(NRH-A)}$, either across genotypes or across treatments, was negatively correlated with LMA (Fig. 5A). Similar relationships were found between LMA and g_m or g_s calculated for the condition of measuring A_{max}, despite lower r^2-values (results not shown). As expected, drought stress induced thicker leaves (increased LMA, Fig. 5A) and the increased LMA led to an increased N_a ($r^2 = 0.40$). But there was a poor correlation between $g_{m(NRH-A)}$ and N_a ($r^2 = 0.08$; Fig. 5B). Instead, the variation in J_{max} and V_{cmax}, either across genotypes or across water-supply treatments, was found to be positively correlated with N_a (Fig. 6A,B), but less correlated with LMA (results not
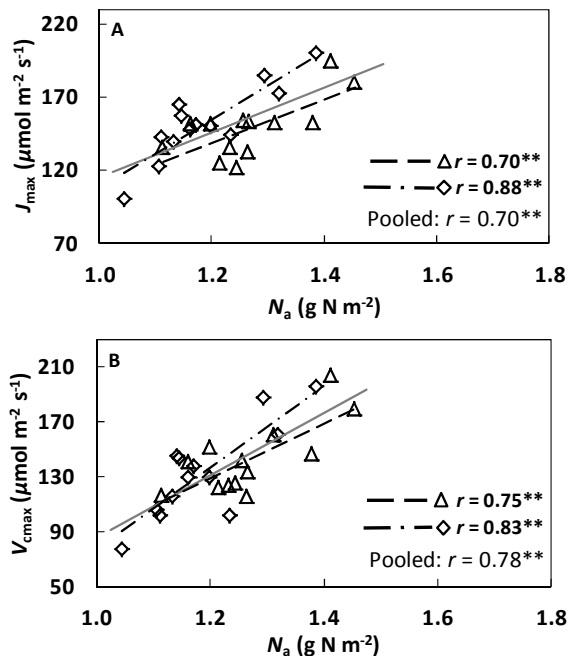
Fig. 6. Relationship between (*A*) electron transport capacity (J_{max}) and leaf nitrogen content (N_a), (B) Rubisco carboxylation capacity (V_{cmax}) and N_a. The symbols and significance levels were as in Fig. 5.

shown). Analysis with an *F*-test demonstrated that generally there was no significant difference between well-watered and drought-stressed plants at grain filling on the relationships of Figs 5-6 ($P > 0.05$), although the slope of the relationship between J_{max} and N_a was significantly lower ($P = 0.012$) for plants under drought.

Ideotype design based on physiological understanding

Given the significant genetic difference in each of the model component traits (Table 1; Fig. 2) and their significant effects on A_{max} and TE (Table 2; Fig. S2 & S3), it is worthwhile to explore the potential to improve *A* and TE using the genetic variation observed. We, therefore, estimated additive effects of individual genome loci, based on Eqn (14). Of the loci differing among the ILs, seven loci were identified to significantly affect the seven primary model parameters (Fig. 1; Table 4). These seven loci were also identified or in close proximity with those mapped for A_{max} using the whole IL population (Chapter 2), suggesting that our selected 11 ILs did represent the population well. There was no one-to-one locus-parameter relationship. Instead, each model parameter was controlled by one to three loci and most loci had an effect

Table 4. The effects of growth stages (flowering, S_f; grain filling, S_g), treatments (T_s, drought stressed; T_w, well watered), and additive effects of QTLs (i.e. a_1, a_2, a_3, a_4, a_5, a_6 and a_7) on seven modelled traits, estimated from regression analysis using Eqn (14): $X_{ijk} = \mu + S_i + T_j + \sum_{n=1}^{N} a_n \times M_{k,n} + e_{ijk}$. QTLs positions and their additive effect coefficients are marked in Fig. 1. For definitions of the traits, see Table in the Appendix. Empty cell in this Table means that the corresponding effect was not significant ($P > 0.05$) and was, therefore, not included in the regression model.

Trait	intercept μ	Growth stage S_f	S_g	Treatment T_s	T_w	The additive effect (a_n) of QTLs [a] a_1	a_2	a_3	a_4	a_5	a_6	a_7
R_d	0.6071	0.2454	0						0.1240	0.2055		
κ_{2LL}	0.3097	0.0219	0	0.0104	0		0.0088		-0.0213			0.0182
J_{max}	151.81	47.81	0				8.42				6.85	
θ	0.8637	-0.0890	0	0.0284	0	0.0277	-0.0304		0.0386			
δ_m	0.8189			-0.1390	0	-0.0890						
V_{cmax}	131.14	37.78	0	18.42	0		12.74				8.84	
δ_s	1.1547	-0.1488	0	-0.3630	0	-0.0899		0.0971		-0.0878		

[a] The positive value for the allelic effect from donor parent 'Haogelao'

on multiple parameters (Table 4), providing a genetic basis of significant correlations between model parameters in Table S2.

The ideotype for high A requires high g_s and g_m and improved photosynthetic efficiency (κ_{2LL} and θ) and capacities (V_{cmax} and J_{max}), while the ideotype for high TE requires low g_s, high g_m and improved photosynthetic efficiency and capacities. So, the ideotype of high A carried the alleles having positive effects on κ_{2LL}, J_{max}, θ, δ_m, δ_s and V_{cmax}, and negative effects on R_d, whereas the ideotype of high TE carried the alleles having positive effects on κ_{2LL}, J_{max}, θ, δ_m and V_{cmax}, and negative effects on δ_s and R_d. The ideotype of high A showed an increase in A of 15.2% (FS), 15.5% (FW), 20.6% (GS) and 17.1% (GW) compared with mean A of 13 ILs (solid curves Fig. 7); the ideotype of high TE showed an increase of 32.2% (FS), 14.8% (FW), 26.1% (GS) and 17.3% (GW) compared with mean TE of 13 ILs (solid curves Fig. 8).

The above estimated improvement in A or TE was moderate, because the same alleles at some loci have contradicted effects on different photosynthesis parameters (Table 4). Assuming that these contradicted effects were not due to pleiotropy, but due to tight gene linkage which could be broken through further rounds of introgression and a higher density marker map to develop near isogenic lines carrying fine mapped QTLs, we evaluated virtual ideotypes for A and TE that only contained positive effects in all the photosynthesis parameters. For A, this virtual ideotype showed an average improvement of 29.9% (FS), 29.3% (FW), 36.4% (GS) and 34.5% (GW) compared with the mean A of 13 ILs (dotted curves Fig. 7). For TE, the virtual ideotype showed an average improvement of 46.9% (FS), 28.2% (FW), 42.0% (GS) and 31.6% (GW) when compared with the mean TE of 13 ILs (dotted curves Fig. 8). When compared with the best genotype we investigated in each stage × treatment combination, for A, the virtual ideotype showed an improvement of 11.0% (FS), 9.6% (FW), 18.7% (GS)
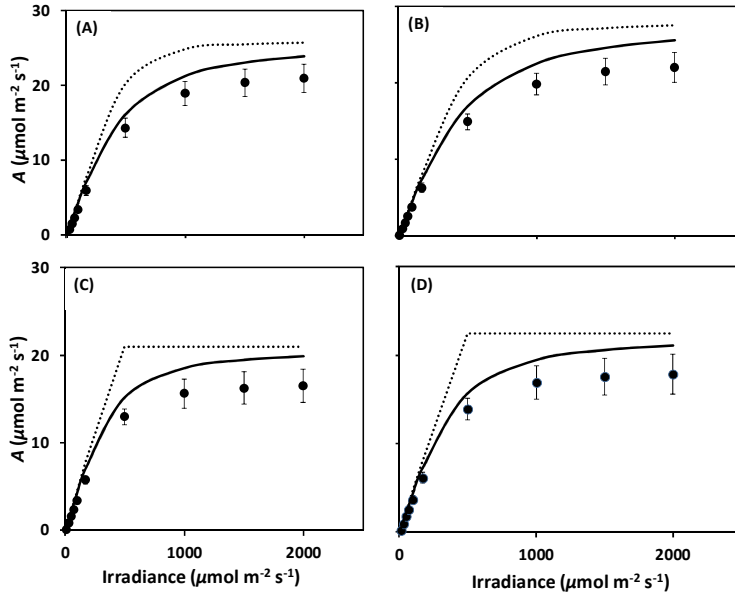
Fig. 7. Constructed response curve of net photosynthetic rate (*A*) to light intensity at ambient CO_2 concentration (380 μmol mol^{-1}) at four stage $\times$ treatment combinations: (A) flowering-drought-stressed treatment (FS); (B) flowering-well-watered environment (FW); (C) grain filling-drought-stressed environment (GS); (D) grain filling-wellwatered environment (GW). Rate of photosynthesis of 13 lines (filled circle, values are means $\pm$ SDs of 13 lines) were calculated from the model using fitted parameter values. The ideotype response (solid curves) and the potential virtual ideotype curves (dotted curves) of photosynthesis were drawn using parameter values, which were calculated by methods described in the Materials & Methods and Results sections.

and 28.5% (GW); for TE, the virtual ideotype showed an improvement of 38.3% (FS), 12.3% (FW), 33.9% (GS) and 15.8% (GW).

The above analysis examined the ideotypes for *A* and TE separately. To explore the potential of selecting a genotype with both improved TE and photosynthesis, A_{max} was plotted against TE (Fig. 9). There were negative correlations for all the stage $\times$ treatment combinations, and the negative correlations were more significant under drought environment than well-watered environment. These relationships suggest that simultaneous improvement of *A* and TE is difficult, especially under drought. We shall later discuss the opportunities of simultaneous selection for improved *A* and TE.

DISCUSSION

Physiological basis of genetic variation in photosynthesis

Our model approach allowed to quantitatively dissect photosynthesis into different physiological components: g_s, g_m, and biochemical efficiency (κ_{2LL}, θ) and
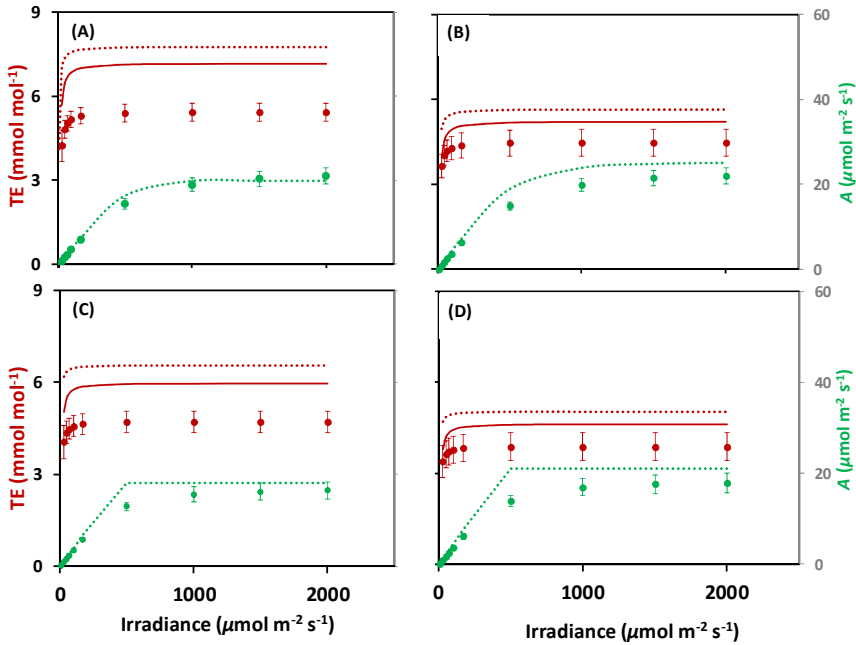
Fig. 8. Constructed transpiration efficiency (TE; dark red colour) and photosynthesis (*A*; green colour) response curve to light intensity at four stage × treatment combinations: (A) flowering-drought-stressed treatment (FS); (B) flowering-well-watered environment (FW); (C) grain filling-drought-stressed environment (GS); (D) grain filling-well-watered environment (GW). All the data were estimated at 380 μmol mol^{-1} CO_2 and 1.5 kPa VPD. The TE (filled dark red circles, values are means ± SDs) and rate of photosynthesis (filled green circles, values are means ± SDs) and of 13 lines were calculated from the model using fitted parameter values. The ideotype response (solid dark red curves), the potential virtual ideotype curves (dotted dark red curves) of TE, and corresponding *A* of the virtual ideotype (dotted green curves) were drawn. These curves were drawn using parameter values calculated by methods described in the Materials & Methods and Results sections.

biochemical capacity (J_{max} and V_{cmax}). In our analysis, most of model parameters showed significant genetic differences (Table 1). For example, parameters κ_{2LL} and θ both affect the electron transport efficiency under limited light. So the genetic variation in κ_{2LL} and θ (Table 1) could be potentially used to improve photosynthetic efficiency before light intensity reaches saturation.

As our previous analysis identified QTLs for A_{max} (Chapter 2), we specifically analysed the relative contribution of photosynthesis parameters (g_s, g_m, V_{cmax} and J_{max}, Table 2) relevant for the condition under which A_{max} was measured. g_s was found to be most associated with genetic variation in A_{max} in our IL population (Table 2, Fig. S2A) under drought. This was in line with reported results showing that mapped QTLs of net photosynthesis (Adachi *et al.*, 2011) were related to g_s. These results are not
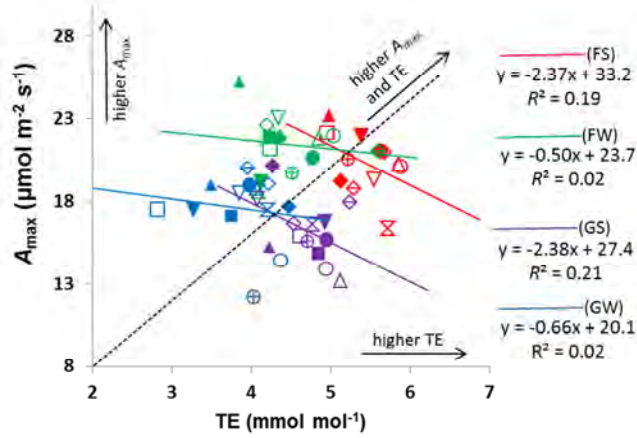
Fig. 9. Relationships between transpiration efficiency (TE) and light-saturated net photosynthesis rate (A_{max}) (380 μmol mol^{-1} CO_2, 1500 μmol m^{-2} s^{-1} light intensity, 25°C, and 1.5 kPa vapour pressure deficit). Symbols and accessions as follows: ▲ IL7, △ IL37, ■ IL42, □ IL69, ● IL84, ○ IL100, ◆ IL130, ◇ IL157, ▼ IL159, ▽ IL161, ⊕ IL164, ⬙ Haogelao, and ⊠ Shennong265. To distinguish between stage × treatment combinations, different colours waere used: flowering-drought-stressed treatment (FS, Red); flowering-well-watered environment (FW, Green); grain filling-drought-stressed environment (GS, Purple); and grain filling-well watered environment (GW, Blue). The linear regression lines (solid lines) are fitted for each stage × treatment combination. The diagonal dashed line is fitted for all the stage × treatment combinations, when forcing the regression line to go through the origin. This dashed line shows the trendline for both high TE and A_{max}.

surprising, given that g_s controls diffusion of CO_2 from ambient air into intercellular airspace and that stomata have evolved into physiological control mechanisms to maximize carbon gain while minimizing water loss (Lawson *et al.*, 2011). However, g_m was also important for the expression of genetic variation in A_{max} (Table 2, Fig. S2B). In fact, under well watered conditions, g_m contributed most to the genetic variation in A_{max} (Table 2).

We found that V_{cmax} and J_{max} contributed comparatively less to genetic variation in A_{max} in each stage × treatment combination (Table 2). This is surprising, given that V_{cmax} and J_{max} reflect Rubisco carboxylation and e$^-$ transport capacities, respectively. The weak correlation between biochemical capacities and A_{max} within each stage × treatment (Fig. S2 D,E) could be due to the small range of variation in V_{cmax} and J_{max}. A comparison of IL161 vs Shengnong265 (whose difference was due to a single introgression on Chromosome 9) showed (Table 3) that V_{cmax} and J_{max} together with g_m and g_s did explain the difference in photosynthesis light and CO_2-response curves (Fig. 4), at least for the flowering stage.

It is known that a long-term environmental adaptation results in a change in leaf morphology, and LMA as a morphological trait has a high plasticity in adjusting to

environmental conditions (Westoby *et al.*, 2002; Poorter *et al.*, 2009). For example, Pons & Pearcy (1994) showed that plants that switched from a high-light environment to low light can substantially (30%-50%) decrease LMA within days. The change of LMA was also shown in our data obtained after the grain-filling stage measurements, where the average LMA for drought stressed leaves was higher than the average for non-stress leaves (Fig. 5A). Our result agreed with the literature (Flexas *et al.*, 2008; Niinemets *et al.*, 2009; Galmés *et al.*, 2011) that g_m decreased with increasing LMA (Fig. 5A). Interestingly, this relation also holds for the genetic variation across 13 lines within either stress or non-stress treatment and the stress treatment did not change the relationship (Fig. 5A). This suggests that LMA plays an important role in the plant's adaptation to environmental conditions as well as in the plant's genotypic strategies within the same environment.

Similar to LMA, N_a also varied between treatments and among genotypes (Fig. 5B). V_{cmax} and J_{max}, rather than g_m or g_s, were linearly correlated with N_a (Figs. 5 & 6). Furthermore, V_{cmax} and J_{max} were less correlated with LMA (results not shown). Again, water supply treatments hardly affected these relationships across the 13 genotypes. Since V_{cmax} and J_{max} affected genetic variation of A_{max} (Table 2), especially at flowering stage (Table 3), elevated capacity of nitrogen accumulation in the leaf should be a preferred trait for improving leaf photosynthetic capacity, as suggested in the literature (Peng *et al.*, 1995; Shiratsuchi *et al.*, 2006; Taylaran *et al.*, 2011).

Physiological basis of genetic variation in transpiration efficiency
TE is another important breeding target for drought tolerance (Condon *et al.*, 2002, 2004). Our data showed that genetic variation in g_s was best correlated with genetic variation in TE in our genetic material (Table 2, Fig. S3A). This was in line with reported results that the gene of TE, *ERECTA* was related to g_s (Masle *et al.*, 2005).

From a theoretical perspective, however, Condon *et al.* (2004) indicated that under certain environment conditions, TE could be improved not only by lowering g_s, but also by higher photosynthetic potential, or a combination of these two. Especially, a greater g_m/g_s ratio results in a higher TE without a negative impact on carboxylation (Barbour *et al.*, 2010; Galmés *et al.*, 2011). We found significant genetic difference for the g_m/g_s ratio in this population ($P < 0.01$, Table 1), and the genetic variation in TE was strongly correlated with the variation in this ratio (Fig. 3). A further improvement of TE may be achieved by improving biochemical activities, resulting in improved *A* with the same transpiration. An ideal plant in drylands would have low g_s, high g_m, and improved biochemical efficiency (Flexas *et al.*, 2010). However, our data showed little association between TE and V_{cmax} or J_{max} (Table 2, Fig. S3).

Potential of using genetic variation to improve photosynthesis and TE

Our model analysis revealed a strong physiological basis of the genetic variation in photosynthesis and in TE; therefore, the model was used to design ideotypes for an improved *A* or TE based on their physiological components. This kind of bottom-up approach was successful in the past for yield component analysis. For example, more insights could be obtained from analysing QTLs or genes for yield components rather than for grain yield per se (Yin *et al.*, 2002), and the component-trait QTLs could be explored to improve yields. Based on this ideotype idea, recent genomic studies have successfully identified genes for one or a few of yield components (reviewed by Xing & Zhang, 2011; Miura *et al*., 2011). However, very few studies were performed using the same approach for photosynthesis.

Based on the genetic variation from our study, we can significantly improve *A* and TE by manipulating alleles of loci influencing different physiological components of photosynthesis (Figs. 7 & 8), suggesting that understanding of the physiological basis of photosynthesis will benefit marker-assisted selection (MAS). Some gene linkage limited the further improvement. For example, a locus from Shenong265 has positive effects on both g_m and g_s, which will benefit breeding for high *A*, while it has a contradictory effect for high TE. High g_s will increase photosynthesis at the expense of high transpiration. Any further improvement of these rice ideotypes of our IL background for higher photosynthetic performance and TE requires further steps of MAS. For example, further backcrossing using markers is needed to reduce the size of introgression segments and develop near isogenic lines carrying fine-mapped QTLs to break any gene linkage. Through this approach, a potential improved ideotype could be achieved as shown by the dotted lines in Fig. 7 and 8.

Can *A* and TE be improved simultaneously?

As expected from existing physiological understanding, our data showed general negative correlations between A_{max} and TE among ILs in each stage $\times$ treatment combination (Fig. 9). This agrees with the observation that selection for higher TE often hampered plant growth and resulted in smaller plants (Blum, 2005). The negative correlations were stronger under drought than in the well-watered environment (Fig. 9), consistent with the result shown in Table 2 that A_{max} was most limited by genetic variation in g_s under drought. So, under drought, any genetic variation resulting in decreased g_s will improve TE but will decrease photosynthesis. Under well-watered conditions, however, A_{max} was most limited by genetic variation in g_m (Table 2); so it is comparatively easier to select for higher TE and higher A_{max}. The small R^2 in Fig. 9 on one hand may reflect the small range of the data set, on the other hand may imply the potential extent to select for both higher A_{max} and TE (i.e. selecting genotypes that

follow the dashed line in the figure). In ideotype design analysis, we assessed the trade-off between *A* and TE, and found that improving *A* would be achieved at the expense of on average 35.3% decrease of TE, when comparing the best virtual ideotype of *A* with the best virtual ideotype of TE. Similarly average *A* would decrease by 12.4%, when comparing the best virtual ideotype of TE with the best virtual ideotype of *A* (data not shown). But if the linkage between g_m and g_s (as shown by the correlation between δ_m and δ_s in Table S2, co-location of QTLs of δ_m and δ_s in Table 4 and Fig. 1) could be broken (reflected by dotted curves in Fig. 8), the best virtual ideotype could have both improved TE (dark red colour) and *A* (green colour) compared with the average of ILs (dotted lines vs filled circles in Fig. 8). Similar results were given by Barbour *et al.*, (2010) for barley varieties, in which variety 'Dash' having higher g_m and comparatively lower g_s resulted in highest *A* and TE across the six varieties they examined. Our analysis using ILs highlights the possibility to improve both *A* and TE within the same genetic background.

CONCLUDING REMARKS

In this study, combined gas exchange and chlorophyll fluorescence data of CO_2 and light response curves of photosynthesis were measured for two stages on leaves of 13 ILs under moderate-drought and well-watered conditions. These curves showed that our previously reported QTLs, especially the major QTL on Chromosome 9 (Fig. 4), identified for the condition of A_{max} measurements (Chapter 2), also affected *A* across all irradiance and CO_2 levels. Using these curves, we estimated seven parameters of a combined conductance-FvCB model as proposed by Yin *et al.* (2009b). We then quantitatively dissected photosynthesis into different physiological components: stomatal conductance, mesophyll conductance, and biochemical efficiency and capacity. Our model method, Eqn (10), presents a novel approach to quantitatively analyse an overall relative limitation of stomatal vs mesophyll diffusion on photosynthesis of a genotype under a given condition.

Our data and analysis confirm the literature reports in several areas. Firstly, we confirmed that g_m strongly declined with an increase in C_i, and increased with an increase in light intensity, a response to CO_2 concentration and light intensity similar to that of g_s (Centritto *et al.*, 2003; Flexas *et al.*, 2007a; Yin *et al.*, 2009b; Douthe *et al.*, 2011). Therefore, there was strict g_m/g_s proportionality (Fig. S1), although independence of g_m on I_{inc} and C_i levels was also found (Tazoe *et al.*, 2009, 2011). Secondly, our results confirm that there was little significant influence of drought on V_{cmax} and J_{max} ($P > 0.01$), suggesting no metabolic impairment but increased diffusional resistances happened under moderate drought (Centritto *et al.*, 2003; Grassi & Magnani, 2005; Galmés *et al.*, 2007). Our result of a decrease in J_{max}/V_{cmax} under

drought is in line with that of Galle *et al.* (2011), suggesting that drought stress could cause down-regulation of linear electron transport (Kohzuma *et al.*, 2009). Thirdly, we confirmed the decrease of photosynthetic parameters with leaf ageing (e.g. Harley *et al.* 1992b; Ethier *et al.* 2006; Flexas *et al.* 2007b). The ageing decreased g_m, R_d, κ_{2LL}, J_{max}, V_{cmax}, g_m/g_s, and increased θ. These changes of parameters may be associated with leaf nitrogen loss through protein degradation as a result of re-translocation of nitrogen to the grains.

However, the main aims of our study were to analyse the effect of genotypes arisen from segregation of photosynthetic QTLs detected in Chapter 2 and to identify the physiological basis of genetic variation and QTLs. Although the effects of leaf stage and water supply on photosynthesis were predominant, the effect of genotype was significant enough to allow the examination of the physiological basis of the genetic variation by the use of the combined conductance-FvCB model. Genetic variation in A_{max} as well as in TE was mainly caused by genetic variation in g_s and g_m (Table 2), in line with significant stomatal and mesophyll limitations when plants face environmental stress (e.g. drought stress, Flexas *et al.*, 2004; Grassi & Magnani, 2005). So, more efforts should be focused on g_s and g_m in breeding programmes for improving photosynthesis and TE. Furthermore, the relationships between photosynthetic parameters (g_m, V_{cmax}, J_{max}) and morpho-physiological measurements (LMA, N_a), which were usually found across environmental treatments (e.g. Harley *et al.* 1992b; Flexas *et al.*, 2008; Galmés *et al.*, 2011), were shown here, for the first time, valid for the variation across genotypes of the same genetic background (Figs. 5A & 6). Therefore, 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.

Based on the genetic variation of physiological components underlying A and TE we explored the ideotype design by constituting alleles which contain loci influencing different components of physiological process of photosynthesis. The suggested virtual ideotypes could be obtained by more rounds of introgression to break any gene linkage within the genome segments of our present ILs. Model calculation showed that these ideotypes can potentially improve A and TE by 17.0% and 25.1%, respectively, compared with the best genotype we investigated. Besides, our analysis using ILs highlights the possibility to improve both A and TE within the same genetic background. Further experimental data with more ILs especially under field conditions can strengthen this conclusion. Of course, improvement of A and TE could also be achieved by broadening the genetic background. Recent advance in genome wide association studies (e.g. Huang *et al.*, 2010) would enhance this approach.

APPENDIX

Symbols and abbreviations

A	Net photosynthesis rate (μmol m^{-2} s^{-1})
A_{max}	Light saturated net photosynthesis at ambient CO_2 and O_2 level (μmol m^{-2} s^{-1})
A_c	Rubisco activity limited net photosynthesis rate (μmol m^{-2} s^{-1})
A_j	Electron transport limited net photosynthesis rate (μmol m^{-2} s^{-1})
C_a	Ambient air CO_2 concentration (μmol mol^{-1}); for model analyses, C_a refers to leaf-surface CO_2 level, with boundary conductance already considered (~6.78 mol m^{-2} s^{-1}, from Li-Cor manual version 6.1)
C_i	Intercellular CO_2 concentration (μmol mol^{-1})
g_s	Stomatal conductance for CO_2 (mol m^{-2} s^{-1})
g_m	Mesophyll conductance (mol m^{-2} s^{-1})
g_{mo}	Residual mesophyll conductance in the g_m model Eqn (7) (mol m^{-2} s^{-1})
$g_{m(NRH-A)}$	g_m estimated by the NRH-A method (Yin & Struik 2009b) based on the data obtained from high C_i of CO_2 response curves and low I_{inc} levels of light response curves at 21% O_2 (mol m^{-2} s^{-1})
g_t	Diffusion conductance from ambient air to the site of carboxylation (mol m^{-2} s^{-1})
g_{to}	Residual diffusion conductance in the g_t model Eqn (9) (mol m^{-2} s^{-1})
I_{inc}	Photon flux density incident on leaves (μmol photon m^{-2} s^{-1})
J	e$^-$ transport rate through PSII used for NADP$^+$ reduction (μmol e$^-$ m^{-2} s^{-1})
J_{max}	Maximum value of J under saturated light (μmol e$^-$ m^{-2} s^{-1})
K_{mc}	Michaelis-Menten constant of Rubisco for CO_2 (μbar)
K_{mo}	Michaelis-Menten constant of Rubisco for O_2 (mbar)
O	Oxygen partial pressure (mbar)
R_d	Day respiration (respiratory CO_2 release in the light other than by photorespiration) (μmol m^{-2} s^{-1})
R_{dk}	Respiratory CO_2 release in the dark (μmol m^{-2} s^{-1})
s	A lumped parameter, see Eqn (5) (-)
$S_{c/o}$	Relative CO_2/O_2 specificity factor for Rubisco (mbar μbar^{-1})
V_{cmax}	Maximum rate of Rubisco activity-limited carboxylation (μmol m^{-2} s^{-1})
δ_m	A parameter in the g_m model, defining $C_c : C_i$ ratio at saturating light (-), see Eqn (7)
δ_t	A parameters in the g_t model, defining $C_c : C_a$ ratio at saturating light (-), see Eqn (9)
δ_s	A parameters in the g_s model, defining $C_i : C_a$ ratio at saturating light (-), see Eqn (11)
κ_{2LL}	Value of conversion efficiency of incident light into J at the strictly limiting light [mol e$^-$ (mol photon)$^{-1}$]
θ	Convexity factor for response of J to I_{inc} (-), see Eqn (4)
Φ_2	Apparent quantum efficiency of PSII e$^-$ flow on PSII-absorbed light basis [mol e$^-$ (mol photon)$^{-1}$]
N_a	Leaf nitrogen per unit area (g N m^{-2} leaf)
Γ_*	C_c based CO_2 compensation point in the absence of R_d (μbar)
FS	Combination of flowering stage and drought-stressed treatment
FW	Combination of flowering stage and well-watered treatment
GS	Combination of grain filling stage and drought-stressed treatment
GW	Combination of grain filling stage and well-watered treatment
LMA	Leaf mass per area (g m^{-2} leaf)
TE	Transpiration efficiency (mmol mol^{-1})

Supplementary Materials in Chapter 3

Table S1. Photosynthesis parameters estimated from gas exchange and chlorophyll fluorescence measurements for all the 13 lines

	Genotype	FvCB model parameters							Other parameters					
		R_d	κ_{2LL}	J_{max}	θ	δ_m	V_{cmax}	δ	R_{dk}	s	g_m(NRH-A)	δ_s	g_m/g_s	J_{max}/V_{cmax}
FS	7	0.311 (0.113)	0.359 (0.030)	199.0 (8.4)	0.775 (0.100)	1.046 (0.090)	172.1 (16.1)	0.501 (0.026)	1.069 (0.508)	0.514 (0.009)	0.075 (0.005)	0.959	1.09	1.16
	37	0.923 (0.177)	0.329 (0.011)	177.6 (2.8)	0.784 (0.037)	1.278 (0.089)	155.1 (6.8)	0.491 (0.018)	1.372 (0.164)	0.465 (0.016)	0.088 (0.008)	0.797	1.60	1.15
	42	-	-	-	-	-	-	-	-	-	-	-	-	-
	69	0.644 (0.234)	0.361 (0.013)	199.3 (3.7)	0.756 (0.046)	0.907 (0.068)	169.8 (8.6)	0.451 (0.018)	1.609 (0.277)	0.514 (0.020)	0.068 (0.006)	0.896	1.01	1.17
	84	1.360 (0.252)	0.396 (0.027)	207.8 (6.8)	0.794 (0.077)	0.756 (0.059)	193.4 (13.5)	0.350 (0.016)	0.978 (0.246)	0.577 (0.022)	0.062 (0.004)	0.652	1.16	1.07
	100	0.989 (0.253)	0.363 (0.027)	185.3 (7.4)	0.767 (0.096)	1.020 (0.111)	160.0 (11.8)	0.420 (0.027)	1.058 (0.230)	0.517 (0.023)	0.082 (0.006)	0.714	1.43	1.16
	130	0.776 (0.243)	0.339 (0.015)	202.4 (4.5)	0.797 (0.049)	0.565 (0.029)	204.8 (20.6)	0.301 (0.011)	1.479 (0.192)	0.485 (0.020)	0.055 (0.004)	0.644	0.88	0.99
	157	0.644 (0.151)	0.346 (0.041)	189.3 (10.4)	0.824 (0.119)	0.949 (0.061)	165.6 (6.8)	0.415 (0.016)	1.327 (0.386)	0.506 (0.012)	0.073 (0.005)	0.737	1.29	1.14
	159	0.760 (0.212)	0.342 (0.029)	233.3 (11.3)	0.826 (0.091)	0.574 (0.032)	224.1 (15.3)	0.294 (0.012)	1.727 (0.234)	0.499 (0.017)	0.063 (0.007)	0.602	0.95	1.04
	161	0.669 (0.193)	0.344 (0.044)	214.8 (14.9)	0.790 (0.151)	0.524 (0.019)	213.3 (12.0)	0.267 (0.008)	1.330 (0.331)	0.488 (0.016)	0.053 (0.003)	0.544	0.96	1.01
	164	0.381 (0.189)	0.335 (0.021)	222.6 (7.8)	0.798 (0.072)	0.532 (0.024)	229.3 (21.1)	0.283 (0.010)	1.511 (0.329)	0.482 (0.015)	0.062 (0.005)	0.607	0.88	0.97
	H	1.235 (0.172)	0.335 (0.011)	186.3 (2.9)	0.870 (0.030)	0.643 (0.039)	208.7 (18.6)	0.325 (0.014)	1.314 (0.406)	0.496 (0.014)	0.050 (0.004)	0.658	0.98	0.89
	S	0.539 (0.177)	0.321 (0.029)	174.6 (7.6)	0.806 (0.098)	0.572 (0.031)	173.3 (14.5)	0.279 (0.013)	1.134 (0.303)	0.461 (0.015)	0.041 (0.003)	0.546	1.05	1.01
	Average	**0.769**	**0.348**	**199.4**	**0.799**	**0.780**	**189.1**	**0.365**	**1.326**	**0.500**	**0.064**	**0.696**	**1.11**	**1.06**
FW	7	0.476 (0.168)	0.363 (0.008)	213.8 (2.7)	0.744 (0.031)	0.969 (0.095)	177.8 (23.1)	0.562 (0.038)	1.642 (0.477)	0.513 (0.013)	0.092 (0.012)	1.337	0.72	1.20
	37	0.459 (0.237)	0.299 (0.020)	172.5 (5.2)	0.793 (0.070)	1.571 (0.114)	127.3 (3.8)	0.677 (0.020)	1.038 (0.257)	0.431 (0.019)	0.100 (0.006)	1.189	1.32	1.36
	42	0.692 (0.239)	0.365 (0.039)	197.2 (11.4)	0.665 (0.167)	0.885 (0.089)	157.0 (10.5)	0.493 (0.028)	1.664 (0.368)	0.500 (0.023)	0.083 (0.012)	1.115	0.79	1.26
	69	0.745 (0.186)	0.346 (0.025)	177.1 (6.9)	0.612 (0.124)	1.316 (0.125)	135.4 (7.8)	0.663 (0.034)	1.571 (0.137)	0.476 (0.015)	0.089 (0.010)	1.337	0.98	1.31
	84	0.855 (0.224)	0.354 (0.030)	196.1 (8.6)	0.738 (0.112)	0.761 (0.041)	168.3 (7.6)	0.403 (0.014)	1.317 (0.389)	0.506 (0.018)	0.059 (0.004)	0.856	0.89	1.17
	100	0.415 (0.148)	0.310 (0.018)	193.6 (6.7)	0.721 (0.084)	1.126 (0.089)	143.9 (5.7)	0.523 (0.018)	1.625 (0.262)	0.445 (0.012)	0.081 (0.010)	0.975	1.15	1.35
	130	0.983 (0.188)	0.358 (0.028)	224.8 (14.9)	0.788 (0.092)	0.561 (0.032)	190.3 (7.8)	0.332 (0.009)	1.299 (0.564)	0.513 (0.016)	0.059 (0.004)	0.815	0.69	1.18
	157	0.088 (0.153)	0.334 (0.043)	204.8 (7.8)	0.752 (0.171)	0.770 (0.043)	182.6 (13.2)	0.445 (0.018)	2.003 (0.313)	0.477 (0.012)	0.093 (0.012)	1.055	0.73	1.12
	159	0.603 (0.214)	0.325 (0.028)	185.4 (7.3)	0.848 (0.080)	0.901 (0.121)	128.0 (8.2)	0.508 (0.018)	1.551 (0.485)	0.481 (0.017)	0.082 (0.007)	1.166	0.77	1.45
	161	0.474 (0.175)	0.332 (0.033)	222.7 (13.5)	0.744 (0.136)	0.697 (0.036)	204.0 (15.4)	0.401 (0.016)	1.606 (0.120)	0.467 (0.014)	0.059 (0.004)	0.943	0.74	1.09
	164	0.395 (0.174)	0.337 (0.035)	207.3 (12.3)	0.720 (0.147)	0.545 (0.026)	178.9 (7.1)	0.319 (0.012)	1.377 (0.495)	0.482 (0.014)	0.063 (0.006)	0.769	0.71	1.16
	H	1.208 (0.169)	0.319 (0.028)	213.6 (10.3)	0.856 (0.085)	0.681 (0.043)	197.6 (11.2)	0.325 (0.013)	1.841 (0.521)	0.492 (0.014)	0.058 (0.003)	0.623	1.09	1.08
	S	0.407 (0.139)	0.326 (0.038)	151.1 (8.1)	0.764 (0.143)	1.044 (0.103)	120.7 (7.2)	0.574 (0.032)	0.896 (0.471)	0.472 (0.012)	0.083 (0.011)	1.275	0.82	1.25
	Average	**0.600**	**0.336**	**196.9**	**0.750**	**0.910**	**162.4**	**0.479**	**1.495**	**0.481**	**0.077**	**1.035**	**0.88**	**1.23**

Continue

Genotype	FvCB model parameters							Other parameters					
	R_d	κ_{2LL}	J_{max}	θ	δ_m	V_{cmax}	δ_t	R_{dk}	s	$g_{m(NRH-A)}$	δ_s	g_m/g_s	J_{max}/V_{cmax}
GS 7	0.059 (0.130)	0.356 (0.042)	152.0 (7.5)	0.817 (0.120)	0.478 (0.022)	160.2 (15.3)	0.296 (0.010)	0.653 (0.576)	0.496 (0.010)	0.034 (0.002)	0.779	0.61	0.95
37	0.413 (0.178)	0.320 (0.028)	124.5 (4.2)	0.912 (0.060)	0.614 (0.024)	122.0 (4.4)	0.323 (0.007)	1.196 (0.135)	0.464 (0.013)	0.037 (0.003)	0.680	0.90	1.02
42	0.290 (0.165)	0.328 (0.019)	151.9 (4.0)	0.802 (0.064)	0.516 (0.028)	146.5 (7.4)	0.291 (0.011)	1.239 (0.416)	0.464 (0.013)	0.033 (0.002)	0.666	0.77	1.04
69	0.005 (0.194)	0.341 (0.055)	135.5 (9.0)	0.847 (0.152)	0.765 (0.036)	123.9 (5.0)	0.418 (0.013)	0.879 (0.274)	0.474 (0.015)	0.040 (0.002)	0.919	0.83	1.09
84	0.109 (0.183)	0.337 (0.037)	132.5 (5.9)	0.846 (0.103)	0.862 (0.042)	115.6 (3.6)	0.435 (0.010)	0.780 (0.585)	0.477 (0.014)	0.045 (0.003)	0.880	0.98	1.15
100	0.445 (0.236)	0.320 (0.051)	121.6 (7.8)	0.870 (0.137)	0.790 (0.081)	125.2 (15.4)	0.405 (0.030)	1.304 (0.779)	0.458 (0.018)	0.034 (0.002)	0.833	0.95	0.97
130	0.835 (0.178)	0.319 (0.021)	194.4 (5.7)	0.897 (0.048)	0.553 (0.023)	203.4 (15.7)	0.333 (0.010)	1.227 (0.201)	0.468 (0.014)	0.049 (0.003)	0.835	0.66	0.96
157	0.817 (0.162)	0.316 (0.059)	151.7 (11.1)	0.886 (0.137)	0.678 (0.042)	151.2 (10.3)	0.380 (0.016)	1.179 (0.335)	0.469 (0.013)	0.040 (0.002)	0.863	0.79	1.00
159	0.869 (0.141)	0.310 (0.043)	153.6 (7.9)	0.902 (0.091)	0.696 (0.035)	141.8 (5.3)	0.366 (0.013)	1.234 (0.197)	0.453 (0.010)	0.049 (0.005)	0.771	0.90	1.08
161	0.317 (0.147)	0.322 (0.053)	153.2 (10.4)	0.863 (0.137)	0.767 (0.039)	133.4 (4.6)	0.450 (0.016)	0.742 (0.305)	0.465 (0.011)	0.056 (0.005)	1.091	0.70	1.15
164	0.494 (0.152)	0.305 (0.033)	135.3 (5.7)	0.890 (0.079)	0.789 (0.072)	116.6 (6.4)	0.419 (0.016)	1.104 (0.159)	0.434 (0.012)	0.045 (0.004)	0.895	0.88	1.16
H	1.055 (0.209)	0.331 (0.025)	179.7 (6.2)	0.890 (0.060)	0.592 (0.037)	179.3 (11.8)	0.307 (0.013)	1.226 (0.138)	0.482 (0.018)	0.050 (0.004)	0.637	0.93	1.00
S	0.365 (0.124)	0.316 (0.020)	151.4 (3.9)	0.852 (0.054)	0.694 (0.035)	141.1 (5.5)	0.376 (0.012)	1.063 (0.125)	0.452 (0.010)	0.054 (0.004)	0.820	0.85	1.07
Average	0.467	0.325	149.0	0.867	0.676	143.1	0.369	1.064	0.466	0.044	0.821	0.83	1.05
GW 7	-0.038 (0.240)	0.335 (0.027)	156.6 (5.4)	0.864 (0.072)	0.727 (0.074)	143.5 (14.2)	0.469 (0.033)	1.130 (0.386)	0.484 (0.019)	0.058 (0.008)	1.318	0.55	1.09
37	0.399 (0.194)	0.291 (0.041)	150.1 (9.1)	0.897 (0.099)	0.947 (0.082)	129.7 (7.4)	0.534 (0.027)	0.866 (0.352)	0.433 (0.016)	0.055 (0.007)	1.225	0.77	1.16
42	0.307 (0.187)	0.348 (0.021)	172.3 (5.0)	0.771 (0.073)	0.477 (0.033)	160.4 (11.9)	0.314 (0.017)	1.254 (0.192)	0.492 (0.015)	0.039 (0.004)	0.915	0.52	1.07
69	0.114 (0.196)	0.337 (0.067)	139.6 (12.4)	0.730 (0.266)	0.856 (0.054)	115.2 (6.1)	0.589 (0.023)	1.303 (0.336)	0.463 (0.017)	0.049 (0.004)	1.890	0.45	1.21
84	0.723 (0.134)	0.351 (0.066)	164.6 (13.6)	0.817 (0.193)	0.768 (0.051)	144.8 (6.9)	0.455 (0.019)	1.506 (0.476)	0.498 (0.011)	0.064 (0.007)	1.120	0.69	1.14
100	0.511 (0.196)	0.305 (0.023)	122.0 (3.9)	0.898 (0.057)	0.784 (0.063)	105.3 (5.1)	0.438 (0.016)	1.144 (0.187)	0.439 (0.016)	0.039 (0.002)	0.995	0.79	1.16
130	0.569 (0.128)	0.303 (0.046)	150.5 (9.1)	0.893 (0.109)	0.809 (0.050)	137.9 (7.3)	0.443 (0.018)	1.012 (0.470)	0.444 (0.010)	0.048 (0.003)	0.978	0.83	1.09
157	0.433 (0.117)	0.321 (0.019)	199.6 (6.0)	0.857 (0.056)	0.452 (0.025)	195.8 (14.0)	0.282 (0.014)	1.479 (0.507)	0.468 (0.009)	0.047 (0.003)	0.747	0.60	1.02
159	0.692 (0.201)	0.325 (0.042)	147.5 (7.5)	0.885 (0.097)	0.685 (0.063)	129.3 (8.4)	0.455 (0.029)	1.241 (0.318)	0.463 (0.015)	0.055 (0.006)	1.358	0.50	1.14
161	0.059 (0.235)	0.297 (0.031)	143.9 (6.0)	0.884 (0.078)	1.340 (0.158)	101.5 (4.7)	0.723 (0.026)	1.291 (0.265)	0.432 (0.019)	0.055 (0.006)	1.570	0.85	1.42
164	0.457 (0.171)	0.275 (0.047)	99.8 (6.9)	0.892 (0.139)	1.024 (0.155)	77.4 (5.2)	0.568 (0.025)	0.750 (0.511)	0.410 (0.014)	0.046 (0.007)	1.277	0.80	1.29
H	1.021 (0.203)	0.306 (0.047)	184.2 (13.8)	0.882 (0.129)	0.633 (0.057)	187.4 (24.6)	0.389 (0.028)	1.255 (0.425)	0.470 (0.017)	0.046 (0.006)	1.006	0.63	0.98
S	-0.147 (0.172)	0.312 (0.050)	142.2 (9.7)	0.823 (0.160)	1.211 (0.139)	102.0 (4.9)	0.635 (0.021)	0.992 (0.349)	0.445 (0.013)	0.059 (0.006)	1.333	0.91	1.39
Average	0.392	0.316	151.8	0.853	0.824	133.1	0.484	1.171	0.457	0.051	1.210	0.69	1.17

FS, flowering stage – drought stressed environment; FW, flowering stage – well watered environment; GS, grain filling stage – drought stressed environment; GW, grain filling stage – well watered environment. For definitions see Appendix.

71

Table S2. Simple correlation coefficients among seven parameters of the photosynthesis model at four stage × treatment combinations (FS, FW, GS, GW) based on data of 11 ILs and their parents. For definitions of the parameters and abbreviations, see table in the Appendix.

		R_d	κ_{2LL}	J_{max}	θ	δ_m	V_{cmax}
FS	κ_{2LL}	0.41					
	J_{max}	-0.18	0.22				
	θ	0.32	-0.37	0.02			
	δ_m	0.07	0.21	-0.53	-0.43		
	V_{cmax}	-0.02	-0.13	0.79**	0.45	-0.84**	
	δ_s	-0.20	0.31	-0.26	-0.46	0.79**	-0.61*
FW	κ_{2LL}	0.23					
	J_{max}	0.25	0.39				
	θ	0.22	-0.44	0.13			
	δ_m	-0.24	-0.44	-0.71**	-0.23		
	V_{cmax}	0.24	0.35	0.92**	0.14	-0.77**	
	δ_s	-0.46	0.08	-0.60*	-0.37	0.70**	-0.66*
GS	κ_{2LL}	-0.57*					
	J_{max}	0.59*	0.01				
	θ	0.70**	-0.65*	0.12			
	δ_m	-0.20	-0.31	-0.55	0.20		
	V_{cmax}	0.57*	0.12	0.94**	0.10	-0.70**	
	δ_s	-0.33	-0.09	-0.16	-0.03	0.63*	-0.32
GW	κ_{2LL}	-0.08					
	J_{max}	0.25	0.50				
	θ	0.35	-0.71**	-0.19			
	δ_m	-0.49	-0.53	-0.62*	0.18		
	V_{cmax}	0.43	0.44	0.95**	-0.10	-0.79**	
	δ_s	-0.49	-0.01	-0.49	-0.34	0.61*	-0.60*

*, ** significant at the level of $P < 0.05$ and 0.01, respectively.
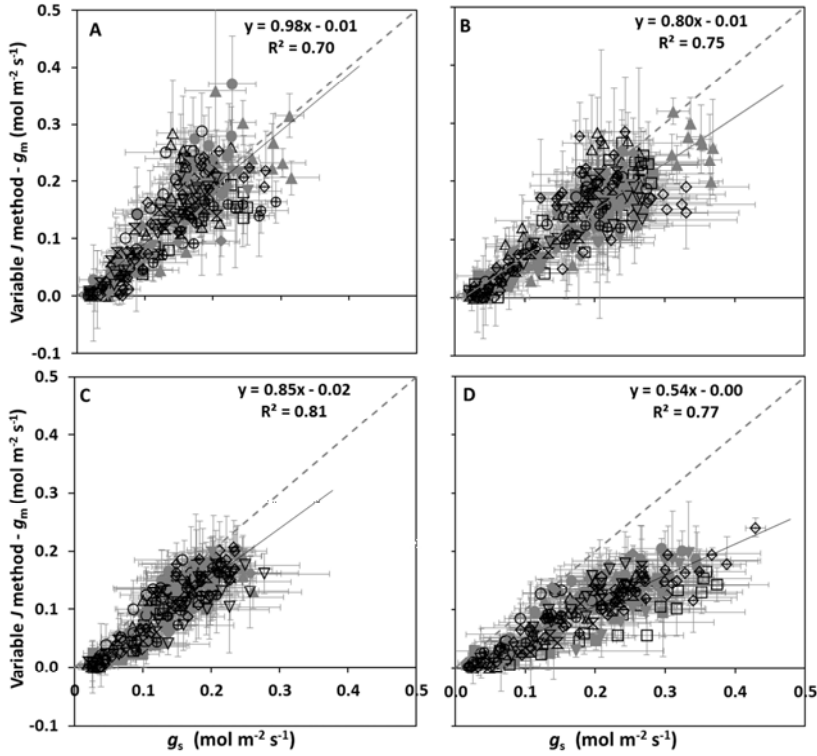
Fig. S1. Relationship between stomatal conductance (g_s) and mesophyll conductance (g_m) calculated using the variable *J* method (Harley *et al.* 1992b) at (A) flowering–drought-stressed treatment (FS), (B) flowering-well-watered treatment (FW), (C) grain filling-drought-stressed treatment (GS), and (D) grain filling-well-watered (GW). Values are means ± standard error of four replicates at each light intensity (10, 30, 50, 70, 100, 170, 500, 1000, 1500, and 2000 μmol m^{-2} s^{-1}) and CO_2 concentration (50, 60, 70, 80, 100, 150, 250, 380, 650, 1000, and 1500 μmol mol^{-1}), based on measurements under the 21% O_2 condition. Symbols and accessions as follows: ▲ IL7, △ IL37, ■ IL42, □ IL69, ● IL84, ○ IL100, ◆ IL130, ◇ IL157, ▼ IL159, ▽ IL161, ⊕ IL164, ⬦ Haogelao, and ⊠ Shennong265. The linear regression lines (solid line) are based on all the data points for each stage × treatment combination. The diagonal (dashed line) is the 1:1 line.
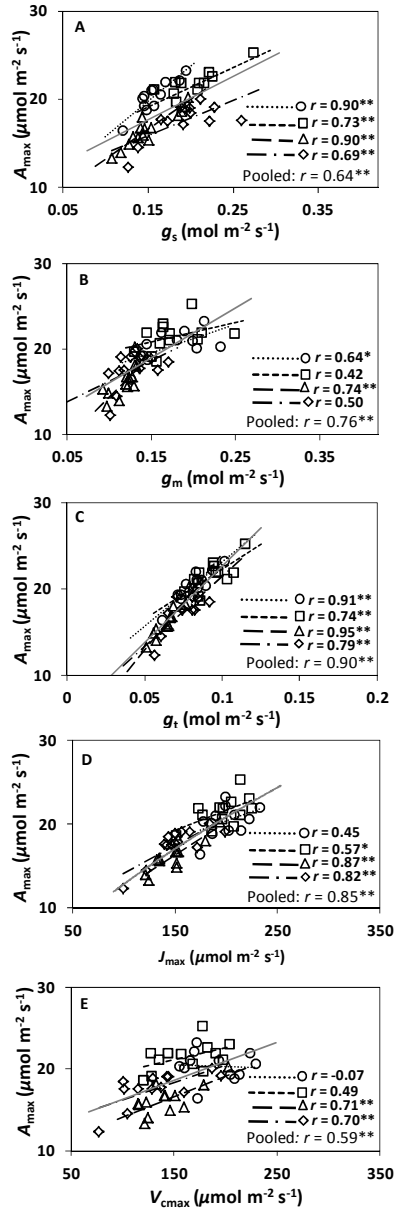
Fig. S2. Relationships between light-saturated net photosynthesis rate (A_{max}) (380 μmol mol^{-1} CO_2, 1500 μmol m^{-2} s^{-1} light intensity, 25°C) and (A) stomatal conductance to CO_2 (g_s), (B) mesophyll conductance to CO_2 (g_m), (C) total diffusion conductance to CO_2, including g_s and g_m (g_t), (D) electron transport capacity (J_{max}), and (E) Rubisco carboxylation capacity (V_{cmax}). Linear regressions were fitted for overall data (grey solid lines) and each stage × treatment combination: flowering-drought-stressed treatment (FS, ○, ·······), flowering-well-watered environment (FW, □, -----), grain filling-drought-stressed environment (GS, △, — —), and grain filling-well watered environment (GW, ◇, — · —). The significance of each correlation was shown as: *, $P < 0.05$; **, $P < 0.01$.

74
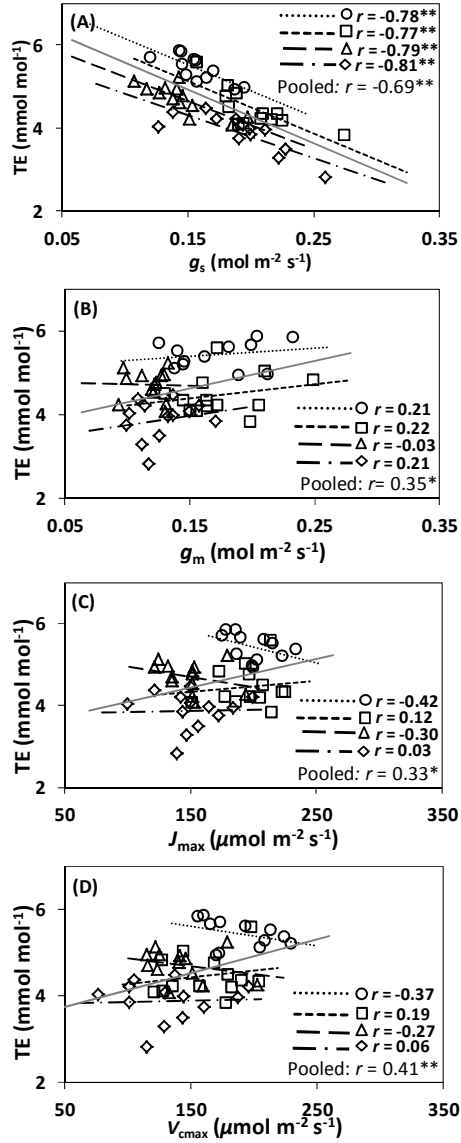
Fig. S3. Relationships between transpiration efficiency (TE) (380 μmol mol^{-1} CO$_2$, 1500 μmol m^{-2} s^{-1} light intensity, 25°C, and vapour pressure deficit of 1.5 kPa) and (A) stomatal conductance to CO$_2$ (g_s); (B) mesophyll conductance to CO$_2$ (g_m); (C) electron transport capacity (J_{max}); (D) Rubisco carboxylation capacity (V_{cmax}). The symbols, abbreviations, and significance levels were as in Fig. 3.

CHAPTER 4

Molecular marker-based crop modelling to predict variation in yield among rice (*Oryza sativa* L.) genotypes grown under well-watered and drought-stress conditions

Junfei Gu [a], Xinyou Yin [a], Chengwei Zhang [b], Huaqi Wang [b], and Paul C. Struik [a]

[a] Centre for Crop Systems Analysis, Department of Plant Science, Wageningen University, P. O. Box 430, 6700 AK Wageningen, The Netherlands
[b] Plant Breeding & Genetics, China Agricultural University, 100193 Beijing, P.R. China

Chapter 4

ABSTRACT

Marker-assisted breeding can be enhanced by accurate, model-based prediction of phenotypic variation in crop yield. We measured input parameters of the crop simulation model GECROS and created molecular marker-based estimates of these parameters from estimated additive allele effects for a rice (*Oryza sativa* L.) population of 96 introgression lines (ILs). We then compared the ability of two versions of the model to predict the yield of ILs under well-watered and droughted conditions, one version with the measured model parameters and one with the marker-based estimates as input. The total variation in yield accounted for was 72% under well-watered conditions and 57% under drought when measured parameters were used, but 52% and 47%, respectively, when marker-based parameters were used. Regression analyses showed that 'total crop nitrogen uptake' had the most significant effect on yield; five other model parameters also significantly influenced yield, but seed dry weight did not. Using the marker-based estimates of model parameters, GECROS also gave a fair prediction of variation in yield among 251 recombinant inbred lines of the same parents under either well-watered or drought conditions. The model-based approach detected more markers than marker selection using multiple regression for yield. Markers most important for determining yield differences among the ILs were on Chromosomes 2 and 3 for well-watered and drought environments, respectively. Further research should aim at (1) upgrading the GECROS model for rice grown under drought, and (2) breaking the putative genetic linkage between high photosynthesis and low yield exhibited in the IL population.

Key words: QTL, ecophysiological crop modelling, model-based breeding, genotype × environment interactions, rice, introgression lines, photosynthesis, recombinant inbred lines, *Oryza sativa.*

INTRODUCTION

An increase in yield of crop cultivars for both favourable and stress-environments is required to feed the growing world population. In rice (*Oryza sativa* L.), breeders have been successful in improving yield during the last 60 years (Peng *et al.*, 2008), through extensive, largely empirical, selection. Developments in genomics provided useful tools and information for dissecting complex traits into single genetic determinants, quantitative trait loci (QTLs). QTLs related to important agronomic traits, such as yield and stress tolerance, have been mapped, cloned and characterized [e.g. Xing *et al.*, 2008; see reviews by Miura *et al.* (2011) and Xing & Zhang (2011)]. These developments have provided a firm basis for further improving yield through marker-assisted selection (MAS) or genetic transformation of crops. However, selection for, or

transformation of, only a few or even a complex of genes may not result into a major yield increase (Sinclair *et al.*, 2004; Yin & Struik, 2008). Furthermore, MAS has hardly been proven successful in breeding for complex traits like yield, which have low heritabilities and exhibit strong genotype × environment (G × E) interactions (Collard & Mackill, 2008).

The complexity of the yield trait stems from its many underlying processes, which are often environmentdependent and show strong feed-back and feed-forward mechanisms during crop growth. Crop yield can be analysed and evaluated using ecophysiological crop growth simulation models that integrate information about processes at lower levels (Yin & Struik, 2008; Hammer *et al.*, 2010; Zhu *et al.*, 2011). Such models are based on crop-physiological knowledge and can quantify causality between relevant physiological processes and responses of these processes to environmental variables (e.g. irradiance, temperature, availability of water and nutrients). By feeding crop models with weather data from other locations, these models could predict yield beyond the environments in which the model parameters were derived and could explain variation in yield of a specific genotype among contrasting environments (Yin *et al.*, 2000a; Sinclair, 2011).

A major challenge for the use of crop models is to predict phenotypic differences between relatively similar lines from a genetic population on the basis of genotype-specific model parameters (Yin *et al.*, 2000a,b; Reymond *et al.*, 2003; Prudent *et al.*, 2011). These model parameters are often referred to as 'genetic coefficients' (Messina *et al.*, 2006; White *et al.,* 2008) and are supposed to be little affected by variation in environment (Yin *et al.*, 2000a). Modelling could thus assist in quantifying the G × E interactions (Yin *et al.*, 2004; Reymond *et al.,* 2004; Hammer *et al.*, 2005; Yin *et al.*, 2005b; Hammer *et al.*, 2006; Chenu *et al.*, 2008; Tardieu & Tuberosa, 2010), and in predicting genotype-to-phenotype relationships (Bertin *et al.,* 2010; Messina *et al.,* 2011).

Advances in the use of molecular markers enable genetic information on physiological traits to be integrated into crop models as determined QTLs for physiologically important parameters, making model parameters genotype specific (Yin *et al.*, 2000b; Reymond *et al.*, 2003, 2004; Nakagawa et al., 2005; Quilot et al., 2005; Uptmoor et al. 2008; Xu *et al.,* 2011). This 'QTL-based modelling' approach can dissect complex traits (e.g. yield) into physiologically relevant component traits, integrate effects of QTLs on the component traits over time and space at whole-crop level, and predict yield of various allele combinations under different environmental conditions (Tardieu & Tuberosa, 2010; Yin & Struik, 2010). In general, the 'QTL-based modelling' approachfollows the different steps described by Yin *et al.* (2000b, 2004): (i) designing and validating an ecophysiological model; (ii) identifying QTLs

and their impact on model parameters following the common marker linkage analysis and QTL mapping approaches; (iii) creating a QTL-based version of the ecophysiological model with QTL allelic information included; (iv) validating whether the QTL-based model can help to understand the G × E interactions within populations.

Such a QTL-based modelling approach was proven to be robust in predicting genetic differences in comparatively simple traits, such as leaf elongation rate in maize, flowering time of barley or rice, and fruit quality in bi-parental crossing populations under different environmental conditions (in terms of vapour pressure deficit, soil moisture content, temperature and photoperiod) (Reymond *et al.*, 2003; Quilot *et al.*, 2004, 2005; Nakagawa *et al.*, 2005; Yin *et al.*, 2005b; Uptmoor *et al.*, 2008, 2012; Bertin *et al.*, 2010; Prudent *et al.*, 2011). Only in a few cases was the QTL-based modelling approach used to predict yield (Yin *et al*., 2000b). To the best of our knowledge, this approach has not been used to incorporate genetic variation in photosynthesis into crop model analysis.

In the present study, we followed the same approach and used the crop model GECROS (Genotype-by-Environment interaction on CROp growth Simulator, Yin & van Laar, 2005) to predict variation in grain yield and biomass of biparental crosses of rice under well-watered and drought-stress environments. Based on the experience of using an older crop model (Yin *et al*., 2000a,b), the GECROS model was designed in such a way that most of its input parameters are close to the traits breeders score for selection. In our previous studies (Chapter 2 and 3), we have analysed genetic control and physiological basis of the genetic variation in leaf photosynthesis and its sensitivity to drought. Whether or not leaf photosynthesis affects biomass and grain yield in our rice populations warrants further analysis, especially considering the inconsistency in the literature with regards to the role of photosynthesis in determining crop biomass or yields (e.g. Richards, 2000; Fischer & Edmeades, 2010). We will, therefore, integrate physiological and genetic effects on leaf photosynthesis with other GECROS model parameters.

We aimed: (1) to examine the ability of the GECROS model with measured model parameter values, or with marker-based estimates of model parameters, to account for yield differences among introgression lines (ILs) of rice; (2) to test the extrapolation ability of the marker-based approach to predict yield variation in an independent population of recombinant inbred lines (RILs) derived from the same parents, and (3) to analyze the relative importance of individual markers in accounting for variation in yield and to examine whether the markers for leaf photosynthesis are also important for grain yield. We focus on the analysis of yield traits of the ILs and of the RILs,

illustrating that the model approach can enhance MAS for inbred breeding to improve yields. Thus, markers, instead of QTLs, are identified for our analysis.

MATERIAL AND METHODS

Plant material and field experiments

The genetic population consisted of 94 advanced backcross ILs and two parents, as described in Chapter 2. The parents were the lowland rice cv. Shennong265 (*Japonica*, recurrent parent) and the upland rice cv. Haogelao (*Indica-Japonica* intermediate, donor parent). Haogelao is drought tolerant, but low-yielding; Shennong265 is drought susceptible, but high-yielding under irrigated conditions. After a cross between the two parents, the resultant F_1 plants were backcrossed with Shennong265 three times, and these BC_3F_1 plants were consecutively self-pollinated five times to construct the genetic population by the single-seed descent method.

Field experiments were conducted to assess model parameters and to measure grain yield and shoot biomass (two major model-output traits). The ILs and the two parents were sown on 10th May, 2009 by direct-seeding at the experimental station of China Agricultural University, Beijing (39°54'N, 116°24'E; elevation 50 m above sea level), China. The mean annual temperature is 13.7°C; the total annual precipitation is 486 mm; the mean daily global radiation is 14 MJ m^{-2} d^{-1}. The soil is classified as a calciaquoll, which contains 23.5% sand, 57.1% silt, and 19.4% clay. The field experiment design followed a randomized complete block design, with two replications, four rows of 2.5 m per plot, 0.30 m between rows, in both rainfed upland and fully irrigated lowland conditions. Seed was hand sown at a depth of 0.03 to 0.04 m. At seedling stage, plants were thinned to a 0.075 m distance between plants within each row (resulting in a plant density of 44.4 plants per m^2). Weeds in both conditions were controlled by a combination of chemical and manual methods. Insects and diseases were controlled chemically. Basal fertilizer application included 48 kg N ha^{-1} (as urea), 120 kg P_2O_5 ha^{-1} and 100 kg K_2O ha^{-1}, and an additional 86 kg N ha^{-1} was applied at the tillering stage and 28 kg N ha^{-1} at the booting stage. For fully irrigated lowland conditions, rice was grown under continual flooding until harvest. For rainfed upland conditions, besides rainfall, irrigation was only applied when necessary at critical stages (i.e. at sowing, 120 mm; at tillering, 150 mm; at booting, 130 mm).

An independent population of 251 RILs derived from the same parents (La, 2004; Zhang, 2006) was sown on 7th of May, 2005 by direct-seeding at the experimental station in Zhuozhou (39.29' N, 115.59' E; elevation 45 m above sea level), China (Zhang, 2006). The mean annual temperature is 13.6°C; the total annual precipitation

is 389 mm; the mean daily global radiation is 14 MJ m^{-2} d^{-1}. The experiment design and management are the same as in the field experiments in Beijing in 2009.

The crop growth model

The model used in this study was the crop growth model GECROS, first described by Yin & van Laar (2005). GECROS is a generic model that operates in daily time steps, simulates the growth and development of the crop on a daily basis, and generates phenotypes for a multitude of traits, based on concepts of the interaction and feedback mechanisms among various contrasting components of crop growth, carbon-nitrogen interaction in particular (Yin & Struik, 2010). The summary information about the latest GECROS model (v3.0) is given in Supplementary Materials (also see Yin, 2013). For a given set of model parameters and environmental conditions, the model produces predictions of grain yield and biomass.

Model inputs, parameterization, and test

The weather inputs for the GECROS model are daily radiation, vapour pressure, maximum temperature, minimum temperature, rainfall and wind speed. These required weather data were collected from a nearby weather station in 2005 (RILs) and 2009 (ILs) at Zhuozhou and Beijing, respectively. Atmospheric CO_2 concentration and the amount of irrigated water were also used as model input.

A complete set of model parameters (Table 1) was determined for each IL from data collected in a well-watered environment in 2009, which include individual seed dry weight (S_W), seed nitrogen concentration (n_{SO}), maximum plant height (H_{max}), the minimum number of days for vegetative growth phase (m_V) or for reproductive (seed fill) phase (m_R) provided both photoperiod and temperature are optimal, and specific leaf area constant (S_{la}). Table 1 also lists total crop nitrogen uptake at maturity (N_{max}) as a model parameter. N_{max} per se, as an accumulative quantity in the crop life cycle, is not considered as a model parameter of the original GECROS. However, there was not sufficient information about the soil, and modelling of nitrogen availability for transition between flooded and nonflooded soil environments is complex and usually full of uncertainties (Gaydon *et al.*, 2012). To reduce an influence of uncertainties in predicting edaphic variables for nitrogen supply, we took a simple approach, using N_{max} as a model parameter. The value of N_{max} was estimated based on dry weight and the nitrogen concentration in plant organs; n_{SO} was determined by means of micro-Kjeldahl digestion and distillation. Nitrogen concentration in straw was assumed to be conservative at 0.463% (see data of Singh *et al.*, 1998), and nitrogen accumulation in the roots was assumed to be 5% of N_{max} (Yin & van Laar, 2005).

Table 1. List of genotype-specific parameters of the GECROS model (see Materials and Methods). DM stands for dry matter.

Trait	Description	Unit
S_W	Seed dry weight	g DM seed^{-1}
n_{SO}	Seed (storage organ) N concentration	g N g^{-1} DM
H_{max}	Maximum plant height	m
m_V	Minimum days for vegetative growth phase	day
m_R	Minimum days for reproductive (seed fill) phase	day
S_{la}	Specific leaf area constant	m^2 leaf g^{-1} DM
N_{max}	Total crop N uptake at crop maturity*	g N m^{-2} ground

* Not an input parameter in the original GECROS (see the texts)

As in GECROS, parameters m_V and m_R are calculated based on a flexible bell-shaped nonlinear function of phenological response to temperature (Yin *et al.*, 2005a), flowering time, and harvest time for each IL. For other non-genotype-specific parameters, default values based on previous studies were used for all lines (Yin & van Laar, 2005).

Robust crop growth models can predict yield based on plant growth potential and whether the supply of carbohydrate and nitrogen can satisfy that potential (Hammer *et al.*, 2010). So, model parameters were estimated from the well-watered experiment (2009, Beijing) as plants could reach their 'potential' growth. To test the model, predicted dry grain yield and dry shoot biomass were compared with measured data for both ILs and RILs. To evaluate the quality of the model outputs, we used the relative root mean square error (*r*RMSE; Wallach *et al.*, 2006), calculated as the root mean square error divided by the mean of the observed value. In addition, the R^2 coefficient of linear regression between predicted and observed values was used to indicate the percentage of phenotypic variation accounted for by the model.

Statistical identification of important markers for model parameters and yield

A total of 130 SSR markers and their position for the IL population were reported previously (Chapter 2; see also the Supplementary material Fig. S1). In order to select markers which could be potentially used for breeding, the effects of markers were analyzed using a two-stage approach. Firstly, using the general linear model (GLM) procedure in the statistical package SAS 9.2 (SAS Inst. Inc.), one-way analysis of variance (ANOVA) was used to test the significance ($P < 0.05$) of markers across the whole genome. Secondly, all significant markers

were put into a multiple regression model (GLM) procedure in SAS, using Eqn 1 where a model parameter value Y (listed in Table 1) or yield of introgression line k, as affected by N markers, was presented as:

$$Y_k = \mu + \sum_{n=1}^{N} a_n M_{k,n} \qquad (1)$$

where μ = the intercept; a_n = the additive effect of the n-th marker; $M_{k,n}$ = genetic score of the n-th marker of the individual introgression line k that takes either the value -1 (allele coming from recurrent parent 'Shennong265') or 1 (donor parent 'Haogelao' allele present). We used the simple additive model (Eqn 1) in which Y reflects the breeding value, as additive genetic effects are predictably transmitted to progeny. For this second step, including all selected markers may lead to non-significant markers in the multiple regression because of the collinearity of markers. Such collinearity can sometimes lead to serious stability problems (Martens & Næs, 1992; Næs & Mevik, 2001). To solve this problem, the non-significant marker with the highest P-value during multiple regression was excluded in the next round of multiple regression. This approach left out one marker at a time, until all markers in the multiple regression became significant ($P < 0.05$) (Ott & Longnecker, 2001).

Using the same Eqn (1), marker-based values of GECROS-model parameters were calculated for each genotype, based on the estimated additive effects for each parameter and the marker allelic information of the ILs or the RILs.

Identifying important yield-determining traits and markers
Multiple linear regression analyses were performed to identify which model parameter in Table 1 influenced yield most. A sensitivity analysis using the GECROS model was performed to identify the contribution of single markers to yield production. This is achieved by following the approach of Yin *et al.* (2000a), i.e. examining yield variation accounted for by the GECROS model when the tested marker was excluded in estimating the marker-based model parameters. First, the baseline simulation was conducted, where IL-specific allelic values for all markers were used as input for simulation. Then, allelic values were fixed, one marker at a time, at zero. The extent to which the percentage of yield variation accounted for by GECROS was decreased relative to the percentage accounted for by the baseline simulation was used to rank the relative importance of the markers in determining grain yields. This model-based identification of markers was compared with the marker analysis based on yield data *per se*.

RESULTS

Variation in yield and physiological model parameters
There was no yield difference between replicates in either well-watered or drought-stressed conditions ($P > 0.05$). The IL population exhibited considerable genotypic

variation in model parameters and grain yield (Fig. 1), showing transgressive segregation. Most model parameters, i.e., S_W, n_{SO}, H_{max}, S_{la}, N_{max} (drought-stressed environment) and yield presented a unimodal distribution. For m_V, m_R and N_{max} (well-watered environment), a bimodal distribution was observed. Parent 'Shennong 265' yielded more than 'Haogelao', even under drought.
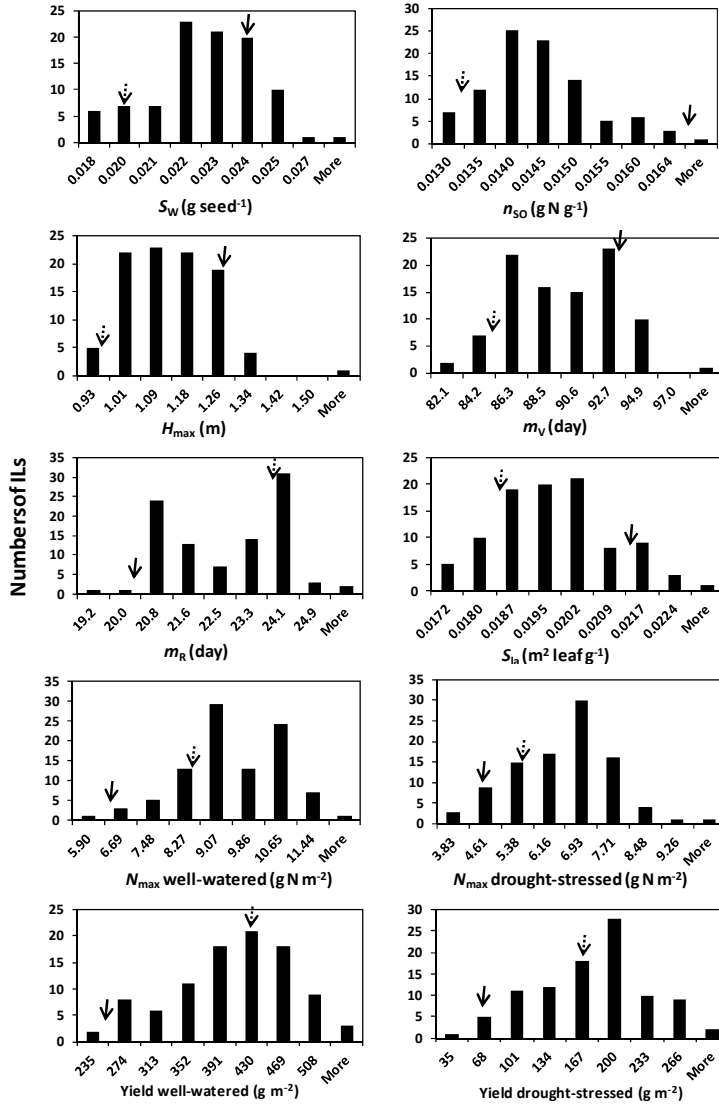
Fig. 1. Frequency distribution of seven model parameters and yield in the population of introgression lines (ILs). Arrows show values for the two parents (full arrow for 'Haogelao' and dotted arrow for 'Shennong265').

Table 2. Linear regression of rice yield (Y) against total crop N uptake (N_{max}) and one other parameter trait of both well-water and drought-stressed input parameters ($n = 96$; for definition of these traits, see Table 1).

Equation	b_0	b_1	b_2	R^2
$Y = b_0 + b_1 N_{max}$	-15.43 / -62.61†	44.49*** / 37.13***		0.576 / 0.592
$Y = b_0 + b_1 N_{max} + b_2 S_W$	-39.73 / -44.54	43.86*** / 37.37***	1356.6 / -881.1	0.577 / 0.593
$Y = b_0 + b_1 N_{max} + b_2 n_{SO}$	466.60 / 86.83	47.11*** / 35.94***	-35698.7*** / -10033.9*	0.749 / 0.614
$Y = b_0 + b_1 N_{max} + b_2 H_{max}$	77.43 / 48.84	49.70*** / 40.49***	-128.5** / -120.7***	0.609 / 0.646
$Y = b_0 + b_1 N_{max} + b_2 m_V$	630.66 / 553.40	39.03*** / 29.32***	-6.73*** / -6.41***	0.668 / 0.714
$Y = b_0 + b_1 N_{max} + b_2 m_R$	-310.43 / -339.14	38.54*** / 29.04***	15.69*** / 14.63***	0.669 / 0.710
$Y = b_0 + b_1 N_{max} + b_2 S_{la}$	161.52 / 92.18	42.74*** / 36.78***	-8334.2* / -7903.0**	0.598 / 0.625

*, **, *** Significant at the 0.05, 0.01, and 0.001 probability levels, respectively.

† Values before slash are for well-watered conditions; values after slash are for drought-stressed conditions.

Contribution of individual model parameters to yield

Effects of individual model parameters on yield were analyzed, for both well-watered and drought-stressed environments as assessed in 2009. Simple single-factor regression based on data of all genotypes revealed that yield correlated with most model parameters. Among them, N_{max} was correlated with yield most. N_{max} alone accounted for 57.6% and 59.2% of the variation in yield under well-watered and drought-stressed environments, respectively (Table 2). N_{max} was also associated with other model parameters. For example, under well-watered conditions, N_{max} correlated with S_w, H_{max}, m_V, and m_R ($r = 0.29$, 0.44, -0.29, and 0.32, respectively; $P < 0.01$). Therefore, N_{max} was used as covariate, when multiple regression was conducted relating yield to each model parameter (Table 2). The results showed yield correlated significantly with all model parameters, except for S_W. Amongst model parameters under well-watered conditions, n_{SO} was best correlated to yield besides N_{max}; under drought-stressed conditions m_V was best correlated to yield.

Performance of the GECROS model

The GECROS model was first evaluated using both well-watered and drought-stressed experiments in 2009. Under well-watered conditions, the model accounted for 72% of the variation in yield (Fig. 2A) and for 78% of the variation in biomass (Fig. 2B), with rRMSE values of 0.10 and 0.09, respectively.

For simulating yield in the drought-stressed environment, first, all parameter values as used for the well-watered environment were applied. This procedure resulted in systematic over-predictions, as the actual nitrogen uptake was much less, resulting in reduced growth under drought. Therefore, observed N_{max} from the drought-stressed
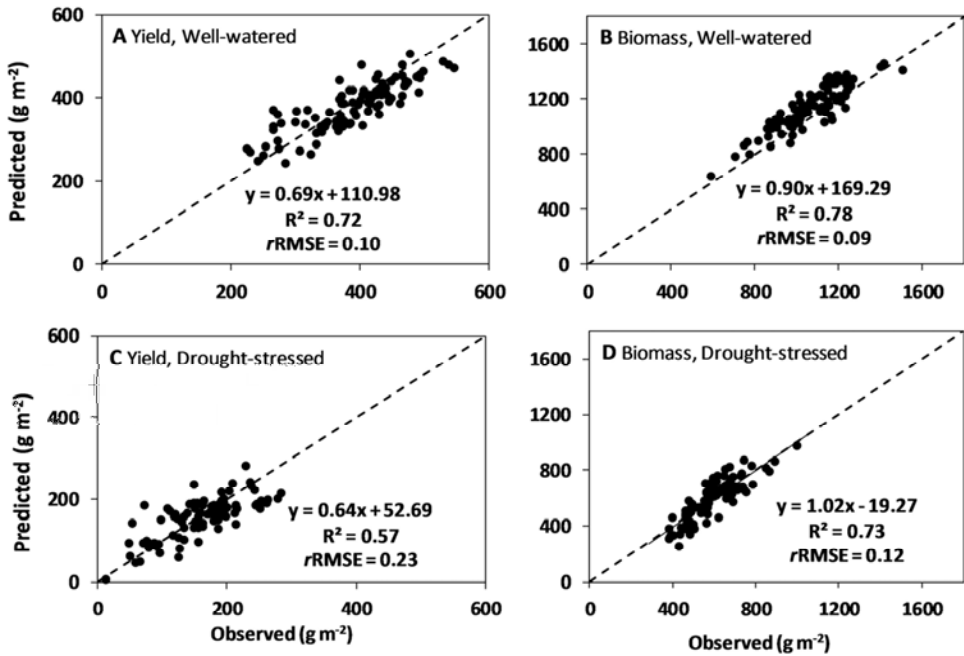
Fig. 2. Comparison between observed and simulated values of grain yield (A, C) and biomass (B, D) of rice introgression lines for both well-watered (A, B) and drought-stressed (C, D) environments. The diagonal line is the 1:1 line.

environment was used. The model accurately predicted biomass, but still over-estimated grain yield by over-estimating the number of grains per m^2 (results not shown). This could be due to the fact that the generic crop model GECROS lacks specific algorithms to account for the impact of high tissue temperature associated with reduced transpirational cooling under drought on spikelet sterility in rice. A further calibration was applied by reducing the seed number (i.e. $\sim$6171 m^2) for all ILs, based on the difference between predicted average population mean and real experimental data. After such a calibration, the model accounted for 57% of the variation in grain yield (Fig. 2C), and 73% of the variation in biomass (Fig. 2D), with rRMSE values of 0.23 and 0.12, respectively.

Both estimations of yield for well-watered and drought-stressed environments were poor compared with the best fit of linear regression in Table 2. This suggests that the input parameters required for GECROS were not all important for defining yield for this IL population, as confirmed by later analysis.
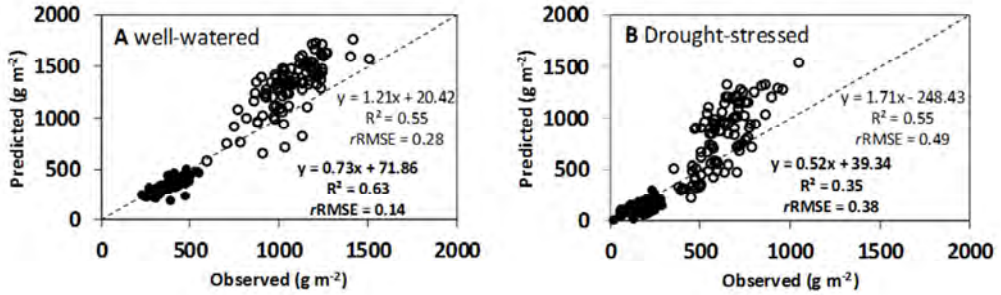
Fig. 3. Comparison between observed and predicted values by introducing photosynthesis parameters for grain yield (filled circles, bold R^2 and rRMSE) and biomass (open circles, non-bold R^2 and rRMSE), under (A) well-watered and (B) drought-stressed environments. The diagonal line is the 1:1 line.

Introducing genetic variation in leaf-level photosynthesis in the GECROS crop model

Based on representative ILs of the same population, We (Chapter 3) reported the QTL/marker effects on individual parameters of the biochemical photosynthesis module in GECROS. We introduced these QTL/marker effects into GECROS. After introducing genetic variation in leaf-level photosynthesis for each IL, the model over-estimated production, especially for biomass, and the variation accounted for by the crop model decreased significantly for both well-watered and drought-stressed conditions (Fig. 3).

Given the poor performance of the model after introducing genetic variation in leaf photosynthesis, the photosynthesis parameters for each IL were replaced by the population mean in all subsequent modelling analyses.

Coupling the effects of identified markers to crop model

First, an analysis was conducted to identify markers conferring for each model parameter. In total 20 markers were detected for all seven model parameters (Tables 3 and 4, Fig. S1). The total fraction of the phenotypic variation accounted for by the markers ranged from 27.3% to 51.7%. Marker RM410 showed multiple effects on n_{SO}, H_{max}, m_V, m_R, and S_{la}; marker RM8030 had multiple effects on S_W, H_{max}, and N_{max} under well-watered conditions; marker RM11 was related to phenology influencing both m_V and m_R; marker RM338 influenced both m_V and N_{max} in the drought-stressed environment; marker RM475 was related to N_{max} in both well-watered and drought-stressed environments.

Secondly, based on the additive effects predicted by the multiple regression analysis and allele information at each detected locus, marker-based values for each of the

Table 3. Coefficients of Eqn (1) used to identify markers conferring for seven physiological model-input parameters and for grain yield, using data from the well-watered conditions in 2009 and also from drought-stressed conditions for total nitrogen uptake. For definition and unit of these parameters, see Table 1. Marker positions were based on the SSR marker linkage map established for the rice introgression lines population (Chapter 2; see Fig. S1). R^2 denotes the percentage of phenotypic variation accounted for by all markers identified for a given parameter or trait.

Trait	μ	Chr.	Location (cM)	Markers	Additive effect (a_n)	P-value	R^2 (%)
S_W		1	124.8	RM1152	0.0010	0.0002	45.2
	0.0222	2	110.9	RM1367	0.0008	0.0007	
		2	139.3	RM8030	-0.0009	0.0035	
		4	123.8	RM2799	-0.0006	0.0039	
n_{SO}		3	79.1	RM251	0.0009	<0.0001	36.8
	0.0148	9	64.4	RM410	0.0003	0.0002	
		12	61.6	RM1261	-0.0004	0.0039	
H_{max}		1	9.5	RM8068	0.037	0.0213	51.7
		2	139.3	RM8030	-0.057	0.0004	
		4	25.5	RM518	-0.035	0.0042	
	1.174	7	43.5	RM432	0.081	<0.0001	
		9	64.4	RM410	0.042	<0.0001	
		10	87.1	RM294A	0.058	0.0021	
m_V		1	124.8	RM1152	-1.38	0.0041	33.6
	90.78	3	108.4	RM338	1.31	0.0487	
		7	47	RM11	1.61	0.0191	
		9	64.4	RM410	1.23	0.0002	
m_R		1	124.8	RM1152	0.52	0.0127	27.3
	21.71	7	47	RM11	-0.81	0.0086	
		9	64.4	RM410	-0.59	0.0001	
S_{la}		1	25.4	RM8145	0.0007	0.0006	31.0
	0.0203	7	81.05	RM3753	0.0006	0.0020	
		9	64.4	RM410	0.0003	0.0098	
N_{max} well-watered		2	92.5	RM475	-0.44	0.0082	37.0
		2	139.3	RM8030	-0.50	0.0052	
	7.83	8	83.7	RM284	-0.53	0.0004	
		9	0.8	RM5799	-0.35	0.0460	
N_{max} drought-stressed		1	98.1	RM306	0.62	0.0011	36.2
		2	92.5	RM475	-0.46	0.0066	
	5.13	3	108.4	RM338	-0.76	0.0042	
		5	20.6	RM7302	-0.47	0.0277	
Yield well-watered		2	92.5	RM475	-21.7	0.0288	56.5
		2	139.3	RM8030	-34.9	0.0009	
	325.3	8	35.7	RM4085	-14.5	0.0431	
		8	83.7	RM284	-22.4	0.0113	
		9	64.4	RM410	-19.5	0.0048	
Yield drought-stressed		3	108.4	RM338	-42.2	0.0002	45.4
		5	20.6	RM7302	-22.2	0.0067	
	137.2	5	132.7	RM538	25.7	0.0328	
		9	64.4	RM410	-18.8	0.0004	
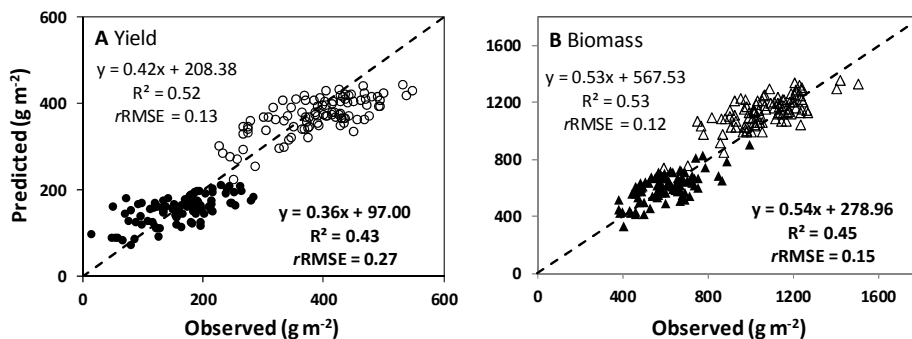
Fig. 4. Comparison between observed values and those predicted by the marker-based trait inputs for 96 rice genotypes of the IL population, for grain yield (A) and biomass (B), for well-watered (empty symbols, non-bold R^2 and rRMSE) and drought stress (filled symbols, bold R^2 and rRMSE). The diagonal line is the 1:1 line.

model parameters were estimated using Eqn (1) for each IL. The performance of GECROS with marker-based estimates of model parameters was examined (Fig. 4). The marker-based GECROS model accounted for 52% of the across-IL phenotypic variation of yield in the well-watered environment and for 43% of the across-IL phenotypic variation in the drought-stressed environment, with rRMSE of 0.13 and 0.27, respectively. These percentages were almost comparable with those percentages accounted for by the markers identified for yield *per se* (Table 3).

The GECROS model using marker-based estimates of model parameters gave less accurate predictions than using measured model parameters (Fig. 4 vs Fig. 2). In both well-watered and drought-stress cases, the marker-based model seemed to over-predict the lower end of observed yield and biomass, and to under-predict the higher end of observed yield and biomass, and, as a result, the range of predicted values was narrower than that of the observed data. This narrower range could be caused by the fact that the detected markers only explained part of the variation of model parameters (Table 3).

We directly compared the predictions of yield and biomass of the two versions of the model, the marker-based model and the model using the measured parameters (Fig. 5). The marker-based predictions correlated well with the original predictions in grain yield for both the well-watered ($r = 0.73$) and the drought-stressed ($r = 0.70$) environment. Similar correlations were obtained in biomass ($r = 0.75$ and 0.71 for the two environments, respectively). These correlations suggest that model parameters estimated using marker information can replace measured input parameters.
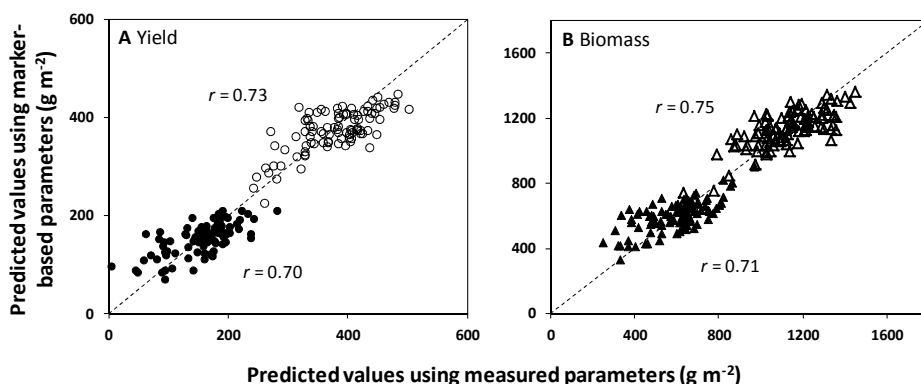
Fig. 5. Correlation between predicted values from marker-based model parameters and those from measured model parameter values for 96 rice genotypes of the IL population for grain yield (A) and biomass (B), for well-watered (empty symbols) and drought-stressed environment (filled symbols)

Extrapolating the prediction of marker-based modelling to a population of recombinant inbred lines from the same parents

For predicting yield differences within an independent population of RILs derived from the same parents, marker-based estimates of model parameters were used according to the same approach as in the IL population using Eqn (1). The model predicted 21% of the phenotypic variation under well-watered conditions and 20% of the phenotypic variation under drought-stressed conditions (Fig. 6), with rRMSE = 0.31 and 0.45, respectively.
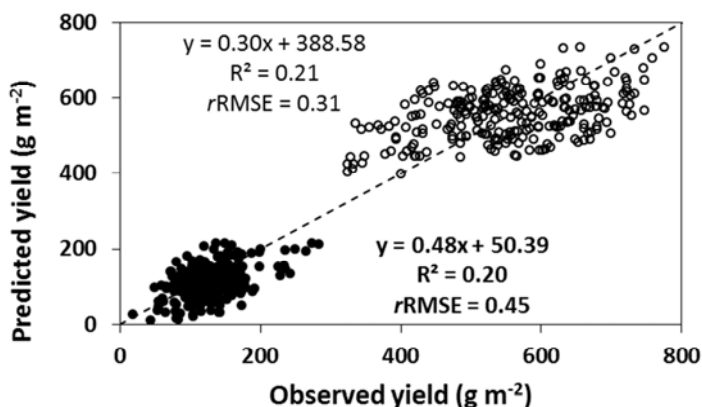
Fig. 6. Comparison between observed values of grain yield and those predicted by additive effects of marker-based parameters for a population of 251 recombinant inbred lines grown under well-watered (empty circles, non-bold R^2 and rRMSE) and drought-stressed (filled circles, bold R^2 and rRMSE) environments.

Model-based sensitivity analysis to identify important yield-defining markers

The sensitivity for a single marker was analysed by excluding the effect of the marker in estimating marker-based model parameters. Marker RM8030 on Chromosome 2 contributed most to the yield of the IL population under well-watered conditions, whereas marker RM338 on Chromosome 3 contributed most under drought-stressed conditions (Table 4). When excluding the additive effect of RM8030, the phenotypic variation accounted for by the GECROS model decreased from 51.6% to 34.2%; for RM338, the phenotypic variation accounted for decreased from 42.6% to 29.8%. The marker ranking obtained through the sensitivity analysis (Table 4) agreed well with the linear regression analysis of rice yield against model parameters (Table 2). As shown by the linear regression N_{max} contributed most to the variation in yield. In accordance with that observation, the most important yield-influencing markers identified by the sensitivity analysis were all related to N_{max} (Table 4).

Table 4 also shows that under well-watered conditions markers RM410 and RM251 related to n_{SO} had a higher ranking than RM5799 related to N_{max}; in drought-stressed conditions marker RM410 and RM432 influencing H_{max} had a higher ranking than RM306 and RM475 influencing N_{max}. Parameters n_{SO} and H_{max} also had statistically significant effects on yield for well-watered and drought-stressed conditions, respectively (Table 2). Therefore, the regression analysis supported our conclusion based on the model-based sensitivity analysis approach.

Our analysis showed that removing some markers could also have no effect on yield prediction or even increase the power of the prediction (Table 4). For example, by removing the additive effect of marker RM251, the prediction for drought-stressed conditions could improve from 42.6% to 46.2% variation accounted for.

Most high-ranking markers found in this approach were consistent with the markers identified for yield *per se* in Table 3, for example, the four highest-ranking markers in the well-watered environment (i.e. RM8030, RM284, RM475 and RM410) and the three highest-ranking markers in the drought-stressed environment (i.e. RM338, RM7302 and RM410). The model-based approach detected 20 markers contributing to yield (Table 4), more than the markers identified from multiple regression analysis for yield (Table 3).

DISCUSSION

Using crop models to predict genotypic differences in yield has constantly been a challenge for crop modellers (Yin *et al*., 2000a; Yin & Struik, 2010). This study examined the ability of the generic crop model GECROS to account for differences in yield within IL and RIL populations of rice. We also used the principles for QTL-based modelling as defined earlier (Yin *et al*., 2000b; 2005b) and integrated genetic

Table 4. Percentage of the phenotypic variation in yield among rice ILs (R^2) accounted for by the marker-based version of the GECROS model when different sets of markers were used to calculate marker-based parameter inputs for GECROS. Marker positions are denoted as 'Chr_cM', that is, Chromosome_centiMorgan, as identified in Table 3.

Chr_cM	Marker	Parameter set																			
1_9.5	RM8068	+†	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
1_25.4	RM8145	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
1_98.1	RM306	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+
1_124.8	RM1152	+	+	+	+	-	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+
2_92.5	RM475	+	+	+	+	+	-	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+
2_110.9	RM11367	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+
2_139.3	RM8030	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
3_79.1	RM251	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+
3_108.4	RM338	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+
4_25.5	RM518	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	-	+	+	+	+
4_123.8	RM2799	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+
5_20.6	RM7302	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	-	+	+	+	+	+
7_43.5	RM432	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
7_47	RM11	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+
7_81.05	RM3753	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+
8_83.7	RM284	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+
9_0.8	RM5799	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+
9_64.4	RM410	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+
10_87.1	RM294A	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	-	-	+
12_61.6	RM1261	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
Well-watered	R^2 (%)	51.6	51.6	53.9	51.6	50.9	46.2	51.7	34.2	47.9	52.6	59.4	51.8	51.6	50.7	51.7	49.4	45.7	48.3	47.3	51.1	53.9
	Marker rank		11	18	11	9	3	14	1	5	17	20	16	11	8	14	7	2	6	4	10	19
Drought-stressed	R^2 (%)	42.6	42.6	41.4	39.5	44.6	37.7	45.5	40.9	46.2	44.1	40.2	33.2	36.9	43.8	41.9	42.6	42.6	35.9	40.5	42.1	
	Marker rank		13	13	10	6	18	5	19	9	20	17	7	2	4	16	11	13	13	3	8	12

† The '+' symbol indicates that IL-specific allelic values (-1 or 1) of the marker was used in calculating marker-based inputs; the '-' symbol indicates that all ILs were assumed to carry an identical allele (i.e. 0) at the marker locus in calculating marker-based inputs.

information on molecular markers identified for individual model parameters of GECROS to predict both yield and biomass. In this discussion section we will mainly analyse the performance of the resulting marker-based GECROS model, and explore how models can help breeders design the ideotypes and support selecting the best genotypes.

Performance of the (marker-based) GECROS model

The GECROS model uses the concept of carbon-nitrogen interactions for simulating crop growth (Yin & van Laar, 2005; Yin & Struik, 2010; Yin, 2013). Unlike those of earlier crop models (Yin *et al*., 2000a), the input parameters of GECROS are mostly related to the traits that breeders usually measure (Table 1), which may facilitate the use of crop modelling in support of breeding (Yin *et al*., 2004).

Our results indicated that using as few as seven parameters (Table 1), the GECROS model showed a good potential to account for observed differences in yield among the 96 ILs including the parents (Fig. 2). More importantly, marker-based GECROS also predicted yield differences among the 96 ILs (Fig. 4).

Since the marker-based model parameters were based on the estimated genetic effects, the marker-based crop model should be able to predict the variation within any progeny from the same parents. This was shown to be the case, using independent lines of the same cross that were not included in the QTL mapping step (Reymond *et al*., 2003). Here we tested this possibility using a different population, i.e. RILs derived from the same parents. The comparatively low percentage of yield variation accounted for the RIL population (Fig. 6) could have been caused by the comparatively larger number of RILs (n = 251), which might have involved segregations that were not revealed by markers found in the smaller IL population. The limited number of markers with small additive effects only accounting for from 27.3% to 51.7% of the phenotypic variation of model-input parameters in the IL population (Table 3) could be another reason.

Despite the promising results there were also problems to overcome when applying this approach. First of all, the model performance was sensitive to nitrogen uptake, as plant nitrogen content not only affects canopy development, but also photosynthesis, and therefore, biomass and yield. Due to the complexity in modelling the transition between flooded and nonflooded soil environments (Gaydon *et al*., 2012) and the lack of information on soil-related parameters needed to simulate nitrogen uptake, N_{max} was directly used as an input parameter of GECROS in this study (Table 1).

Secondly, the drought treatment changed the sink-source relationships which required adjustments to be made. In our simulation, model parameters were first estimated from the well-watered experiment (2009, Beijing). However, a further

calibration was found necessary for simulating spikelet numbers, when the model was applied to the drought-stress environment (see the Results). Drought environments reduce transpirational cooling, leading to high tissue organ temperature and high spikelet sterility in rice; this effect can be highly genotype-dependent (Jagadish *et al.*, 2007). The generic model GECROS relates potential seed number to carbon and nitrogen accumulation in the vegetative phase, and does not have algorithms to account for this effect of organ temperature (Yin & van Laar, 2005). Obviously, there is a need to better account for the final spikelet number of rice when applying the model for predictions under stress.

Role of crop models in designing ideal plant types
Identifying the most determinant yield-defining traits may help to design the 'new plant type' (Peng *et al.*, 2008). Crop modelling can dissect complex traits into physiological components. Using the crop model GECROS, yield was connected to, and dissected into seven model parameters. By dissecting complex traits into physiologically meaningful components, it is possible to assess genetic variation for each component and evaluate its relative importance by sensitive analyses or regression analyses. Regression analyses showed that N_{max} had the most significant effect on yield (Table 2). This is in line with the result that the important yield-determining markers identified by GECROS-based sensitivity analysis (Table 4) were mainly those for N_{max} (Table 3). Similarly, Prudent *et al.* (2011) combining a fruit sugar model and QTL analysis, identified key elementary processes and genetic factors underlying tomato fruit sugar concentration. All these results show that the dissection approach based on physiological models can point where the QTLs for complex traits come about (Yin *et al.*, 2002), thereby revealing biological insights into complex traits. At the same time this dissection approach suggests how to create the best combination of component traits for an ideal plant type that will perform best under given conditions.

Marker-based GECROS also indicates markers that had a least impact on grain yield (Table 4). Such an analysis may suggest whether or not the model has incorporated right parameters in predicting yield differences among genotypes in a certain environment. Removing some markers even increase the power of the prediction, as was the case of marker RM251 when using the model to drought-stressed conditions (Table 4). This could have been caused by the fact that marker RM251 only influenced the traits (i.e. n_{SO}), which did not have a significant effect on yield of our IL population under drought-stressed conditions.

The model-based dissection approach should not be considered to replace, but only to complement the yield *per se* approach, as the latter identified markers (e.g. RM4085

for well-watered, and RM538 for drought-stress environment, Table 3; Fig. S1) that were not detected by the model-based sensitivity analysis. This arises from the possibility that markers under statistical threshold of detection for component traits can be detected when the aggregated complex trait itself was analysed (Yin *et al*., 2002). The other possibility is that some yield-influencing traits were not incorporated into the GECROS model as input parameters.

Marker-based modelling to improve the efficiency of marker-assisted selection
Combined with conventional breeding approaches, MAS has been used to integrate major genes or QTLs with large effect into widely grown varieties (Jena & Mackill, 2008). But so far MAS only had a moderate impact on breeding for complex traits, for which many small effect genes are involved and highly environment-dependent (e.g. yield, drought tolerance) (Collard & Mackill, 2008). One of the advantages of QTL-based models is that they can be used to evaluate the contribution of a single QTL to yield (Chenu *et al.*, 2009). This could potentially assist in finding the most important markers for MAS.

We showed that the existing GECROS model can be a useful tool to enhance the efficiency of MAS especially for complex traits (i.e. yield). The markers were first identified for various yield-determining physiological traits that are input parameters of GECROS (Table 3). The relative importance of these markers were then ranked by performing marker-based model sensitivity analysis (Table 4). Such an analysis detected markers that breeders can prioritize in their MAS programmes for specific environments. Compared with identification of markers/QTLs for yield *per se*, the model-based approach provided breeders with more information for MAS. This analysis confirms the assertion that rather than looking only for QTLs for a complex trait (yield) itself, determining QTLs for underlying component traits will provide more genetic information (Yin *et al.,* 2002; Tardieu & Tuberosa, 2010; Prudent *et al*., 2011). Notably, the GECROS model-based approach identified some markers that were otherwise unidentified by analysis of yield *per se* (e.g. marker RM432 for drought-stress environment) (Table 3). This approach provided breeders with more choice of markers for selection. It remains to be tested through actual breeding whether this additional information does indeed result in better genotypes.

QTL/marker-based modelling combined with sensitivity analysis (Table 4) can also directly evaluate a single QTL/marker's effect on yield level, which could be used to evaluate specific genotype *in silico*, thus potentially reducing labour extensive selection in the field. Since crop modelling quantifies causality between relevant physiological processes and responses of these processes to environmental variables, we could resolve the commonly observed G × E interaction (Yin *et al*., 2004; Hammer

et al., 2006; Tardieu, 2011), even the QTL × environment interaction (Hemamalini *et al.*, 2000; Asins, 2002). For example, Chenu *et al.* (2009) using the crop model APSIM-Maize simulated that a QTL accelerating leaf elongation will increase yield in an environment with water deficit before flowering, but reduced yield under terminal drought stress. Our model analysis showed that the marker RM338 contributed the greatest to yield under stressed environment, but had no effect at all under well-watered environment (Table 4). This modelling analysis will greatly improve the MAS efficiency for traits which greatly influenced by environment factors.

Leaf photosynthesis in relation to crop biomass and yield
The above discussions were based on our results without considering the genetic variation in leaf photosynthesis. As a follow-up of our previous analyses (Chapter 2 and 3) that reported significant genetic variation in photosynthesis in our IL population, we introduced genetic variation in leaf photosynthesis into the crop model to examine whether this would improve the predicting power of GECROS for different genotypes.

Photosynthesis, being the source of organic carbon, is expected to be correlated with yield, as evidenced by Fischer & Edmeades (2010). Our analysis gave the results opposite to this expectation (Fig. 3). This is seemingly in line with the result that introducing leaf nitrogen content led to poorer prediction of the variation in grain yield among barley RILs (Yin *et al.*, 2000a) and with the statement that increasing leaf photosynthesis is not a useful strategy to increase crop yield (e.g. Richards, 2000). Our result could be caused by the fact that in our IL population the photosynthesis-increasing allele of the major QTL stemmed from the lower yielding parent 'Haogelao' (Chapter 2) – the parent that had lower yield in both well-watered and drought environments (Fig. 1). Furthermore, the genetic variation in parameters of the biochemical steady-state leaf photosynthesis model was only studied at flowering and at grain filling in the greenhouse (Chapter 3), which might be different from the acclimated real-life differences between leaves in the dynamic conditions in the field (Archontoulis *et al.*, 2012). More extensive measurements and increased temporal resolution may be required.

Nevertheless, our results reflect the complex hierarchy from leaf-level photosynthesis to crop yield. Our ILs differ genetically in many respects other than photosynthesis (Fig. 1), and the variation in photosynthesis may only play a comparatively small role in this population. A simulation study focusing on the impact of natural genetic variation in leaf photosynthesis on crop productivity would be needed to exclude potential confounding effects due to variation in other physiological processes.

CONCLUSIONS

Our results are promising for applying marker-based modelling in support of breeding: the approach integrates genetic information for model parameters to predict a complex yield trait across different environments and genetic make-ups. The approach can prioritize markers in the MAS programmes for specific environments. Compared with identification of markers for yield *per se*, our analysis provided breeders with more information for MAS, although the approach should be considered to be complementary to the analysis-of-yield-*per-se*. Further improvement could be achieved by upgrading crop models for rice especially when grown under drought stress. It is also necessary to try more generations of introgression to break the putatively tight genetic linkage between high photosynthesis and low yield) and use a higher-density linkage map and a larger population size to improve the genetic resolution by increasing the power of the analyses.

Supplementary Materials in Chapter 4

Description of crop growth model GECROS

GECROS is a generic crop model operating in daily time steps (with phenology and photosynthesis related processes simulated in shorter time steps). The model simulates crop growth and development over time and generates phenotypes for many different traits. For a detailed description of its first version, see Yin & van Laar (2005). The latest version of the model (v3.0) was described as the Supplementary material of Yin (2013). Here, we only describe key features of this latest version that are related to modelled processes relevant to this paper.

Instantaneous leaf photosynthesis (A) was calculated from the analytical algorithms that are based on the model of Farquhar *et al.* (1980), coupled with a phenomenological CO_2-diffusion conductance model (for overview, see Yin & Struik, 2009 and references therein; Yin *et al.*, 2009). The analytical cubic polynomials simultaneously solve stomatal conductance (g_s), internal [CO_2], and leaf photosynthesis rate (A). The obtained g_s was used in the Penman-Monteith equation (Monteith, 1973) for surface energy balance to model leaf transpiration and leaf temperature. Leaf temperature was then used for re-calculating leaf photosynthesis and transpiration. The effects of leaf nitrogen (N) content on photosynthesis, g_s and transpiration are reflected by the effects of leaf N on parameters of the photosynthesis model. Furthermore, an option is allowed for mesophyll conductance (g_m) to vary in proportion with g_s in response to all environmental factors, given recent reports that g_m may resemble g_s in response to various environmental variables (e.g. Flexas *et al.*, 2008; Yin *et al.*, 2009; Gu *et al.*, 2012).

Spatial extension from leaf to canopy photosynthesis and transpiration was established using the sun/shade model of de Pury & Farquhar (1997). Temporal extension from instantaneous rates to daily total was performed using the five-point Gaussian integration (Goudriaan, 1986) to account for (a)symmetric diurnal course of radiation and temperature. These approaches for spatial and temporal extensions apply to the case in the absence of drought stress.

In the presence of drought stress, the available water is partitioned between sunlit and shaded leaves according to the relative share of their potential transpiration to obtain their instantaneous actual transpiration. The actual transpiration is transformed into the actual level of g_s using the Penman-Monteith equation, and the actual g_s was then used as input to an analytical quadratic model, to estimate the instantaneous actual photosynthesis of the sunlit and shaded leaves. The Gaussian integration is again used to obtain the daily total of the actual photosynthesis.

Crop respiration was modelled, based on the framework of Cannell & Thornley (2000) that recognises individual relationships between respiration and each process it supports.

GECROS uses two equations of Yin & Schapendonk (2004) for simulating the partitioning of C and N, respectively, between shoots and roots. They were based on the classical root-shoot functional balance theory, with an incorporation of the mechanism that plants control root-shoot partitioning in order to maximise their relative C gain.

The intra-shoot nitrogen partitioning is based on a pre-defined maximum grain N concentration of a genotype and a minimum N concentration in the stems. If the N requirements for the grains and stems are met from the current N uptake, the remaining shoot nitrogen goes to the photosynthetically active plant parts (including leaf blades, leaf sheaths, photosynthetically active parts of the stems and ears), whose surface area determines the green-surface area index (GAI). If the requirements for the grains are not met, remobilisation of N takes place, first from the reserves and then from the leaves and the roots, until the reserves are depleted and N concentrations in the leaves and roots reach their minimum values. This remobilisation advances leaf and root senescence.
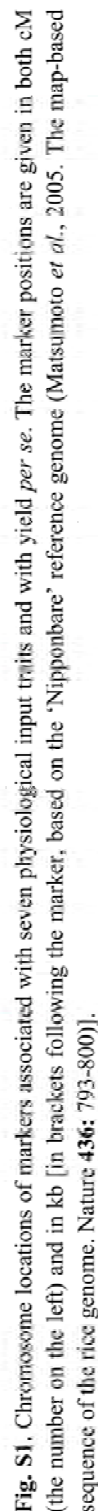
Maximum stem weight of the crop is assumed to be proportional to maximum plant height, whilst maximum single grain weight is set as genotypic parameter. Potential grain number per m^2 is co-determined by carbon (C) and N accumulation during vegetative growth. Daily demand for C by stems and grains is simulated using the differential form of an equation for describing any asymmetric sigmoid pattern of a determinate growth (Yin *et al.*, 2003). The remaining shoot-carbon goes either to the leaves or to the C reserve pool in the stems, depending on whether GAI becomes limited by nitrogen. The GAI can be either C or N limited; it is calculated following the principles described by Yin *et al.* (2000). If C reserves are present, C is made available to the grains, when current photosynthesis does not satisfy their C demand.

For simulating phenological development, development stage is defined as 0 at seedling emergence, 1 at start of grain filling and 2 at physiological grain maturity. The intervals from stage 0 to 1 and from 1 to 2 depend on the genotype-specific number of days at optimum temperature. A flexible bell-shaped non-linear function (Yin *et al.*, 1995) is used to describe the temperature response of development rate. This rate has a value of 0 when the hourly temperature is below the base temperature or above the ceiling temperature; it is 1 when it is equal to the optimum temperature. Development rate is also affected by daylength during the photoperiod sensitive part of the vegetative phase (but the daylength effect was not used for the simulations in the current study).

References in Supplementary Materials:

Cannell MGR, Thornley JHM. 2000. Modelling the components of plant respiration: Some guiding principles. *Annals of Botany* **85**, 45-54.

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

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.

Flexas J, Ribas-Carbó M, Diaz-Espejo A, Galmes J, Medrano H. 2008. Mesophyll conductance to CO_2: current knowledge and future prospects. *Plant, Cell and Environment* **31**, 602-621.

Goudriaan J. 1986. A simple and fast numerical method for the computation of daily totals of crop photosynthesis. *Agricultural and Forest Meteorology* **38**, 249-254.

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

Monteith JL. 1973. Principles of environmental physics. Edward Arnold, London, 241pp.

Yin X. 2013. Improving ecophysiological simulation models to predict the impact of elevated atmospheric CO_2 concentration on crop productivity. *Annals of Botany*, in press (DOI: 10.1093/aob/mct016).

Yin X, Kropff MJ, McLaren G, Visperas RM. 1995. A nonlinear model for crop development as a function of temperature. *Agricultural and Forest Meteorology* **77**, 1-16.

Yin X, Schapendonk AHCM, Kropff MJ, van Oijen M, Bindraban PS. 2000. A generic equation for nitrogen-limited leaf area index and its application in crop growth models for predicting leaf senescence. *Annals of Botany* **85**, 579-585.

Yin X, Goudriaan J, Lantinga EA, Vos J, Spiertz HJ. 2003. A flexible sigmoid function of determinate growth. *Annals of Botany* **91**, 361-371.

Yin X, Schapendonk AHCM. 2004. Simulating the partitioning of biomass and nitrogen between root and shoot in crop and grass plants. *NJAS-Wageningen Journal of Life Sciences* **51**, 407-426.

Yin X, van Laar HH. 2005. Crop systems dynamics: an ecophysiological simulation model for genotype-by-environment interactions. Wagningen Academic Publishers, Wageningen, The Netherlands.

Yin X, Struik PC. 2009. C_3 and C_4 photosynthesis models: An overview from the perspective of crop modelling. *NJAS-Wageningen Journal of Life Science* **57**, 27-38.

Yin X, Struik PC, Romero P, Harbinson J, Evers JB, van der Putten PEL, Vos J. 2009. Using combined measurements of gas exchange and chlorophyll fluorescence to estimate parameters of a biochemical C_3 photosynthesis model: a critical appraisal and a new integrated approach applied to leaves in a wheat (*Triticum aestivum*) canopy. *Plant, Cell & Environment* **32**, 448-464

Fig. S1. Chromosome locations of markers associated with seven physiological input traits and with yield *per se*. The marker positions are given in both cM (the number on the left) and in kb [in brackets following the marker, based on the 'Nipponbare' reference genome (Matsumoto *et al.*, 2005. The map-based sequence of the rice genome. Nature **436**: 793-800)].

Can exploiting natural genetic variation in leaf photosynthesis contribute to increasing rice productivity? A simulation analysis

Junfei Gu, Xinyou Yin, Tjeerd-Jan Stomph, Paul C. Struik

Centre for Crop Systems Analysis, Department of Plant Science, Wageningen University, P. O. Box 430, 6700 AK Wageningen, The Netherlands

ABSTRACT

Rice productivity can be limited by available photosynthetic assimilates from leaves. However, the lack of significant correlation between crop yield and leaf photosynthetic rate (A) is noted frequently. The effects of engineering for an improved photosynthesis have been reported to damp down gradually when moving up from leaf to crop level, because of complicated constraints and feedback mechanisms. Here we examined the extent to which natural genetic variation in A can contribute to increasing rice productivity. Using the mechanistic model GECROS we analysed the impact of genetic variation in A on crop biomass production, based on the quantitative trait loci for various photosynthetic components within a rice introgression-line population. We showed that genetic variation in A of 25% can be scaled up equally to crop level, resulting in an increase in biomass of 22-29% across different locations and years. This was probably because the genetic variation in A resulted not only from Rubisco-limited photosynthesis but also from electron transport-limited photosynthesis; as a result, photosynthetic rates could be improved for both light-saturated and light-limited leaves in the canopy. Rice production could be significantly improved by mining the natural variation in existing germplasm, especially the variation for parameters that determine light-limited photosynthesis.

Key words: Canopy photosynthesis, crop model, GECROS, genetic variation, *Oryza sativa* L., photosynthesis, rice.

INTRODUCTION

Cereal yield is determined by the accumulated photosynthetic assimilates over the entire growing season that are partitioned into the caryopses. Improvements in crop management and genetic gain in harvest index are largely responsible for the increased cereal yields over the last decades (Austin, 1999; Peng *et al.*, 2008). However, it has been argued that cereal production is now approaching a plateau and further increases in yield will necessitate an increase in photosynthesis (Austin, 1994; Mitchell & Sheehy, 2006; Lawson *et al.*, 2012).

Crop photosynthesis accumulated for the entire growing season depends on the ability of the crop to build up and maintain a canopy for capturing incoming light, but also on the photosynthetic capacity and efficiency of leaves. There may be chances to increase the light capture by improving early leaf area growth rate or by introducing 'stay green' genes to extend the growing season (Long *et al.*, 2006). For rice, however, leaf area dynamics and canopy architecture may have been effectively optimized for maximum light capture through breeding (Horton, 2000). Any further increase in photosynthesis of the rice crop may have to come from improved leaf photosynthesis.

Photosynthesis per unit leaf area seems to have been improved already as suggested by experimental comparisons of old and modern varieties of cereals, including rice, in concert with improvements in harvest index and grain number (Fischer & Edmeades, 2010).

Given the fast developing biotechnology, opportunities for improving leaf-level photosynthesis via genetic engineering have been extensively explored, by either experimental approaches or theoretical computation. Approaches include, for example, designing more efficient Rubisco (Mueller-Cajar & Whitney, 2008; Whitney & Sharwood, 2008); exploiting existing inter-specific variation in Rubisco efficiency (Zhu *et al.*, 2004a); increasing RuBP regeneration and light reaction (Miyagawa *et al.*, 2001; Peterhansel *et al.*, 2008; Rott *et al.*, 2011); increasing mesophyll conductance (Uehlein *et al.*, 2008); introducing CO_2-concentrating mechanism into C_3 crops (Price *et al.*, 2008); introducing CO_2-concentrating mechanism with Kranz anatomy into C_3 crops (von Caemmerer *et al.*, 2012); short-circuiting photorespiration (Maurino & Peterhansel, 2010); and increasing the rate of transition from photoprotection (Zhu *et al.*, 2004b). Long *et al.* (2006) estimated that these ambitious approaches, if successful, would need research efforts of 10-30 years, depending on the avenues to be used.

Leaf photosynthesis could be improved not only through transgenic biotechnology, but also through the exploitation of natural variation with a conventional breeding approach. Parry *et al.* (2011) indicated that mining existing genetic variation could be the most efficient method for short term improvements (< 5 years). Recently, quantitative trait loci (QTLs) related to different photosynthetic parameters have been successfully mapped (Takai *et al.*, 2009; Adachi *et al.*, 2011; Chapter 2). Furthermore, in Chapter 3, we using the biochemical photosynthesis model of Farquhar *et al.* (1980) as adapted by Yin *et al.* (2009b), successfully dissected genetic variation of leaf photosynthesis present in an introgression line (IL) population into different biophysical and biochemical components. Their analysis showed that by using genetic variation in all components, leaf-level photosynthesis could potentially be increased by ca 20% through marker assisted selection.

However, photosynthesis rate per unit area of leaf does not correlate well with biomass produced (Evans & Dunstone, 1970; Teng *et al.*, 2004). This has led to a common notion that increasing leaf photosynthesis is not a useful strategy to increase crop yield (Richards, 2000; Zhao *et al.*, 2008). Actually, this notion was confirmed by our own work on the IL population: among the many physiological parameters examined, leaf photosynthesis was not important in determining the differences in crop yield among the ILs observed in a field experiment, either under drought or under well-watered conditions (Chapter 4). This lack of persistence of variation across scales

is probably due to the complex hierarchy from leaf-level photosynthesis to crop yield and to interaction and feedback mechanisms occurring between physiological components within the individual plant, between plants of the same crop and between the crop and the environment. Moreover, leaf photosynthesis and crop yields are both associated with some random experimental error. Together, the complexities and the experimental noise might mask the potential contribution of the small within-population variation in leaf photosynthesis to the variation in final crop yield. Therefore, modelling has been a useful tool to investigate the potential of improved photosynthesis on crop productivity (Day & Chalabi, 1988; Long *et al.,* 2006).

In this paper we use the process-based crop model GECROS (Yin & van Laar, 2005) to examine the extent to which exploiting the natural genetic variation in leaf photosynthesis components can contribute to variation in canopy photosynthesis and in crop yield in rice. The GECROS model combines sufficient physiological rigour for complex phenotypic responses with genotype-specific parameters. We use this model to scale up variation in leaf photosynthesis components as detected in our previous study (Chapter 3) to variation in canopy photosynthesis and in biomass productivity across the entire growing season for contrasting environments. Input parameter values for model simulation are only those derived from our previous results on quantitative trait loci (QTLs) for various leaf photosynthesis parameters, while other input parameters of the GECROS model are maintained the same across rice genotypes. In this way, potential confounding effects due to variation in other physiological processes can be avoided to exclusively illustrate the potential impact of natural genetic variation in leaf photosynthesis on crop productivity. We specifically hypothesise for potential larger persistence during scaling up than observed in previous studies from the literature if photosynthesis can be improved irrespective of light level, and test this hypothesis using the IL population segregating for QTLs related to both light-saturated and light-limited photosynthesis parameters.

MATERIALS AND METHODS
Based on the genetic variation found in an IL population of rice (Chapter 2 and 3), the crop model GECROS (Yin & van Laar, 2005) was used to evaluate the expression of genetic variation in leaf photosynthesis in terms of variation of canopy photosynthesis and crop biomass production.

Crop growth model GECROS
GECROS is a generic model that operates in daily time steps, simulates the growth and development of the crop over time, and generates phenotypes for a multitude of traits, based on concepts of the balance, interaction and feedback mechanisms among

various contrasting components of crop growth. A detailed description of GECROS and its algorithms can be found in Yin & van Laar (2005). In this section, only key features, related to carbon partitioning, nitrogen (N) demand and partitioning, phenological development, and photosynthesis are described; the corresponding algorithms are given in Appendix 1.

Nitrogen demand is the maximum of the deficiency driven and the growth-activity driven demand (Eqn 1-6). The deficiency driven demand is the amount of nitrogen required to restore the critical minimum nitrogen concentration. The growth-activity driven demand is based on the optimum nitrogen/carbon ratio for the maximum relative carbon gain.

Root-shoot partitioning for nitrogen and carbon responds to environmental factors, based on the root-shoot functional balance theory (Charles-Edwards, 1976). The intra-shoot nitrogen partitioning (Eqn 7) is based on a pre-defined maximum grain nitrogen concentration of a genotype and a minimum nitrogen concentration in the stems. If the nitrogen requirements for the grains and stems are met from the current nitrogen uptake, the remaining shoot nitrogen goes to the leaves, which include the photosynthetically active parts of the stems, sheaths and ears. If the requirements for the grains are not met, remobilisation of nitrogen first from the reserves and then from the leaves and the roots takes place, until the reserves are depleted and the nitrogen concentrations in the leaves and roots reach their minimum values. This remobilisation stimulates leaf and root senescence. If the grain nitrogen requirements are not met by shoot nitrogen and remobilisation, the grain nitrogen concentration declines. Intra-shoot carbon partitioning to the stems (including sheaths) and to the grains are determined according to their expected daily carbon demands, which are described by the differential form of a sigmoid function for asymmetric determinate growth (Yin *et al.*, 2003b). The remaining shoot-carbon goes either to the leaves, or to the carbon reserve pool in the stems when the green-surface area index (GAI) becomes nitrogen limited. The GAI is calculated according to the principles described by Yin *et al.* (2000c), as either the carbon or the nitrogen limited GAI. The carbon reserves, if any, become available to the grains, when current photosynthesis does not satisfy the carbon demand by grains.

In GECROS, phenological development is calculated by Eqn 8-10. Development stages are defined as 0 at seedling emergence, 1 at start of grain filling and 2 at physiological grain maturity. The intervals from stage 0 to 1 and from 1 to 2 depend on the genotype-specific number of days at optimum temperature. A flexible bell-shaped non-linear function (Yin *et al.*, 1995) is used to describe the temperature response of development rate, which has a value of zero when the hourly temperature is below the base temperature or above the ceiling temperature and one when it is

equal to the optimum temperature. In case of photoperiod-sensitive genotypes, development rate is also affected by daylength during the photoperiod sensitive phase of the vegetative interval.

Canopy photosynthesis sub-model of GECROS

To compute canopy photosynthesis as driver of crop growth, GECROS uses the two-leaf approach that divides the canopy into sunlit and shaded fractions, based on solar height; each fraction is modelled separately with a single-layer leaf model (Eqn 11-16; de Pury & Farquhar, 1997; Wang & Leuning, 1998). For both canopy fractions, the photosynthetically active nitrogen is calculated using a base value of leaf nitrogen (below which photosynthesis is zero) and a leaf nitrogen extinction coefficient to describe an exponential profile in the canopy for vertical decline in nitrogen (Yin *et al.*, 2000c). For example, photosynthetically active nitrogen for the entire canopy (N_c), for the sunlit leaf fraction of the canopy ($N_{c,su}$) and for the shaded leaf fraction of the canopy ($N_{c,sh}$), can be estimated by Eqn 17-19. To estimate the photosynthesis parameters for the entire canopy, we introduced the nitrogen dependency through a linear function (Eqn 20) (Harley *et al.*, 1992b). The photosynthetic rate of each canopy fraction was then computed using a leaf model as described below. The canopy photosynthesis model was also decoupled from GECROS in order to simulate the variation among the ILs in canopy photosynthesis alone, without the crop growth feedback loops.

Leaf photosynthesis sub-model

In GECROS, prediction of the rate of photosynthesis at leaf level is based on the models of Farquhar *et al.* (1980) as modified by Yin *et al.* (2009b) (Eqns 21-24). A phenomenological model of the Leuning type (Leuning 1995), Eqn 25, was introduced (Yin & Struik, 2009a) for quantifying stomatal conductance g_s. A similar equation, Eqn 26, was used to describe mesophyll conductance g_m (Yin *et al.*, 2009b). Parameters δ_m and δ_s in Eqns 25 and 26 can be used to estimate the g_m:g_s ratio (Chapter 3). Eqns 27 and 28 were used to predict the effects of vapour pressure difference on the conductances. Leaf temperature, which affects the rates of most biochemical reactions of photosynthesis (Eqns 29-30), is also predicted in GECROS by coupling the g_s and leaf photosynthesis models (Yin & Struik 2009a) with the Penman-Monteith equation. The kinetic constants (K_{mc} and K_{mo}) of Rubisco required for leaf photosynthesis were taken from Bernacchi *et al.* (2002), and Rubisco specificity ($S_{c/o}$) was set at 3.02 mbar μbar^{-1} (Chapter 3); and both were assumed to be conservative across all the ILs. In GECROS, day respiration (R_d) was assumed to be scaled with the maximum Rubisco activity (V_{cmax}). Genetic variation in leaf

photosynthesis parameters like V_{cmax} was based on results from Chapter 3 (also see below).

Genetic input parameters for leaf photosynthesis of introgression lines
The above leaf model contains six photosynthesis parameters κ_{2LL}, J_{max}, θ, δ_m, δ_s and V_{cmax}. Genome regions or QTLs were previously assigned for these parameters, based on 11 representative lines of the IL population (see Fig. 1 in Chapter 3). The additive effects of the QTLs estimated therein were used to estimate genotype-specific values of the six parameters for individual lines of the IL population, based on the marker allelic information for these ILs. The QTL model takes into account both the origin of the allele at a detected locus (conditional on the genotype score of the nearest marker) and the effect of the alleles at this locus on the parameter value. A parameter value X of introgression line k, containing N QTLs, was presented as:

$$X_k = \mu + \sum_{n=1}^{N} a_n \times M_{k,n}$$

where μ = the intercept; a_n = the additive effect of the n-th QTL; $M_{k,n}$ = genetic score of the n-th QTL of the individual introgression line k that takes either the value -1 (allele coming from recurrent parent 'Shennong265') or 1 (donor parent 'Haogelao' allele present). All the calculations were based on the results from Chapter 3. Note that while the population contains 96 ILs, only 38 genotypes were identified based on the allelic information of the QTLs for the six photosynthesis parameters.

Other input parameter values and model simulation
All other parameter values were calibrated for rice in Chapter 4 and were used here across all individual ILs. Simulation was performed for the conditions from 13th May 2008 and 14th May 2009 onwards at Shangzhuang Experimental Station (39°54'N, 116°24'E; elevation of 50 m above sea level) of China Agricultural University, in Beijing, North China (coded as BJ08 and BJ09, respectively). To test the genetic variation in response to different climate conditions, simulations were also carried out for the dry season (from 10th Jan) in Los Baños (14°11'N, 121°15'E; elevation of 21 m above sea level), International Rice Research Institute, Philippines from 2001 to 2005 (PH01, PH02, …, PH05, respectively). The time course of weather variables in these environments are shown in Supplementary Fig. S1.

Data analysis
In order to quantitatively evaluate the genetic variation in the IL population, genetic variation was calculated as $((X_{max} - X_{min}) / \overline{X}) \times 100$ (%) where X_{max} and X_{min} stands for maximum and minimum value, respectively, and $\overline{X}$ stands for the mean, of all the lines
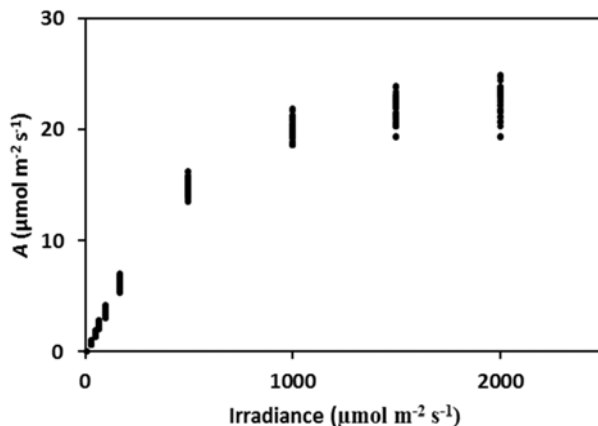
Fig. 1. The modelled light response curves of net photosynthesis rate (*A*) to light intensity at ambient CO_2 concentration (380 μmol mol^{-1}) and a leaf temperature of 25℃ in a population of 38 introgression lines.

in the population. Note that although the simulations were conducted only for the 38 genotypes, the population mean of the simulated output traits was still calculated on the basis of 96 ILs, i.e. the weighted mean given the number of IL repeats in each genotype. For each individual introgression line, the genetic gain or loss was calculated as $((X_i - \overline{X})/\overline{X}) \times 100$ (%), where X_i stands for the output parameter value for the *i*-th IL. The correlations and multiple analyses were calculated by PROC CORR, PROC GLM, respectively in SAS 9.2 (SAS Inst. Inc.).

RESULTS AND DISCUSSION

Genetic variation in leaf photosynthesis

For this paper we were only interested in the genetic variation in leaf photosynthesis, so the 96 ILs (including the two parents) were divided into 38 unique genotypes, based on origin of the alleles at the seven loci of leaf photosynthesis parameters reported in Chapter 3. Light response curves of leaf photosynthesis for the IL population were constructed based on detected QTLs for each of the six parameters (Fig. 1). Associated with QTLs detected for photosynthetic efficiency under limiting light (κ_{2LL}, θ), diffusional conductance (δ_m, δ_s), maximal rate of electron transport (J_{max}), and maximum rate of Rubisco activity (V_{cmax}), genetic variation in leaf level photosynthesis at various light intensities was considerable. At low light (100 μmol m^{-2} s^{-1}) level, genetic variation amounted to 31.4%, whereas it was 18.7% at intermediate light (500 μmol m^{-2} s^{-1}) level, and 25.4% at saturated light (2000 μmol m^{-2} s^{-1}) level (Fig. 1).
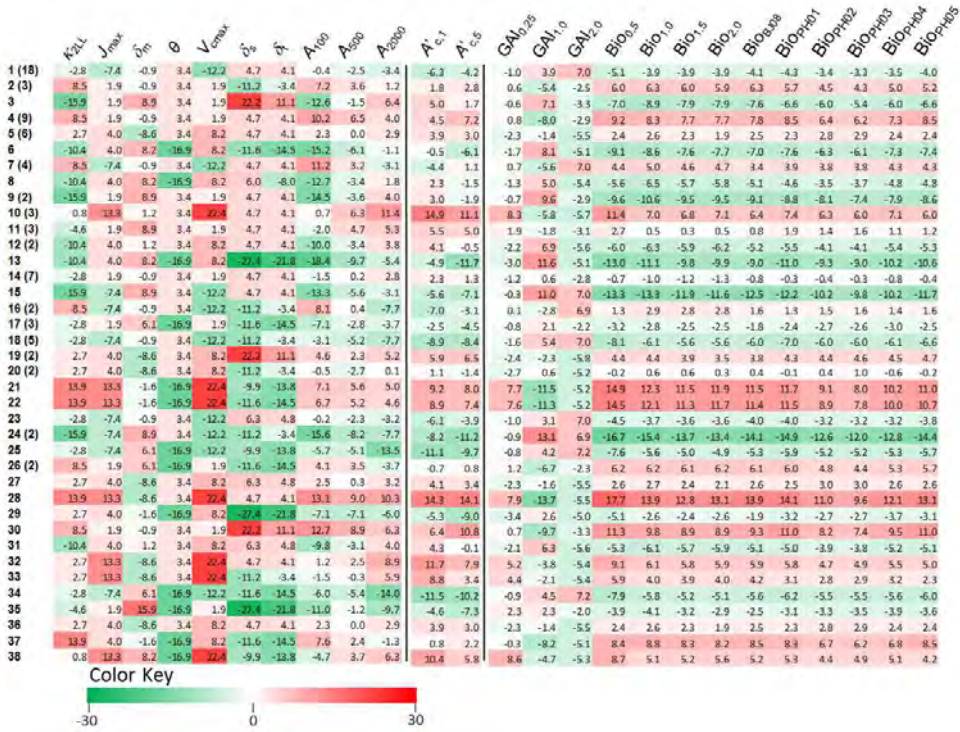
Fig. 2. Heat map of genetic gain or loss (%) of 38 genotypes at leaf, canopy and crop level, from left to right as separated by the vertical black lines. Each row represents a unique genotype based on alleles at detected loci (Gu *et al.*, 2012b), each column represents a model parameter. The first column shows in total 38 different genotypes out of 96 introgression lines with number of repeats in brackets. The values increase from green via white to red. The genetic gain or loss for each introgression line (%) was calculated as $(X_i - \overline{X})/\overline{X}$, where X_i stands for the parameter value for *i*-th genotype, and $\overline{X}$ stands for the population mean. In this figure, the vertical direction shows the genetic gain or loss for all genotypes at the leaf, canopy, and crop level; the horizontal direction shows how each line performed with regards to the photosynthetic components at leaf level, the canopy photosynthesis at low or high GAI, the GAI dynamics, the biomass dynamics, and the harvested biomass at different locations and years. κ_{2LL}, value of conversion efficiency of incident light into e⁻ transport at the strictly limiting light; J_{max}, maximum value of e⁻ transport under saturated light; θ, convexity factor for response of e⁻ transport to irradiance; δ_m, δ_s, and δ_t, parameters defines chloroplast /intercellular [CO$_2$] ratio, intercellular /ambient [CO$_2$] ratio, and chloroplast /ambient [CO$_2$] ratio at saturating light, respectively; V_{cmax}, maximum rate of Rubisco activity-limited carboxylation; A_{100}, A_{500}, and A_{2000}, leaf photosynthesis at low light (100 μmol m⁻² s⁻¹), intermediate light (500 μmol m⁻² s⁻¹), and saturated light (2000 μmol m⁻² s⁻¹), respectively; $A'_{c,1}$ and $A'_{c,5}$, daily canopy photosynthetic gain at high light level for green-surface area index (GAI) = 1 and GAI = 5, respectively; GAI$_{0.25}$, GAI$_{1.0}$ and GAI$_{2.0}$, value of GAI at development stage 0.25 (seedling), 1.0 (flowering), 2.0 (harvest), respectively; Bio$_{0.5}$, Bio$_{1.0}$, Bio$_{1.5}$ and Bio$_{2.0}$, total biomass at development stage 0.5 (seedling), 1.0 (flowering), 1.5 (grain filling), and 2.0 (harvest) for Beijing in 2009, respectively; Bio$_{BJ08}$, Bio$_{PH01}$, Bio$_{PH02}$, Bio$_{PH03}$, Bio$_{PH04}$ and Bio$_{PH05}$, total harvest biomass at location Beijing year 2008, and location Philippines year 2001 to 2005, respectively.

Estimated genetic variation in canopy photosynthesis

Canopy photosynthesis (A'_c) was spatially integrated from base to the top of the canopy. Some parameters (e.g. V_{cmax}, J_{max}, g_s, and g_m) were adjusted to change with depth in the canopy, based on the modelled profile of leaf nitrogen content in the canopy. Spatially integrated canopy photosynthesis must consider the heterogeneous radiation in canopies and the non-linear response of photosynthesis to irradiance. At the same time, the heterogeneous light condition within the canopy is affected by solar angles, incident photon flux during the day, and different GAI values. All these complexities shed doubt whether relations between photosynthetic parameters and leaf photosynthesis apply when scaling up to canopy level photosynthesis. Using the canopy model decoupled from GECROS, we simulated A'_c with various GAIs and under different environmental conditions.

As shown by Figs 2 and 3, leaf photosynthesis at saturating light (A_{2000}) correlated well with daily canopy carbon gain at high light intensity when GAI was 1 or 5 (i.e. $A'_{c,1}$ & $A'_{c,5}$). Moreover, component parameters of leaf photosynthesis show similar correlations with A_{2000}, $A'_{c,1}$ and $A'_{c,5}$. In Fig. 2, we found more or less the same genetic gain or loss at leaf (A_{2000}), and canopy level ($A'_{c,1}$ and $A'_{c,5}$) for each genotype. For A_{2000}, genetic variation was 25.6%, which is comparable with 26.5% for $A'_{c,1}$ and 25.8% for $A'_{c,5}$ (Table 1). All these results suggest that genetic variation in leaf photosynthesis in this IL population scales up well to canopy level.
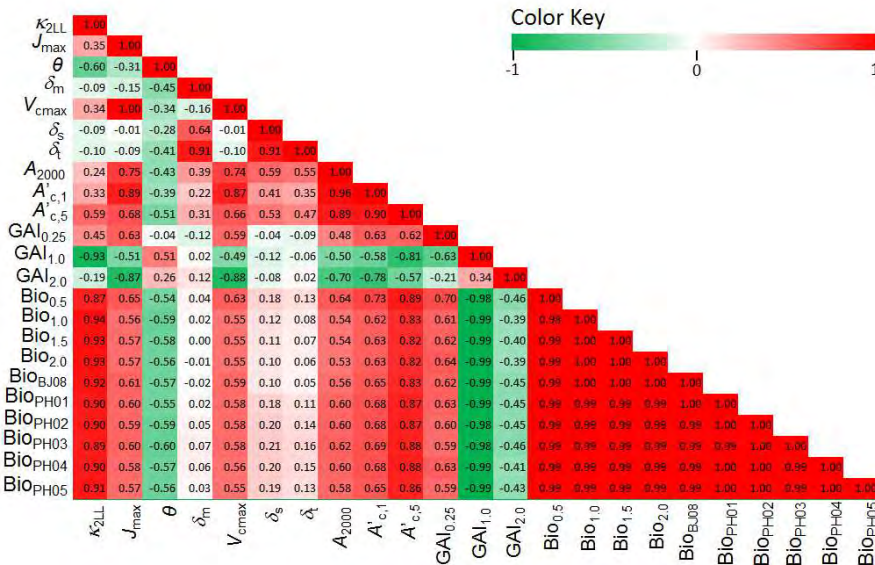

Fig. 3. Pearson correlation coefficient heat map of parameters from leaf level, canopy level to crop level. Correlations are scaled by the colour of the corresponding cell. Parameters are represented in the same order on the *x*- and *y*-axes. The meaning of symbols and abbreviations is the same as in Fig. 2.

Table 1. Minimum, maximum and population mean of traits, and observed genetic variation of introgression lines at leaf level, canopy level and crop level. For traits, see explanation in Fig. 2. Genetic variation was calculated as $(X_{max} - X_{min}) / \overline{X}$, where X_{max} and X_{min} stands for maximum and minimum value in the population, respectively; $\overline{X}$ stands for the population mean.

	Trait	Min	Max	Population mean	Genetic variation (%)
Leaf level	κ_{2LL}	0.272	0.369	0.324	29.8
	J_{max}	136.5	167.1	147.5	20.7
	θ	0.723	0.916	0.790	24.5
	δ_m	0.730	0.908	0.878	20.3
	V_{cmax}	109.6	152.7	124.8	34.6
	δ_s	0.806	1.355	1.109	49.5
	δ_l	0.383	0.544	0.489	32.8
	A_{100}	2.9	4.1	3.6	31.4
	A_{500}	13.4	16.2	14.9	18.7
	A_{2000}	19.3	24.9	22.4	25.4
Canopy level	$A'_{c,1}$	25.5	33.2	28.9	26.4
	$A'_{c,5}$	78.7	101.7	89.2	25.8
Crop level	$GAI_{0.25}$	1.86	2.09	1.93	12.0
	$GAI_{1.0}$	4.64	6.09	5.38	26.9
	$GAI_{2.0}$	0.89	1.01	0.94	13.0
	$Bio_{0.5}$	672	949	806	34.4
	$Bio_{1.0}$	1897	2554	2241	29.3
	$Bio_{1.5}$	2142	2799	2482	26.5
	$Bio_{2.0}$	2097	2740	2422	26.5
	Bio_{BJ08}	2092	2775	2436	28.0
	Bio_{PH01}	2049	2748	2409	29.0
	Bio_{PH02}	2352	2988	2692	23.6
	Bio_{PH03}	2370	2952	2693	21.6
	Bio_{PH04}	2240	2881	2569	24.9
	Bio_{PH05}	2113	2794	2470	27.6

We also examined the light response of daily canopy photosynthesis under various environmental conditions. For that purpose, the genotype-specific light response of daily A'_c at various air temperatures, vapour pressures, and GAIs were calculated for the site with 39°54'N latitude (Beijing) on day 107 with a day-length of 13.1 hour (a typical day in the season there) (Fig. 4; Table 2) assuming a uniform light distribution over the day. At an air temperature of 25 °C, a GAI of 5, and a vapour pressure of 1.5 kPa, the genetic variation of the daily A'_c among the ILs was 30.3% at low light (100 μmol m^{-2} s^{-1}, ~PAR = 1.04 MJ m^{-2} d^{-1}) level, 31.0% at intermediate light (500 μmol m^{-2} s^{-1}, ~PAR = 5.18 MJ m^{-2} d^{-1}) level, and 25.8% at high light (2000 μmol m^{-2} s^{-1}, ~PAR = 20.7 MJ m^{-2} d^{-1}) level (Table 2). When adjusting air temperature, canopy photosynthesis changed dramatically, especially when canopy photosynthesis was compared for an air temperature of 40°C with that at 10°C or 25°C. Still, the genetic
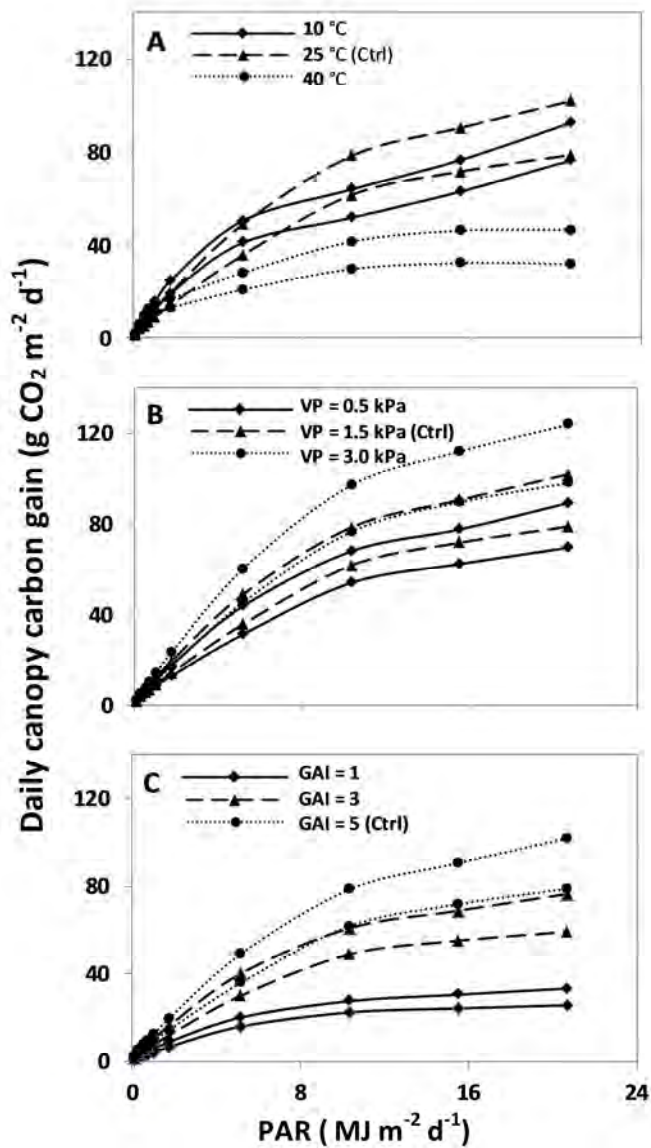
Fig. 4. Canopy photosynthetic responses to photosynthetically active radiation (PAR) under different air temperature (A), vapour pressure (VP; B) and green-surface area index (GAI; C) for the introgression line population. For clarity, only maximum and minimum values of the IL population are shown. The control is for location Beijing, the 107th day of the year 2009, with day length of 13.11 hours, and assuming a uniform light distribution over the day, a constant air temperature of 25 °C, an ambient CO_2 concentration of 380 μmol mol^{-1}, a green-surface area index of 5 and a vapour pressure of 1.5 kPa. For each individual simulation, control parameters were used for the whole introgression line population except for the tested parameters that were varied.

Table 2. Genetic variation in daily canopy photosynthesis (A'_c) at different combinations of air temperature, vapour pressure (VP), and green-surface area index (GAI).

		PAR (MJ m^{-2} d^{-1})	A'_c (min) (g CO$_2$ m^{-2} d^{-1})	A'_c (max) (g CO$_2$ m^{-2} d^{-1})	Population mean	Genetic variation (%)
Control *(air temp. = 25°C, GAI = 5, VP = 1.5 kPa)*		1.035	9.17	12.44	10.81	30.3
		5.175	35.75	48.94	42.60	31.0
		20.700	78.72	101.71	89.16	25.8
Air temperature *(GAI = 5, VP = 1.5 kPa)*	10°C	1.035	12.11	15.84	13.92	26.8
		5.175	41.10	50.47	44.20	21.2
		20.700	76.41	92.44	82.26	19.5
	40°C	1.035	11.44	14.98	12.67	28.0
		5.175	20.91	28.19	24.06	30.3
		20.700	31.95	46.57	38.36	38.1
Vapour pressure (VP) *(air temp. = 25°C, GAI = 5)*	0.5 kPa	1.035	8.53	11.57	10.02	30.3
		5.175	31.31	43.68	37.89	32.7
		20.700	69.43	89.03	78.38	25.0
	3.0 kPa	1.035	10.66	14.36	12.60	29.3
		5.175	45.34	60.10	53.07	27.8
		20.700	97.88	123.57	109.55	23.5
Green-surface area index (GAI) *(air temp. = 25°C, VP = 1.5 kPa)*	1	1.035	4.10	5.56	4.84	30.2
		5.175	15.84	19.98	17.73	23.4
		20.700	25.54	33.17	28.86	26.4
	3	1.035	7.56	10.27	8.91	30.5
		5.175	29.90	39.99	35.08	28.8
		20.700	59.07	76.28	66.68	25.8

variation within the IL population was relatively stable for the different levels of light intensity (range 25.8-31.0%), despite its large variation with temperature (Table 2).

Genetic variation in green-surface area index and biomass production
We used GECROS to assess how the spatial and temporal integration of genetic variation in leaf photosynthesis resulted in genetic variation in GAI and total biomass (including dry weight of all living and dead shoot and root materials) in this IL population. Simulated grain yield was not analysed here because of the multiple uncertainties related to the quantification of grain numbers in the model (the model assumes that grain number can be carbon- or nitrogen-determined; also see Chapter 4).

   Fig. 5 illustrates for Beijing, 2009, assuming a total nitrogen uptake of 15 g N m^{-2}, that there was considerable genetic variation in both GAI and total biomass, throughout the growing season. For GAI, there was more genetic variation at flowering (DS = 1.0: 26.9%) than at an early vegetative growth stage (DS = 0.25: 12.0%), or at harvest stage (DS = 2.0: 13.0%) (Table 1). The reason for the larger genetic variation at flowering could be the fully developed canopy and the large amount of nitrogen held in the canopy. At early vegetative stage, the canopy was not fully developed yet, and therefore the genetic variation was not fully expressed. At
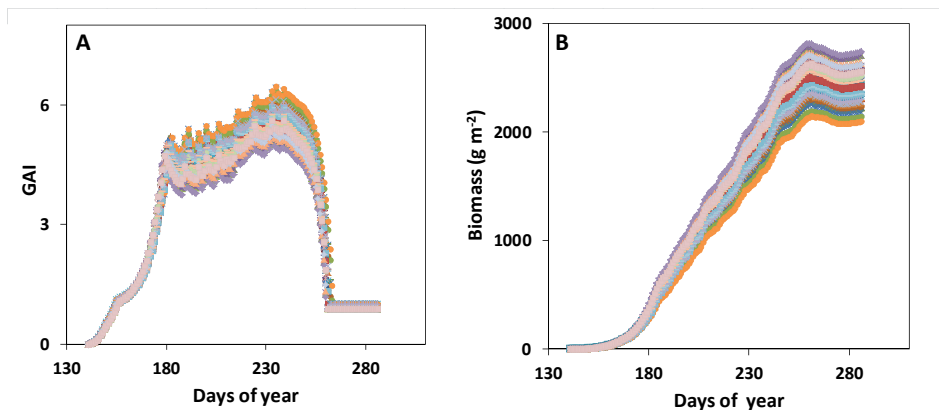
Fig. 5. Time courses of calculated (A) green-surface area index (GAI; m^2 m^{-2}) and (B) total biomass (g m^{-2}) of 38 genotypes in Beijing during the year 2009.

harvest stage, leaves were senescing and most nitrogenous compounds in the leaves were decomposed and the N translocated to the grains. For biomass, the genetic variation at DS 0.5 (~ tillering stage), 1.0 (flowering), 1.5 (mid-grain filling) and 2.0 (harvest) was 34.4%, 29.3%, 26.5% and 26.5%, respectively.

There was a negative relationship between GAI and total biomass, especially at flowering stage (see also Fig. 3). The negative relationship could be due to the feedback mechanisms in the model. A high rate of canopy photosynthesis is associated with high J_{max} and V_{cmax} (Fig. 3), which means more nitrogen content per leaf area, whereas in the model GAI is co-determined by available carbon assimilates and canopy nitrogen content (Yin *et al.*, 2000c). The feedback mechanism in the GECROS model -that current high photosynthesis may dilute leaf nitrogen in the subsequent days-, can lead the model to predict an accelerated leaf senescence, a phenomenon generally also observed experimentally (Fangmeier *et al.*, 2000; Ainsworth & Long, 2005). In Fig. 4C, after full canopy closure had been reached (i.e. GAI = 3), canopy photosynthesis hardly increased with a further increase in GAI. This may explain why a plant with improved leaf photosynthesis and a comparatively small GAI, still has an advantage in biomass production.

In view of the fact that environmental factors considerably influence both leaf and canopy photosynthesis, we show the dynamic pattern of GAI and biomass during an entire growing season (Beijing 2009), and biomass production at different latitudes and in different years (Beijing 2008, 2009 and Los Baños 2001-2005) (Fig. 2). The simulations carried out for both Beijing and Los Baños have a similar tendency across different growth stages growing seasons and locations, despite of variation in the

climatic variables across years and locations (Supplementary Fig. S1). In fact, genetic variation for each level mostly ranged between 20 and 30% (Table 1).

Our model simulation already showed that genetic variation in leaf photosynthesis among the 38 genotypes resulted in an average 25.9% increase in biomass production. Further improvement is still possible. Based on the seven QTL regions, there are in total $2^7 = 128$ possible genotypes. By evaluating the 128 possible genotypes, the potential ideotype that combines all positive alleles for photosynthesis was simulated to have a biomass production advantage of 5.6%, 7.4%, 8.8%, 7.0%, 5.1%, 7.6%, and 8.2%, when compared with the best IL of the 38 genotypes, for BJ08, BJ09, PH01, PH02, PH03, PH04, PH05, respectively.

On contribution of leaf photosynthesis components to biomass productivity

There has been a long-standing controversy as to whether an increase in leaf-level photosynthesis would increase yield (Evans & Dunstone, 1970; Borrás *et al.*, 2004; Long *et al.*, 2006). It is commonly assumed that even when there is an improvement in leaf photosynthesis components, the effects will be diluted through biological hierarchy, resulting in only a small effect at canopy level or crop level. For example, by assuming a widely observed inverse relationship between maximum catalytic rates of carboxylation per active site (k^c_c) and Rubisco specificity ($S_{c/o}$), Zhu *et al.* (2004a) showed that replacing the average Rubisco of terrestrial C_3 plants by Rubisco with an optimal $S_{c/o}$ would only increase canopy photosynthesis by 3%. Sinclair *et al.* (2004) presented an calculation for soybean, starting with an assumed 50% increase in the production of mRNA for synthesis of Rubisco, but ending with only 6% increase or even a 6% decrease in yield depending on whether there is extra nitrogen accumulation possible or not. Yin & Struik (2008) assessed the impact of a successful introduction of the full C_4 system into rice. The simulation resulted in ca 25% yield increase, lower than the originally expected 50% increase.

These results are not surprising, since photosynthesis in the canopy can be either light saturated or light limited. Light-limited photosynthesis is electron transport-limited, whereas light-saturated photosynthesis is generally Rubisco-limited, particularly at lower [CO_2]. In Zhu's calculation, the optimal specificity of Rubisco will increase leaf photosynthesis at high light level (>400 µmol m^{-2} s^{-1}), this effect was weakened by an opposite effect at low light (due to the negative relationship between $S_{c/o}$ and k^c_c). In Sinclair's simulation, 50% more mRNA for synthesis of the subunits of Rubisco only increased light-saturated photosynthesis. If there were no additional N inputs, the required investment in Rubisco will cause less nitrogen being available for subsequent transfer to the seed and the seed becomes nitrogen limited (i.e. 6% decrease in yield). In Yin & Struik's analysis on the potential benefit from introducing

Table 3. Regression coefficients (with standard errors between brackets), intercept, and total variation accounted for (R^2) for multiple linear regression analysis of total biomass (Bio) as a function of κ_{2LL}, J_{max}, θ, δ_m, V_{cmax}, and δ_s (i.e. Bio $= b_0 + b_1\kappa_{2LL} + b_2J_{max} + b_3\theta + b_4\delta_m + b_5V_{cmax} + b_6\delta_s$), based on data of an introgression line population. Regressions were made for locations Beijing, 2008 (BJ08); Beijing, 2009 (BJ09) and Los Baños, Philippines from 2001 to 2005 (PH01, PH02, …, PH05, respectively). The meaning of symbols are same as in Fig. 2.

Trait	Regression coefficient (probability of significance)						Intercept (b_0)	R^2 (%)
	b_1	b_2	b_3	b_4	b_5	b_6		
Bio$_{BJ08}$	**5483.1 (<1.0 × 10^{-8})**	**11.3 (0.0035)**	**604.3 (2.6 × 10^{-6})**	**176.9 (0.0012)**	-3.8 (0.1520)	**224.3 (<1.0 × 10^{-8})**	-1427.6	99.5
Bio$_{BJ09}$	**5806.6 (<1.0 × 10^{-8})**	**8.2 (0.0400)**	**774.6 (9.0 × 10^{-8})**	**259.7 (2.8 × 10^{-5})**	-1.9 (0.4813)	**213.0 (<1.0 × 10^{-8})**	-1500.9	99.4
Bio$_{PH01}$	**5806.2 (<1.0 × 10^{-8})**	**10.5 (0.0069)**	**864.7 (<1.0 × 10^{-8})**	**219.7 (0.0001)**	-2.9 (0.2606)	**340.5 (<1.0 × 10^{-8})**	-1902.0	99.5
Bio$_{PH02}$	**5433.8 (<1.0 × 10^{-8})**	-1.4 (0.7690)	**791.4 (3.8 × 10^{-6})**	**258.9 (0.0005)**	5.0 (0.1549)	**321.2 (<1.0 × 10^{-8})**	-692.3	99.0
Bio$_{PH03}$	**5040.2 (<1.0 × 10^{-8})**	-0.5 (0.9356)	**709.0 (0.0003)**	**274.8 (0.0023)**	4.2 (0.3296)	**298.6 (<1.0 × 10^{-8})**	-527.9	98.3
Bio$_{PH04}$	**5572.8 (<1.0 × 10^{-8})**	7.8 (0.0597)	**803.5 (1.0 × 10^{-7})**	**265.0 (3.7 × 10^{-5})**	-1.4 (0.6133)	**334.2(<1.0 × 10^{-8})**	-1440.8	99.3
Bio$_{PH05}$	**5820.8 (<1.0 × 10^{-8})**	**7.0 (0.0414)**	**796.0 (<1.0 × 10^{-8})**	**186.1 (0.0003)**	-1.2 (0.6272)	**359.4 (<1.0 × 10^{-8})**	-1503.6	99.6

Coefficient values significant at a level of $P < 0.05$ are in bold.

the C_4 system into C_3 rice through the CO_2-concentrating mechanism, light-saturated photosynthesis significantly improved while the light-limited photosynthesis at leaf level remained unchanged. This may explain the much lower increase in yield than originally expected.

Using our simulation results, we performed multiple regression analysis to indicate which parameters are most important for determining final biomass. Multiple regression analysis showed that the leaf-level genetic variation in κ_{2LL} contributed most to the variation in total biomass, followed by δ_s, θ, δ_m and J_{max} (Table 3), which somewhat differed from the results of the simple correlation analysis (Fig. 3). Since κ_{2LL} contributes to electron transport at limiting light, this result was in line with the assumption that in a canopy most leaves were in the state of electron transport-limited photosynthesis. This is further supported by the fact that V_{cmax} did not significantly contribute to the explanation of observed variance in any of the simulations (Table 3). This may explain the results of our simulation, which unlike most earlier simulation studies, incorporated the genetic variation in both light-saturated and light-limited photosynthesis parameters. Our results are in line with the report of Day & Chalabi (1988) that a same percentage increase in quantum use efficiency, relative to light-saturated photosynthetic capacity, resulted in a more significant increase of canopy photosynthesis.

CONCLUDING REMARKS

In our previous research (Chapter 3), we not only found QTLs for light-saturated photosynthesis (e.g. QTLs for V_{cmax}, δ_m, δ_s), but also QTLs contributing to light-limited photosynthesis (e.g. QTLs for κ_{2LL}, θ). This could also be the explanation why genetic variation in our genetic population showed to scale up equally from leaf, to canopy to crop-level (Fig. 2). Our results showed a very promising approach to increase plant production through conventional marker-assisted breeding. This is in line with Parry *et al.* (2011), who suggested that improving photosynthesis through mining existing germplasm is the most efficient way. There have been reports on considerable genetic variation between rice cultivars (Adachi *et al.*, 2011; Chapter 2) to be utilized in a breeding programme. Further progress could be enhanced through recent advances in genome wide association studies (Huang *et al.*, 2010), by exploiting a broader genetic background.

APPENDIX 1. Main equations used to simulate carbon partitioning, nitrogen demand and partitioning, phonological development, and canopy and leaf photosynthesis.

Meaning of the symbols, derivation of the algorithms, and a detailed description of the GECROS and its algorithms, can be found in Yin & van Laar (2005).

Important equations in GECROS

$$N_{demD} = W_S(n_{cri} - n_{act})(1 + N_R/N_S)/\Delta t \tag{1}$$

$$n_{cri} = n_{cri0}e^{-0.4\vartheta} \tag{2}$$

$$N_{demA} = C_R \sigma_C^2/(d\sigma_C/d\kappa) \tag{3}$$

$$\sigma_C = (\Delta C/\Delta t)/C_S \tag{4}$$

$$d\sigma_C/d\kappa = [\sigma_{C(\kappa+\Delta\kappa)} - \sigma_{C(\kappa)}]/\Delta\kappa \tag{5}$$

$$N_{dem} = \min[N_{maxup}, \max(N_{demD}, N_{demA})] \tag{6}$$

$$C_{\vartheta i} = \omega_i C_{max} \frac{(2\vartheta_e - \vartheta_m)(\vartheta_e - \vartheta_i)}{\vartheta_e(\vartheta_e - \vartheta_m)^2} \left(\frac{\vartheta_i}{\vartheta_e}\right)^{\frac{\vartheta_m}{(\vartheta_e - \vartheta_m)}} \tag{7}$$

$$\omega_i = \begin{cases} g(T)/m_V & \vartheta \leq \vartheta_1 \text{ or } 1 \geq \vartheta \geq \vartheta_2 \\ g(T)h(D_{lp})/m_V & \vartheta_1 < \vartheta < \vartheta_2 \\ g(T)/m_R & \vartheta > 1 \end{cases} \tag{8}$$

$$g(T) = \left[\left(\frac{T_C - T}{T_C - T_O}\right)\left(\frac{T - T_b}{T_O - T_b}\right)^{\left(\frac{T_O - T_b}{T_C - T_O}\right)}\right]^{C_t} \tag{9}$$

$$h(D_{lp}) = 1 - P_{sen}(D_{lp} - M_{op}) \tag{10}$$

Canopy photosynthesis sub-model

$$\phi_{su,i} = e^{-k_b L_i} \tag{11}$$

$$\phi_{su} = \frac{1}{L}\int_0^L e^{-k_b L_i} dL_i = (1 - e^{-k_b L})/(k_b L) \tag{12}$$

$$\phi_{sh} = 1 - \phi_{su} \tag{13}$$

$$I_c = (1 - \rho_{cb})I_{b0}\left(1 - e^{-k_b' L}\right) + (1 - \rho_{cb})I_{d0}\left(1 - e^{-k_d' L}\right) \tag{14}$$

$$I_{c,su} = (1 - \sigma)I_{b0}(1 - e^{-k_b L}) + (1 - \rho_{cb})I_{d0} \frac{k_d' \left[1 - e^{-(k_d' + k_b)L}\right]}{k_d' + k_b}$$

$$+ I_{b0} \left\{ (1 - \rho_{cb}) \frac{k_d'\left[1 - e^{-(k_d' + k_b)L}\right]}{k_d' + k_b} - (1 - \sigma)\frac{1 - e^{-2k_b L}}{2} \right\} \tag{15}$$

$$I_{c,sh} = I_c - I_{c,su} \tag{16}$$

$$N_c = n_0(1 - e^{-k_n L})/k_n - n_b L \tag{17}$$

$$N_{c,su} = n_0\left[1 - e^{-(k_n + k_b)L}\right]/(k_n + k_b) - n_b(1 - e^{-k_b L})/k_b \tag{18}$$

$$N_{c,sh} = N_c - N_{c,su} \tag{19}$$

$$P^{25} = S_{Na}(N_a - n_b) \quad \text{(nitrogen dependence for } V_{cmax} \text{ and } J_{max} \text{ at 25 °C)} \tag{20}$$

Leaf photosynthesis sub-model

$$A_c = \frac{(C_c - \Gamma_*)V_{cmax}}{C_c + K_{mc}(1 + O/K_{mo})} - R_d \tag{21}$$

$$A_j = \frac{(C_c - \Gamma_*)J}{4C_c + 8\Gamma_*} - R_d \tag{22}$$

$$A = \min(A_c, A_j) \tag{23}$$

$$J = \left(\kappa_{2LL}I_{inc} + J_{max} - \sqrt{[\kappa_{2LL}I_{inc} + J_{max}]^2 - 4\theta J_{max}\kappa_{2LL}I_{inc}}\right)/(2\theta) \tag{24}$$

$$g_s = g_{so} + \delta_s(A + R_d)/(C_c - \Gamma_*) \tag{25}$$

$$g_m = g_{mo} + \delta_m(A + R_d)/(C_c - \Gamma_*) \tag{26}$$

$$g_t = \delta_t(A + R_d)/(C_c - \Gamma_*) \tag{27}$$

$$\delta_t = \frac{a_1}{1 + VPD*D_0} \tag{28}$$

$$\text{Parameter} = \text{Parameter}_{25}e^{\frac{(T_l - 25)E}{[298R(T_l + 273)]}} \quad \text{(for } R_d, V_{cmax}, K_{mc}, K_{mo}) \tag{29}$$

$$J_{max} = J_{max25}e^{\frac{(T_l - 25)E}{[298R(T_l + 273)]}} \times \frac{1 + e^{(298S - D)/(298R)}}{1 + e^{[(T_l + 273)S - D]/[R(T_l + 273)]}} \tag{30}$$

Table. Symbols (with units) used in Eqn (1) – (30) in the APPENDIX.

Symbols	Description
a_1	an empirical coefficient, see Eqn (28) (-)
A	Net photosynthesis rate (μmol m^{-2} s^{-1})
A_c	Rubisco activity limited net photosynthesis rate (μmol m^{-2} s^{-1})
A_j	Electron transport limited net photosynthesis rate (μmol m^{-2} s^{-1})
c_t	curvature factor in Eqn (9) (-)
C	total carbon in live material of the crop (g C m^{-2} ground)
C_c	CO_2 concentration at the carboxylation site of Rubisco (μmol mol^{-1})
C_{max}	maximum carbon content of stem or seed at the end of its growth (g C m^{-2} ground)
C_R	carbon in live root (g C m^{-2} ground)
C_S	carbon in live shoot (g C m^{-2} ground)
$C_{\vartheta i}$	carbon demand for growth of an organ at stage ϑ_i (g C m^{-2} ground d^{-1})
D	Deactivation energy (J mol^{-1})
D_0	an empirical coefficient, see Eqn (28) (kPa)
D_{lp}	daylength for photoperiodic response of phenology (h)
E	activation energy (J mol^{-1})
$g(T)$	function for phonological response to temperature (-)
g_m	Mesophyll conductance (mol m^{-2} s^{-1})
g_{mo}	Residual mesophyll conductance in the g_m model Eqn (7) (mol m^{-2} s^{-1})
g_s	Stomatal conductance for CO_2 (mol m^{-2} s^{-1})
g_{so}	residual Stomatal conductance for CO_2 (mol m^{-2} s^{-1})
g_t	Diffusion conductance from ambient air to the site of carboxylation (mol m^{-2} s^{-1})
$h(D_{lp})$	function for phenological response to photoperiod (-)
I_{b0}	incident direct-beam radiation above canopy (J m^{-2} grounds^{-1})
$I_{c,sh}$	absorbed radiation by shaded leaves of canopy (J m^{-2} grounds^{-1})
$I_{c,su}$	absorbed radiation by sunlit leaves of canopy (J m^{-2} grounds^{-1})
I_c	absorbed radiation by canopy (J m^{-2} grounds^{-1})
I_{d0}	incident diffuse radiation above canopy (J m^{-2} grounds^{-1})
I_{inc}	Photon flux density incident on leaves (μmol photon m^{-2} s^{-1})
J	e$^-$ transport rate through PSII used for NADP$^+$ reduction (μmol e$^-$ m^{-2} s^{-1})
κ	nitrogen-carbon ratio in crop (g N g^{-1} C)
k_b	direct-beam radiation extinction coefficient (m^2 ground m^{-2} leaf)
k_b	direct-beam radiation extinction coefficient (m^2 ground m^{-2} leaf)
k_b'	scattered-beam radiation extinction coefficient (m^2 ground m^{-2} leaf)
k_d'	diffuse radiation extinction coefficient (m^2 ground m^{-2} leaf)
k_n	nitrogen extinction coefficient (m^2 ground m^{-2} leaf)
K_{mc}	Michaelis-Menten constant of Rubisco for CO_2 (μbar)
K_{mo}	Michaelis-Menten constant of Rubisco for O_2 (mbar)
L	green leaf area index of canopy (m^2 leaf m^{-2} ground)
L_i	L counted from the top to the i-th layer of canopy (m^2 leaf m^{-2} ground)
m_R	minimum number of days for seed filling phase (d)
m_V	minimum number of days for vegetative growth phase (d)
M_{op}	maximum or minimum optimum photoperiod (h)
n_0	canopy top-leaf nitrogen (g N m^{-2} leaf)
n_{act}	actual nitrogen concentration in living shoot (g N g^{-1} dw)
n_b	minimum leaf nitrogen for photosynthesis (g N m^{-2} leaf)
n_{cri}	critical shoot nitrogen concentration (g N g^{-1} dw)
n_{cri0}	initial critical shoot nitrogen concentration (g N g^{-1} dw)
N_a	leaf nitrogen content per area (g N m^{-2} leaf)

Table continued

Symbols	Description
$N_{c,sh}$	photosynthetically active nitrogen in shade leaves of canopy (g N m^{-2} leaf)
$N_{c,su}$	photosynthetically active nitrogen in sunlit leaves of canopy (g N m^{-2} leaf)
N_c	total photosynthetically active nitrogen in canopy (g N m^{-2} ground)
N_{maxup}	maximum crop nitrogen uptake (g N m^{-2} ground d^{-1})
N_R	nitrogen in live root (g N m^{-2} ground)
N_S	nitrogen in live shoot (g N m^{-2} ground)
O	Oxygen partial pressure (mbar)
P^{25}	value for V_{cmax} or J_{max} at 25 ℃ (μmol m^{-2} s^{-1})
P_{sen}	photoperiod sensitivity of phonological development (h^{-1})
R	universal gas constant (=8.314 J K^{-1} mol^{-1})
R_d	Day respiration (respiratory CO_2 release in the light other than by photorespiration) (μmol m^{-2} s^{-1})
S	Entropy term (J K^{-1} mol^{-1})
T	diurnal temperature (°C)
T_b	base temperature for phenological development (°C)
T_C	ceiling temperature for phenological development (°C)
T_l	leaf temperature (ºC)
T_O	optimum temperature for phenological development (°C)
VPD	vapour pressure deficit (kPa)
V_{cmax}	Maximum rate of Rubisco activity-limited carboxylation (μmol m^{-2} s^{-1})
W_S	weight of live shoot (g dw m^{-2} ground)
δ_m	A parameter in the g_m model, defining $C_c : C_i$ ratio at saturating light (-)
δ_s	A parameters in the g_s model, defining $C_i : C_a$ ratio at saturating light (-)
δ_t	A parameters in the g_t model, defining $C_c : C_a$ ratio at saturating light (-)
κ_{2LL}	Value of conversion efficiency of incident light into J at the strictly limiting light [mol e$^-$ (mol photon)$^{-1}$]
ρ_{cb}	canopy beam radiation reflection coefficient (-)
σ_C	relative shoot activity (g C g^{-1} C d^{-1})
ω_i	development rate at stage ϑ_i (d^{-1})
ϑ_1	development stage at which plant starts to become sensitive to photoperiod (-)
ϑ_2	development stage at which plant ends to respond to photoperiod (-)
ϑ_e	development stage at the end of growth of stem or seed (-)
ϑ_i	development stage during the growth of stem or seed (-)
ϑ_m	development stage at the time of maximal growth rate of stem or seed (-)
ϕ_{sh}	fraction of shaded leaves in in a canopy (-)
$\phi_{su,i}$	fraction of sunlit leaves at canopy depth L_i (-)
ϕ_{su}	fraction of sunlit leaves in a canopy (-)
Γ_*	C_c based CO_2 compensation point in the absence of R_d (μbar)
θ	Convexity factor for response of J to I_{inc} (-)
σ	leaf scattering coefficient (-)
ϑ	development stage (-)

Supplementary Materials in Chapter 5
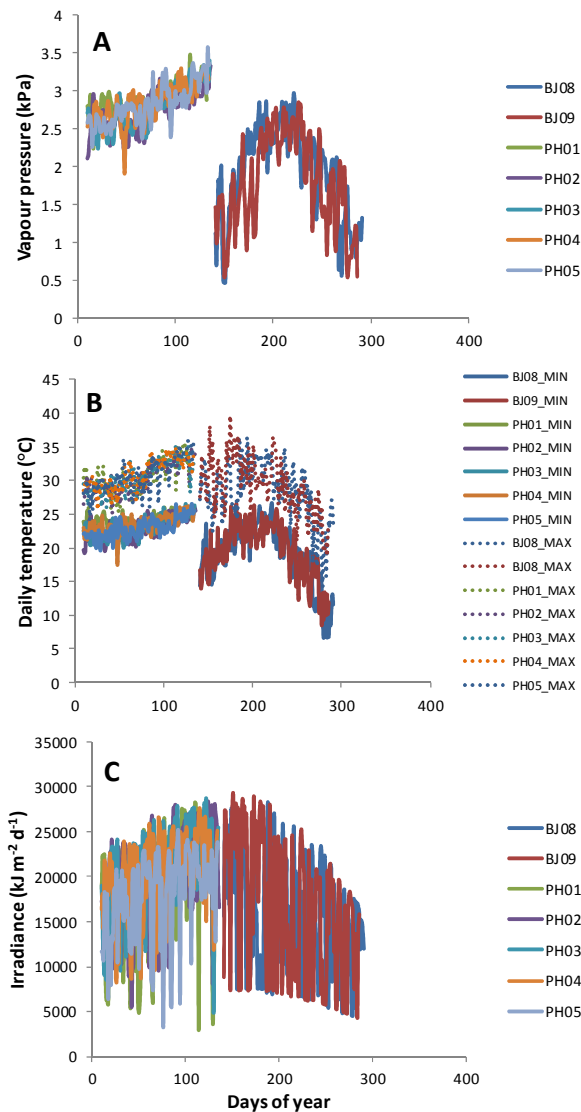
Fig. S1. Variation during the growing season of (A) vapour pressure, (B) daily maximum (MAX) and minimum (MIN) temperature and (C) irradiance for a rice crop at Beijing (BJ), year 2008, 2009 (08, 09) (summer season) and at Los Baños, Philippines (PH), 2001-2005 (01, 02, 03, 04, 05) (dry season).

CHAPTER 6

General discussion

Improving grain yield of rice (*Oryza sativa* L.) for both favourable and stressful environments is the main breeding objective. Increasing rice grain yields will help ensure food security. Crop modelling has long been considered a useful tool to assist breeding (Loomis *et al*., 1979; Whisler *et al*., 1986; Shorter *et al*., 1991; Boote *et al*., 1996, 2001). However, to date the contribution has been small (Miflin, 2000; Tardieu, 2010).

Probably this small contribution was because crop physiologists and modellers did not fully consider the genetic basis of model-input parameters (Stam, 1998), although they often refer to these model-input parameters as 'genetic coefficients' (White & Hoogenboom, 1996; Mavromatis *et al*., 2001; White *et al*., 2008). The development of molecular genetics provided a new method for relating crop model input parameters to their genetic determinants, quantitative trait loci (QTLs). With QTL allelic information and estimated QTL effects, model input trait values can be calculated based on the genotype and used as inputs to crop models thus replacing original measured input trait values. This QTL-based approach can dissect complex traits (e.g. yield) into physiologically relevant component traits, integrate effects of QTL of the component traits over time and space at whole-crop level, and predict yield of various allele combinations under different environmental conditions (Yin & Struik, 2011). Such a QTL-based modelling approach was first introduced by Yin *et al.* (1999a,b; 2000b), to predict differences in yield among relatively similar lines from a genetic population. They showed, however, that improved models were needed in order to make this approach really successful and robust (Yin *et al*., 2004). Later, this approach was used to study crop traits such as leaf elongation rate in maize, flowering time, and fruit quality (Reymond *et al.*, 2003; Quilot *et al.*, 2004; Nakagawa *et al.*, 2005; Quilot *et al.*, 2005; Yin *et al.*, 2005b; Uptmoor *et al.*, 2008; Bertin *et al.*, 2010; Prudent *et al.*, 2011). These later studies showed that this approach was robust in predicting genetic differences in bi-parental crossing populations under different conditions (in terms of vapour pressure deficit, soil moisture content, temperature and photoperiod). But most studies focused on specific traits, only a few used the QTL-based modelling approach to predict complex traits like yield (Yin *et al.*, 2000b). To the best of my knowledge, this approach has not been applied to analyse leaf photosynthesis, especially not to photosynthetic response to stress.

Because of the importance of photosynthesis and its sensitivity to drought, I firstly amalgamate ecophysiological photosynthesis modelling and QTL analysis to study the genetic variation in photosynthesis in a backcross introgression line population developed from a cross between lowland and upland rice cultivars and grown under well-watered and drought conditions. Secondly QTL analysis was further extended to other physiological parameters of rice. Molecular marker-based estimates of these

traits from estimated additive allele effects were used as input into the mechanistic crop model GECROS (Genotype-by-Environment interaction on CROp growth Simulator, Yin & van Laar, 2005). Thirdly, the genetic variation at the level leaves was upscaled up, again by using the GECROS model, to examine whether such a variation can be expressed into the variation of a similar magnitude for biomass production at the crop level. I hope that through this thesis study, useful knowledge and insights were created for closing the gap between genotype and phenotypes across various scales.

In this general discussion, first I will summarize and discuss the results presented in this thesis. I will then evaluate how crop modelling and QTL/marker-based modelling can assist crop breeding and genetics. Finally, I will discuss the future prospects of integrating crop physiology, crop modelling, plant genetics and plant breeding.

RESEARCH FINDINGS OF THIS THESIS

Genetic variation of photosynthesis in the field

Because of the primary importance of photosynthesis in determining crop growth, quantifying genetic variation and identifying QTLs related to enhanced photosynthesis in rice populations are important steps in improving rice productivity. But photosynthesis is greatly influenced by environmental factors. Moreover, the microclimate unavoidably fluctuates under natural field conditions (Flood *et al.*, 2011), making photosynthesis a difficult variable to assess for a large population. For example, it was shown in Chapter 2 that during flowering and grain filling the light-saturated photosynthesis fluctuated more than three-fold, because of fluctuations in temperature and humidity of the air. This will hamper QTL analysis of photosynthesis measured under field conditions. It is quite possible to draw wrong conclusions if the random noise caused by variation in the micro-environment is not removed. In this thesis, both a statistical covariant model and a photosynthesis model were used to standardize observations to the same temperature and vapour pressure deficit (Chapter 2). This approach showed its value and could thus be used also in other situations where the same problem occurs.

Teng *et al.* (2004) and Zhao *et al.* (2008) did not correct for microclimate fluctuations, when measuring photosynthesis under field conditions. They found QTLs for photosynthesis that did not co-localize with QTLs for yield. In contrast, the main QTL associated with variation in photosynthesis among individuals of an introgression line population studied in Chapter 2 was found to be near marker RM410 on Chromosome 9, the same marker that indicated the position of a QTL accounting for variation in yield of this population (Chapter 4). This contrasting result could be due to

differences in genetic backgrounds in the experiments, but may also be caused by ignoring the effects of microclimate on photosynthesis in the studies by Teng *et al.* (2004) and Zhao *et al.* (2008). In our experiments, consistent results across environments and growth stages were obtained, and co-location of physiologically tightly related QTLs was observed. The major QTL near marker RM410 was consistent across both developmental stages and both drought and well-watered conditions. This QTL controlling multiple photosynthetic traits identified under field environment was then successfully confirmed in an independent greenhouse experiment (Chapter 2). These results also suggest that photosynthesis at different stages and under different treatments is, at least to some extent, influenced by the same genetic factors.

Physiological basis for genetic variation in leaf photosynthesis
Only few studies have investigated the physiological basis for reported genetic variation in leaf photosynthesis (Masle *et al.*, 2005; Adachi *et al.*, 2011; Scafaro *et al.*, 2011; Taylaran *et al.*, 2011). The physiological basis of the genetic variation in photosynthesis is therefore still unclear. Dissecting photosynthesis into different physiological processes will help:

- To identify the genetic variation in each biophysical and biochemical component of photosynthesis;
- To evaluate the potential of utilizing the genetic variation in these components for improving photosynthesis (*A*) and transpiration efficiency (TE).

Therefore, combined gas exchange and chlorophyll fluorescence data were collected to assess CO_2 and light response curves of 13 introgression lines grown under moderate drought and well-watered conditions (Chapter 3). These 13 lines were carefully selected as representatives of the population, based on the QTLs for leaf photosynthesis reported in Chapter 2. Using these curves, seven parameters of a combined conductance-FvCB (Farquhar, von Caemmerer, & Berry, 1980) model as proposed by Yin *et al.* (2009b) were estimated. 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. Genetic variation in light saturated photosynthesis and TE 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. Our results also showed for the first time 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. 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 by 17.0% and 25.1%, respectively, compared with the best genotype of the 13 lines investigated. Chapter 3 also presented a novel approach to quantitatively analyse an overall relative limitation of g_m versus g_s on photosynthesis of a genotype under a given condition. 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 (Condon *et al.*, 2002, 2004; Blum, 2005; Barbour *et al.*, 2010). This result would be especially interesting for breeding for semi-arid environments.

Projection of leaf photosynthesis to rice production

Plant growth is driven by the availability of carbohydrates and other assimilates. Rice productivity can be limited by available photosynthetic assimilates from leaves; so an increased photosynthesis is expected to result in higher yields. But recent studies reveal that the relationship between leaf photosynthesis and crop yield is not as straightforward as expected. For example, numerous free air CO_2 elevation studies show that higher rates of photosynthesis do not lead to a commensurate increase in biomass and yield (Ainsworth & Long, 2005). Studies of natural genetic diversity in fact often reveal a negative correlation between leaf photosynthesis and biomass or yield (e.g. Jahn *et al.*, 2011). The effects of engineering for an improved photosynthesis have been projected to damp down gradually when moving up from leaf to crop level, presumably because of complicated constraints and feedback mechanisms (Sinclair *et al.*, 2004; Zhu *et al.*, 2004a; Yin & Struik, 2008). My thesis seemed to support these statements, as there was no correlation between leaf photosynthetic rate and grain yield among individuals of the rice population (Chapter 4).

To examine the extent to which natural genetic variation in photosynthesis can contribute to increasing biomass production and yield of rice, I used the mechanistic model GECROS of Yin & van Laar (2005), and incorporated quantitative information from QTLs for various photosynthetic components to model genetic variation within a rice introgression line population (Chapter 5). It was shown that genetic variation in photosynthesis of 25% can be scaled up equally to crop level, resulting in an increase in biomass of 22-29% depending on location and year. The analysis suggests that the genetic variation in photosynthesis resulted not only from Rubisco-limited photosynthesis but also from electron transport-limited photosynthesis; consequently,

photosynthetic rates could be improved for both light-saturated and light-limited leaves in the canopy. Rice production could thus be significantly improved by mining the natural variation in existing germplasm.

APPLICATION OF THE MODELLING APPROACH

Yin & Struik (2010) summarized the potential added value of robust physiological modelling for classical quantitative genetics and expressed their opinions on the perspectives for modelling gene-trait-crop relationships. Here, based on own experiments and analysis, I will address how modelling can assist crop breeding and genetics.

Models can support the Quantitative Trait Loci (QTLs) 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 (Chapter 2). Using model-based standardization, several QTLs related to photosynthesis were found under field conditions. Ecophysiological models can thus play a role in improving the quality of data on traits that are difficult to phenotype. 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.

Models can dissect complex traits into physiological components

Physiological modelling can dissect complex traits (e.g. photosynthesis or yield) into physiological component traits. In Chapter 3, a photosynthesis model was used to dissect photosynthesis into: (1) stomatal conductance g_s, (2) mesophyll conductance

g_m, and (3) electron transport capacity and Rubisco carboxylation capacity. In Chapter 4, using the crop model GECROS, yield was connected to, and dissected into seven physiological input parameters. 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, in Chapter 3, genetic variation in light-saturated photosynthesis and transpiration efficiency was found to be mainly associated with variation in g_s and g_m. In Chapter 4, the physiological input trait 'total crop nitrogen uptake at maturity' was found to have the most significant effect on yield. 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.

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 QTLs 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 & Struik 2011). 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 & Dunstone, 1970; Teng *et al.*, 2004; Zhao *et al.*, 2008; Jahn *et al.*, 2011; Chapter 4). The reason for this discrepancy could be that plants differed genetically in many respects other than photosynthesis. Hence, I used the crop model GECROS, and found that the natural genetic variation in leaf photosynthesis within our experimental materials 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 QTLs for a specific organ-level trait at crop level in a different environment. 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 QTLs for each physiological component in photosynthesis, it was shown that the ideotype for leaf-level photosynthesis and *TE* could potentially be improved by 17.0% and 25.1%, respectively (Chapter 3).

QTL-based modelling can quantify constraints in breeding

Model simulation could inspire breeders. However, Stam (1998) and Koornneef & Stam (2001), from a geneticist's perspective, expressed their concerns that the ignorance of the inheritance of the model-input traits 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. 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. For example, in Chapter 3, trade-offs were shown between improving photosynthesis and TE either due to tight linkage or to pleiotropic effects of QTLs related to g_m and g_s. If the linkage between g_m and g_s, or co-location of QTLs 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 programs thus strengthening the scientific basis for breeding program design.

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 QTLs with large effect into widely grown varieties (Jena & 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 (Collard & Mackill, 2008). 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). This thesis also showed that the existing GECROS model can be a useful tool to enhance marker-assisted breeding through a model-based ideotype design (Chapter 4). Using the principles for QTL-based modelling as defined earlier (Yin *et al.*, 2000b; 2004; 2005b), marker-based crop modelling was performed in Chapter 4 to rank the markers identified for various yield-determining physiological traits that are input parameters of GECROS. Such an analysis detected markers that breeders can prioritize in their MAS programmes for specific environments. Chapter 4 showed that

compared with identification of markers through multiple regression for yield *per se*, the model-based approach identified additional QTLs and could be complementary to the analysis of yield *per se*.

FUTRUE PROSPECTS OF INTEGRATION OF CROP MODELLING AND PLANT GENETICS

In this thesis, QTLs were identified for a set of physiological parameters associated with leaf photosynthesis (Chapter 2), phenological development rates and morphological traits (Chapter 4). These traits were used as input parameters of the crop model GECROS. With their marker/QTL-based estimates as input to the model, the QTL effects for traits typically at the single-organ level over a short time scale, were projected for their impact on crop growth during the whole growing season in the field (Chapter 4). This thesis provided a strong case where the information from functional genomics can be brought up to crop level via modelling.

I first focused on leaf photosynthesis by using a detailed biochemical photosynthesis model (Chapters 2 and 3), and then scaled up to crop level by the GECROS model which uses the concept of carbon-nitrogen interactions for a balanced modelling of crop growth (Chapters 4 and 5). This analysis showed that the modelling strategy can promote communication across scales from the level of leaves, through canopy to crop level.

Systems simulation modelling has long been suggested as a powerful tool to understand crop yield formation and to assist crop improvement programmes (e.g., Loomis *et al*., 1979). However, modelling studies at the crop level using some knowledge of fundamental plant biology (e.g., biochemistry, genomics) are currently still sporadic (Yin & Struik, 2008). Some model algorithms are based on untested or empirical hypotheses, or even missing. For example, it was shown in Chapter 4 that a better account for the final spikelet number of rice is needed when applying the model to drought environments, when high tissue organ temperature and high spikelet sterility can be expected (Jagadish *et al.*, 2007). This indicates that model components related to sink formation still need to be improved especially for predictions under stress environments.

In this thesis I used the approach of Yin *et al*. (2000b; 2004) and performed QTL-based physiological modelling of leaf photosynthesis and crop productivity in rice. The approach could be expanded. Gene-based crop modelling has already practiced by White & Hoogenboom (1996), Messina *et al.* (2006), and White (2008), on an empirical basis though. The fast development of genomics with second-generation genome sequencing and genome-wide association studies may enhance opportunities for developing gene-based modelling. The advance of transcriptomics, proteomics,

metabolomics, and phenomics may enhance the link between genome data, metabolic pathways and processes, physiological component processes, and crop yield and production. Accordingly, different temporal, spatial and structural scales are required for different components, pathways, and processes of the model (Yin & Struik, 2008). In the end, such 'crop systems biology' approach (Yin & Struik, 2008, 2010) should enable *in silico* assessment of crop response to genetic fine-tuning under defined environmental scenarios, thereby providing a powerful tool in support of breeding for complex crop traits.

References

Adachi S, Tsuru Y, Nito N, Murata K, Yamamoto T, Ebitani T, Ookawa T, Hirasawa T. 2011. Identification and characterization of genomic regions on chromosomes 4 and 8 that control the rate of photosynthesis in rice leaves. *Journal of Experimental Botany* **62**, 1927-1938.

Ainsworth EA, Long SP. 2005. What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2. *New Phytologist* **165**, 351-372.

Ando T, Yamamoto T, Shimizu T, Ma XF, Shomura A, Takeuchi Y, Lin SY, Yano M. 2008. Genetic dissection and pyramiding of quantitative traits for panicle architecture by using chromosomal segment substitution lines in rice. *Theoretical and Applied Genetics* **116**, 881-890.

Archontoulis SV, Yin X, Vos J, Danalatos NG, Struik, PC. 2012. Leaf photosynthesis and respiration of three bioenergy crops in relation to temperature and leaf nitrogen: how conserved are biochemical model parameters among crop species? *Journal of Experimental Botany* **63**: 895-911.

Ashikari M, Matsuoka M. 2006. Identification, isolation and pyramiding of quantitative trait loci for rice breeding. *Trends in Plant Science* **11**, 344-350.

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M. 2005. Cytokinin oxidase regulates rice grain production. *Science* **309**, 741-745.

Asins M. 2002. Present and future of quantitative trait locus analysis in plant breeding. *Plant Breeding* **121**, 281-291.

Austin R. 1999. Yield of wheat in the United Kingdom: recent advances and prospects. *Crop Science* **39**: 1604-1610.

Austin RB. 1994. Plant breeding opportunities. In: KJ Boote, JM Bennett, TR Sinclair & GM Paulsen (eds) Physiology and Determination of Crop Yield. P 567-585. ASA, CSSA and SSSA. Madison, WI, USA.

Azcón-Bieto J, Farquhar GD, Caballero A. 1981. Effects of temperature, oxygen concentration, leaf age and seasonal variations on the CO_2 compensation point of *Lolium perenne* L. *Planta* **152,** 497-504.

Babu RC, Nguyen BD, Chamarerk V, Shanmugasundaram P, Chezhian P, Jeyaprakash P, Ganesh SK, Palchamy A, Sadasivam S, Sarkarung S, Wade LJ, Nguyen HT. 2003. Genetic analysis of drought resistance in rice by molecular markers: association between secondary traits and field performance. *Crop Science* **43,** 1457–1469.

References

Ball JT, Woodrow IE, Berry JA. 1987. A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In: Biggins J (ed) Progress in Photosynthesis Research, Vol. IV. Martinus-Nijhoff Publishers, Dordrecht, The Netherlands, pp 221–224.

Barbour MM, Warren CR, Farquhar GD, Forrester GUY, Brown H. 2010. Variability in mesophyll conductance between barley genotypes, and effects on transpiration efficiency and carbon isotope discrimination. *Plant, Cell & Environment* **33,** 1176-1185.

Bernacchi CJ, Portis AR, Nakano H, Caemmerer Sv, Long SP. 2002. Temperature response of mesophyll conductance. Implications for the determination of Rubisco enzyme kinetics and for limitations to photosynthesis in vivo. *Plant Physiology* **130**, 1992–1998.

Bernacchi CJ, Singsaas EL, Pimentel C, Portis AR, Long SP. 2001. Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell and Environment* **24,** 253-259.

Bernier J, Kumar A, Ramaiah V, Spaner D, Atlin G. 2007. A large-effect QTL for grain yield under reproductive-stage drought stress in upland rice. *Crop Science* **47,** 507-516.

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

Bindraban PS, Sayre KD, Solis-Moya E. 1998. Identifying factors that determine kernel number in wheat. *Field Crops Research* **58**, 223-234.

Blum A. 2005. Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive? *Australian Journal of Agricultural Research* **56**, 1159-1168.

Boote KJ, Jones JW, Pickering NB. 1996. Potential uses and limitations of crop models. *Agronomy Journal* **88**, 704-716.

Boote KJ, Kropff MJ, Bindraban PS. 2001. Physiology and modelling of traits in crop plants: implications for genetic improvement. *Agricultural Systems* **70**, 395-420.

Borrás L, Slafer GA, Otegui ME. 2004. Seed dry weight response to source–sink manipulations in wheat, maize and soybean: a quantitative reappraisal. Field Crops Research **86:** 131-146.

Bouman B, Humphreys E, Tuong T, Barker R. 2007. Rice and water. *Advances in Agronomy* **92**, 187-237.

Bouman B, Van Keulen H, Van Laar H, Rabbinge R. 1996. The 'School of de Wit' crop growth simulation models: A pedigree and historical overview. *Agricultural Systems* **52**, 171-198.

Boyer JS, Westgate ME. 2004. Grain yields with limited water. *Journal of Experimental Botany* **55**, 2385-2394.

Bradbury M, Baker NR. 1984. A quantitative determination of photochemical and non-photochemical quenching during the slow phase of the chlorophyll fluorescence induction curve of bean leaves. *Biochimca et Biophysica Acta* **765,** 275–281.

Bräutigam A, Weber APM. 2011. Do metabolite transport processes limit photosynthesis? *Plant Physiology* **155**, 43-48.

Buckley TN, Mott KA. 2002. Stomatal water relations and the control of hydraulic supply and demand. *Progress in Botany* **63,** 309–325.

Casson SA, Hetherington AM. 2010. Environmental regulation of stomatal development. *Current Opinion in Plant Biology* **13,** 90-95.

Cattivelli L, Rizza F, Badeck F-W, Mazzucotelli E, Mastrangelo AM, Francia E, Marè C, Tondelli A, Stanca AM. 2008. Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. *Field Crops Research* **105**, 1-14.

Centritto M, Loreto F, Chartzoulakis K. 2003. The use of low [CO_2] to estimate diffusional and non-diffusional limitations of photosynthetic capacity of salt-stressed olive saplings. *Plant, Cell & Environment* **26**, 585-594.

Centritto M, Lauteri M, Monteverdi MC, Serraj R. 2009. Leaf gas exchange, carbon isotope discrimination, and grain yield in contrasting rice genotypes subjected to water deficits during the reproductive stage. *Journal of Experimental Botany* **60,** 2325-2339.

Charles Edwards DA. 1976. Shoot and Root Activities During Steady-state Plant Growth. *Annals of Botany* **40:** 767-772.

Chaves MM. 1991. Effects of water deficits on carbon assimilation. *Journal of Experimental Botany* **42,** 1–16.

Chaves MM, Flexas J, Pinheiro C. 2009. Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Annals of Botany* **103**, 551–560.

Chaves MM, Pereira JS, Maroco J, Rodrigues ML, Ricardo CPP, Osório ML, Carvalho I, Faria T, Pinheiro C. 2002. How plants cope with water stress in the field? Photosynthesis and growth. *Annals of Botany* **89,** 907–916.

Chen J, Ding J, Ouyang Y, Du H, Yang J, Cheng K, Zhao J, Qiu S, Zhang X, Yao J, Liu K, Wang L, Xu C, Li X, Xue Y, Xia M, Ji Q, Lu J, Xu M, Zhang Q.

2008. A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica–japonica hybrids in rice. *Proceedings of the National Academy of Sciences* **105**, 11436-11441.

Cheng SH, Cao LY, Zhuang JY, Chen SG, Zhan XD, Fan YY, Zhu DF, Min SK. 2007. Super hybrid rice breeding in China: Achievements and prospects. *Journal of Integrative Plant Biology* **49**, 805-810.

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

Chenu K, Chapman SC, Tardieu F, McLean G, Welcker C, Hammer GL. 2009. Simulating the yield Impacts of organ-level quantitative trait loci associated with drought response in maize: a "gene-to-phenotype" modeling approach. *Genetics* **183**, 1507-1523.

Cochard H, Lluis C, Le RX, Thierry A. 2002. Unraveling the effects of plant hydraulics on stomatal closure during water stress in walnut. *Plant Physiology* **128,** 282–290.

Collard BCY, Mackill DJ. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 557-572.

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD. 2002. Improving intrinsic water use efficiency and crop yield. *Crop Science* **42**, 122-131.

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD. 2004. Breeding for high water-use efficiency. *Journal of Experimental Botany* **55**, 2447-2460.

Cook MG, Evans LT. 1983. Some physiological aspects of the domestication and improvement of rice (Oryza spp.). *Field Crops Research* **6**, 219-238.

Cowan IR. 1977. Stomatal behavior and environment. *Advances in Botanical Research* **4,** 117–228.

Day W, Chalabi ZS. 1988. use of models to investigatethe link between the modification of photosynthetic characteristics and improved crop yields. *Plant Physiology Biochemistry* **26**: 511-518.

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

De Wit CT. 1959. Potential photosynthesis of crop surfaces. *Netherlands Journal of Agricultural Science* **7**, 141-149.

Dekkers JCM, Hospital F. 2002. Multifactorial genetics: The use of molecular genetics in the improvement of agricultural populations. *Nature Reviews Genetics* **3**, 22-32.

Delfine S, Alvino A, Zacchini M, Loreto F. 1998. Consequences of salt stress on conductance to CO_2 diffusion, Rubisco characteristics and anatomy of spinach leaves. *Australian Journal of Plant Physiology* **25,** 395-402.

Ding J, Lu Q, Ouyang Y, Mao H, Zhang P, Yao J, Xu C, Li X, Xiao J, Zhang Q. 2012. A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proceedings of the National Academy of Sciences* **109**, 2654-2659.

Dingkuhn M, De Datta SK, Dorffling K, Javellana C. 1989. Varietal differences in leaf water potential, leaf net CO_2 assimilation, conductivity, and water use efficiency in upland rice. *Australian Journal of Agricultural Research* **40,** 1183–1192.

Dingkuhn M, de Vries FWTP, Miezan K. 1993. Improvement of rice plant type concepts: Systems research enables interaction of physiology and breeding. In: de Vries PF, Teng P, Metselaar K, eds. *Systems approaches for agricultural dvelopment*. Dordrecht, the Netherlands: Kluwer Academic Publishers, 19-35.

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A. 2004. Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of Hd1. *Genes & Development* **18**, 926-936.

Douthe C, Dreyer E, Epron D, Warren CR. 2011. Mesophyll conductance to CO_2, assessed from online TDL-AS records of $^{13}CO_2$ discrimination, displays small but significant short-term responses to CO_2 and irradiance in *Eucalyptus* seedlings. *Journal of Experimental Botany* **62,** 5335-5346.

Eshed Y, Zamir D. 1994. Introgressions from *Lycopersicon pennellii* can improve the soluble-solids yield of tomato hybrids. *Theoretical and Applied Genetics* **88,** 891-897.

Eshed Y, Zamir D. 1995. An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* **141,** 1147-1162.

Ethier GJ, Livingston NJ, Harrison DL, Black TA, Moran JA. 2006. Low stomatal and internal conductance to CO_2 versus Rubisco deactivation as determinants of the photosynthetic decline of ageing evergreen leaves. *Plant, Cell & Environment* **29**, 2168–2184.

Evans JR. 1989. Photosynthesis and nitrogen relationships in leaves of C_3 plants. *Oecologia* **78,** 9–19.

References

Evans JR. 1999. Leaf anatomy enables more equal access to light and CO_2 between chloroplasts. *New Phytologist* **143,** 93-104.

Evans JR, Kaldenhoff R, Genty B, Terashima I. 2009. Resistances along the CO_2 diffusion pathway inside leaves. *Journal of Experimental Botany* **60**, 2235-2248.

Evans LT, Dunstone RL. 1970. Some physiological aspects of evolution in wheat. *Australian Journal of Biological Sciences* **23**, 725-742.

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q. 2006. *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoretical and Applied Genetics* **112**, 1164-1171.

Fangmeier A, Chrost B, Högy P, Krupinska K. 2000. CO_2 enrichment enhances flag leaf senescence in barley due to greater grain nitrogen sink capacity. *Environmental and Experimental Botany* **44:** 151-164.

Farquhar GD, Richards RA. 1984. Isotopic composition of plant carbon correlates with water-use efficiency of wheat genotypes. *Australian Journal of Plant Physiology* **11**, 539-552.

Farquhar GD, Sharkey TD. 1982. Stomatal Conductance and Photosynthesis. *Annual review of plant physiology* **33**, 317-345

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.

Farquhar GD, von Caemmerer S, Berry JA. 2001. Models of photosynthesis. *Plant Physiology* **125**, 42-45.

Fischer RA, Edmeades GO. 2010. Breeding and cereal yield progress. *Crop Science* **50:** 85-98.

Flexas J, Bota J, Loreto F, Cornic G, Sharkey TD. 2004. Diffusive and metabolic limitations to photosynthesis under drought and salinity in C_3 plants. *Plant Biology* **6**, 269-279.

Flexas J, Ribas-Carbó M, Diaz-Espejo A, Galmés J, Medrano H. 2008. Mesophyll conductance to CO_2: current knowledge and future prospects. *Plant, Cell & Environment* **31,** 602-621.

Flexas J, Diaz-Espejo A, Galmés J, Kaldenhoff R, Medrano H, Ribas-Carbó M. 2007a. Rapid variations of mesophyll conductance in response to changes in CO_2 concentration around leaves. *Plant, Cell & Environment* **30,** 1284–1298.

Flexas J, Ortuño MF, Ribas-Carbo M, Diaz-Espejo A, Flórez-Sarasa ID, Medrano H. 2007b. Mesophyll conductance to CO_2 in *Arabidopsis thaliana*. *New Phytologist* **175**, 501-511.

Flexas J, Diaz-Espejo A, Berry J, Cifre J, Galmés J, Kaldenhoff R, Medrano H, Ribas-Carbó M. 2007c. Analysis of leakage in IRGA's leaf chambers of open gas exchange systems: quantification and its effects in photosynthesis parameterization. *Journal of Experimental Botany* **58**, 1533–1543.

Flexas J, Galmés J, Gallé A, Gulías J, Pou A, Ribas-Carbo M, Tomàs M, Medrano H. 2010. Improving water use efficiency in grapevines: potential physiological targets for biotechnological improvement. *Australian Journal of Grape and Wine Research* **16,** 106-121.

Flood PJ, Harbinson J, Aarts MGM. 2011. Natural genetic variation in plant photosynthesis. *Trends in Plant Science* **16**, 327-335.

Fujino K, Sekiguchi H, Matsuda Y, Sugimoto K, Ono K, Yano M. 2008. Molecular identification of a major quantitative trait locus, qLTG3–1, controlling low-temperature germinability in rice. *Proceedings of the National Academy of Sciences* **105**, 12623-12628.

Fukuoka S, Saka N, Koga H, Ono K, Shimizu T, Ebana K, Hayashi N, Takahashi A, Hirochika H, Okuno K. 2009. Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science* **325**, 998-1001.

Galle A, Florez-Sarasa I, Aououad HE, Flexas J. 2011. The Mediterranean evergreen Quercus ilex and the semi-deciduous Cistus albidus differ in their leaf gas exchange regulation and acclimation to repeated drought and re-watering cycles. *Journal of Experimental Botany* **62**, 5207-5216.

Galmés J, Medrano H, Flexas J. 2007. Photosynthetic limitations in response to water stress and recovery in Mediterranean plants with different growth forms. *New Phytologist* **175,** 81-93.

Galmés J, Conesa MÀ, Ochogavía JM, Perdomo JA, Francis DM, Ribas-Carbó M, Savé R, Flexas J, Medrano H, Cifre J. 2011. Physiological and morphological adaptations in relation to water use efficiency in Mediterranean accessions of Solanum lycopersicum. *Plant, Cell & Environment* **34**, 245-260.

Gaydon DS, Probert ME, Buresh RJ, Meinke H, Suriadi A, Dobermann A, Bouman B, Timsina J. 2012. Rice in a cropping systems – modelling transitions between flooded and non-flooded soil environments. *European Journal of Agronomy* **30,** 9-24.

Genty B, Briantais J-M, Baker NR. 1989. The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta* **990,** 87-92.

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H. 2002. A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). *Science* **296**, 92-100.

References

Grassi G, Magnani F. 2005. Stomatal, mesophyll conductance and biochemical limitations to photosynthesis as affected by drought and leaf ontogeny in ash and oak trees. *Plant, Cell & Environment* **28**, 834-849.

Gu J. 2007. Development of introgression lines of upland rice (*Oryza sativa* L.) and QTL mapping of root traits. MSc thesis. China Agricultural University.

Hammer GL, Chapman S, Van Oosterom E, Podlich DW. 2005. Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems. *Crop and Pasture Science* **56:** 947-960.

Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, van Eeuwijk F, Chapman S, Podlich D. 2006. Models for navigating biological complexity in breeding improved crop plants. *Trends in Plant Science* **11:** 587-593.

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

Harley PC, Loreto F, Marco GD, Sharkey TD. 1992a. Theoretical considerations when estimating the mesophyll conductance to CO_2 flux by analysis of the response of photosynthesis to CO_2. *Plant Physiology* **98,** 1429-1436.

Harley PC, Thomas RB, Reynolds JF, Strain BR. 1992b. Modelling photosynthesis of cotton grown in elevated CO_2. *Plant, Cell and Environment* **15**, 271-282.

Hattori Y, Nagai K, Furukawa S, Song XJ, Kawano R, Sakakibara H, Wu J, Matsumoto T, Yoshimura A, Kitano H. 2009. The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. *Nature* **460**, 1026-1030.

Havaux M. 1992. Stress tolerance of photosystem II in vivo: antagonistic effects of water, heat, and photoinhibition stresses. *Plant Physiology* **100,** 424-432.

Haverkort AJ, Kooman PL. 1997. The use of systems analysis and modelling of growth and development in potato ideotyping under conditions affecting yields. *Euphytica* **94**, 191-200.

Hayashi N, Inoue H, Kato T, Funao T, Shirota M, Shimizu T, Kanamori H, Yamane H, Hayano-Saito Y, Matsumoto T. 2010. Durable panicle blast-resistance gene *Pb1* encodes an atypical CC-NBS-LRR protein and was generated by acquiring a promoter through local genome duplication. *The Plant Journal* **64**, 498-510.

Hellewell KB, Rasmusson DC, Gallo-Meagher M. 2000. Enhancing yield of semidwarf barley. *Crop Science* **40**, 352-358.

Hemamalini G, Shashidhar H, Hittalmani S. 2000. Molecular marker assisted tagging of morphological and physiological traits under two contrasting

moisture regimes at peak vegetative stage in rice (Oryza sativa L.). *Euphytica* **112**, 69-78.

Herold A. 1980. Regulation of photosynthesis by sink activity - the missing link. *New Phytologist* **86**, 131-144.

Horton P. 2000. Prospects for crop improvement through the genetic manipulation of photosynthesis: morphological and biochemical aspects of light capture. *Journal of Experimental Botany* **51:** 475-485.

Houle D, Govindaraju DR, Omholt S. 2010. Phenomics: the next challenge. *Nature* **11**, 855-866.

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. *Theoretical and Applied Genetics* **95**, 313-320.

Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, Xia G, Chu C, Li J, Fu X. 2009. Natural variation at the *DEP1* locus enhances grain yield in rice. *Nature Genetics* **41,** 494-497.

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 Q-F, Li J, Han B. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics* **42**, 961-967.

Huang Y. 2001. Rice ideotype breeding of Guangdong Academy of Agricultural Sciences in retrospect. *Guangdong Agricutural Science* **3**, 2-6.

IPCC. 2007. Contribution of working groups I, II and III to the forth assessment report of the intergovernmental panel on climate change. Core Writing Team, Pachauri, R.K., Reisinger, A (Eds.) IPCC, Geneva, Switzerland.

Ishii R. 1995. Cultivar differences. In: Matuso T, Kumazawa K, Ishii R, Ishihara K, Hirata H, eds. Science of the rice plant 2. Physiology. Tokyo: Food and agriculture policy research center, 566–572.

Israelsson M, Siegel RS, Young J, Hashimoto M, Iba K, Schroeder JI. 2006. Guard cell ABA and CO_2 signaling network updates and Ca^{2+} sensor priming hypothesis. *Current Opinion in Plant Biology* **9,** 654-663.

Izawa T. 2007. Adaptation of flowering-time by natural and artificial selection in Arabidopsis and rice. *Journal of Experimental Botany* **58**, 3091-3097.

Jagadish S, Craufurd P, Wheeler T. 2007. High temperature stress and spikelet fertility in rice (*Oryza sativa* L.). *Journal of Experimental Botany* **58:** 1627-1635.

References

Jahn CE, Mckay JK, Mauleon R, Stephens J, McNally K, Bush DR, Leung H, Leach JE. 2011. Genetic variation in biomass traits among 20 diverse rice varieties. *Plant Physiology* **155,** 157-168.

Jansen RC. 1995. Genetic mapping quantitative trait loci in plants: a novel statistical approach. PhD dissertation, Wageningen University, The Netherlands.

Jena K, Mackill D. 2008. Molecular markers and their use in marker-assisted selection in rice. *Crop Science* **48**, 1266-1276.

Ji X, Shiran B, Wan J, Lewis DC, Jenkins CLD, Condon AG, Richards RA, Dolferus R. 2010. Importance of pre-anthesis anther sink strength for maintenance of grain number during reproductive stage water stress in wheat. *Plant, Cell and Environment* **33**, 926-942.

Jiang GH, He YQ, Xu CG, Li XH, Zhang Q. 2004. The genetic basis of stay-green in rice analysed in a population of doubled haploid lines derived from an *indica* by *japonica* cross. *Theoretical and Applied Genetics* **108,** 688–698.

Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X, Qian Q, Li J. 2010. Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nature Genetics* **42**, 541-544.

Jin J, Huang W, Gao JP, Yang J, Shi M, Zhu MZ, Luo D, Lin HX. 2008. Genetic control of rice plant architecture under domestication. *Nature Genetics* **40**, 1365-1369.

Keating BA, Carberry PS, Hammer GL, Probert ME, Robertson MJ, Holzworth D, Huth NI, Hargreaves JNG, Meinke H, Hochman Z, McLean G, Verburg K, Snow V, Dimes JP, Silburn M, Wang E, Brown S, Bristow KL, Asseng S, Chapman S, McCown RL, Freebairn DM, Smith CJ. 2003. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy* **18**, 267-288.

Khush GS. 1995. Breaking the yield frontier of rice. *GeoJournal* **35**, 329-332.

Khush GS. 2005. What it will take to feed 5.0 billion rice consumers in 2030. *Plant Molecular Biology* **59**, 1-6.

Khush GS, Coffman WR, Beachell HM. 2001. The history of rice breeding: IRRI's contribution. In: Rockwood WG, ed. *Rice Research and Production in the 21st Century: Symposium Honoring Robert F. Chandler, Jr.* Los Baños, Philippines: International Rice Research Institute, 117–135.

Kohzuma K, Cruz JA, Akashi K, Hoshiyasu S, Munekage YN, Yokota A, Kramer DM. 2009. The long-term responses of the photosynthetic proton circuit to drought. *Plant, Cell & Environment* **32,** 209-219.

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M. 2006. An SNP caused loss of seed shattering during rice domestication. *Science* **312**, 1392-1396.

Koornneef M, Stam P. 2001. Changing paradigms in plant breeding. *Plant Physiology* **125**, 156-159.

Kramer PJ, Boyer JS. 1995. Water relation of plants and soils. Academic Press, London.

Kropff M, Haverkort A, Aggarwal P, Kooman P. 1995. Using systems approaches to design and evaluate ideotypes for specific environments. In: Bouma J, Kuyvenhoven A, Bouman BAM, Luyten JC, Zandstra HG, eds. *Eco-regional approaches for sustainable land use and food production*. Dordrecht, the Netherlands: Kluwer Academic Publishers, 417-435.

Kumar A, Bernier J, Verulkar S, Lafitte HR, Atlin GN. 2008. Breeding for drought tolerance: Direct selection for yield, response to selection and use of drought-tolerant donors in upland and lowland-adapted populations. *Field Crops Research* **107,** 221-231.

La H. 2004. The QTL mapping of traits related to drought tolerance and genetic transformation of herbicide tolerance in rice (*Oryza sativa* L.). Dissertation, China Agricultural University.

Lafitte HR, Price AH, Courtois B. 2004. Yield response to water deficit in an upland rice mapping population: associations among traits and genetic markers. *Theoretical and Applied Genetics* **109,** 1237–1246.

Lanceras JC, Pantuwan G, Jongdee B, Toojinda T. 2004. Quantitative trait loci associated with drought tolerance at reproductive stage in rice. *Plant Physiology* **135,** 384–399.

Lander E, Kruglyak L. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics* **11,** 241-247.

Lawlor DW. 1995. The effects of water deficit on photosynthesis. In: Smirnoff N (ed) Environment and plant metabolism. Flexibility and acclimation. BIOS Scientific Publishers, Oxford, pp 129–160.

Lawlor DW, Cornic G. 2002. Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. *Plant, Cell and Environment* **25**, 275-294.

Lawson T, Kramer DM, Raines CA. 2012. Improving yield by exploiting mechanisms underlying natural variation of photosynthesis. *Current Opinion in Biotechnology* **23:** 215-220.

References

Lawson T, von Caemmerer S, Baroli I. 2011. Photosynthesis and Stomatal Behaviour. Progress in Botany 72. In: Lüttge U, Beyschlag W, Büdel B, Francis D, eds, Vol. 72: Springer Berlin Heidelberg, 265-304.

Leuning R. 1990. Modelling stomatal behaviour and photosynthesis of Eucalyptus grandis. *Australian Journal of Plant Physiology* **17,** 159–175.

Leuning R. 1995 A critical appraisal of a combined stomatal photosynthesis model for C_3 plants. *Plant, Cell and Environment* **18,** 339–355.

Li C, Zhou A, Sang T. 2006. Rice domestication by reducing shattering. *Science* **311**, 1936-1939.

Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, Wang X, Liu X, Teng S, Hiroshi F. 2003. Control of tillering in rice. *Nature* **422**, 618-621.

Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J. 2011. Natural variation in GS5 plays an important role in regulating grain size and yield in rice. *Nature Genetics* **43**, 1266-1269.

Lin HX, Yamamoto T, Sasaki T, Yano M. 2000. Characterization and detection of epistatic interactions of 3 QTLs, Hd1, Hd2, and Hd3, controlling heading date in rice using nearly isogenic lines. *Theoretical and Applied Genetics* **101,** 1021-1028.

Lin YR, Schertz KF, Paterson AH. 1995. Comparative analysis of QTL affecting plant height and maturity across Poaceae, in reference to an interspecific sorghum population. *Genetics* **141,** 391-411.

Lincoln S, Daly M, Lander E. 1993. MAPMAKER/EXP 3.0 and MAPMAKER/QTL 1.1. Technical report. Whitehead institute of Medical Research, Cambridge, Mass.

Lloyd J, Syvertsen JP, Kriedemann PE, Farquhar GD. 1992. Low conductances for CO_2 diffusion from stomata to the sites of carboxylation in leaves of woody species. *Plant, Cell & Environment* **15**, 873-899.

Long SP, Bernacchi CJ. 2003. Gas exchange measurements, what can they tell us about underling limitations to photosynthesis? Procedures and source error. *Journal of Experimental Botany* **54,** 2393-2401.

Long SP, Zhu XG, Naidu S, Ort DR. 2006. Can improvement in photosynthesis increase crop yields? *Plant, Cell and Environment* **29,** 315-330.

Loomis RS, Rabbinge R, Ng E. 1979. Explanatory models in crop physiology. *Annual Review of Plant Physiology* **30,** 339-367.

Lu C, Zhang J. 1999. Effects of water stress on photosystem II photochemistry and its thermostability in wheat plants. *Journal of Experimental Botany* **50,** 1199-1206.

Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics* **10**, 565-577.

Mae T. 1997. Physiological nitrogen efficiency in rice: Nitrogen utilization, photosynthesis, and yield potential. *Plant and Soil* **196**, 201-210.

Makino A, Mae T, Ohira K. 1984. Changes in photosynthetic capacity in rice leaves from emergence through senescence. Analysis from ribulose-1,5-bisphosphate carboxylase and leaf conductance. *Plant and Cell Physiology* **25,** 511–521.

Makino A, Mae T, Ohira K. 1985. Photosynthesis and ribulose-1,5-bisphosphate carboxylase/oxygenase in rice leaves from emergence through senescence. Quantative analysis by carboxylation/oxygenation and regeneration of ribulose 1,5-bisphosphate. *Planta* **166,** 414–420.

Martens H, Næs T. 1992. *Multivariate calibration*. Chichester, UK: J. Wiley and Sons.

Masle J, Gilmore SR, Farquhar GD. 2005. The *ERECTA* gene regulates plant transpiration efficiency in *Arabidopsis*. *Nature* **436**, 866-870.

Masumoto C, Ishii T, Kataoka S, Hatanaka T, Uchida N. 2004. Enhancement of rice leaf photosynthesis by crossing between cultivated rice, *Oryza sativa* and wild rice species, *Oryza ryfipogon*. *Plant Production Science* **7,** 252–259.

Matsumoto T, Wu J, Kanamori H, Katayose Y, Fujisawa M, Namiki N, Mizuno H, Yamamoto K, Antonio BA, Baba T. 2005. The map-based sequence of the rice genome. *Nature* **436**, 793-800.

Maurino VG, Peterhansel C. 2010. Photorespiration: current status and approaches for metabolic engineering. *Current Opinion in Plant Biology* **13:** 248-255.

Mavromatis T, Boote K, Jones J, Irmak A, Shinde D, Hoogenboom G. 2001. Developing genetic coefficients for crop simulation models with data from crop performance trials. *Crop Science* **41**, 40-51.

McCormick AJ, Cramer MD, Watt DA. 2006. Sink strength regulates photosynthesis in sugarcane. *New Phytologist* **171**, 759-770.

McCouch SR, Cho YG, Yano M, Paul E, Blinstrub M. 1997. Report on QTL nomenclature. *Rice Genetics Newsletter* **14,** 11-13.

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom R, DeClerck G, Schneider D, Cartinhour S, Ware D, Stein L. 2002. Development and Mapping of 2240 New SSR Markers for Rice (*Oryza sativa* L.). *DNA Research* **9,** 199-207.

McCown RL, Hammer GL, Hargreaves JNG, Holzworth DP, Freebairn DM. 1996. APSIM: a novel software system for model development, model testing

and simulation in agricultural systems research. *Agricultural Systems* **50**, 255-271.

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

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

Miflin B. 2000. Crop improvement in the 21st century. *Journal of Experimental Botany* **51**, 1-8.

Mir RR, Zaman-Allah M, Sreenivasulu N, Trethowan R, Varshney RK. 2012. Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. *Theoretical and Applied Genetics* **125**, 625-645.

Mitchell PL, Sheehy JE. 2006. Supercharging rice photosynthesis to increase yield. *New Phytologist* **171:** 688-693.

Miura K, Ashikari M, Matsuoka M. 2011. The role of QTLs in the breeding of high-yielding rice. *Trends in Plant Science* **16:** 319-326.

Miura K, Ikeda M, Matsubara A, Song XJ, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M. 2010. OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nature Genetics* **42**, 545-549.

Miyagawa Y, Tamoi M, Shigeoka S. 2001. Overexpression of a cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. *Nature* Biotechnology **19:** 965-969.

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding* **3**, 87-103.

Money NP. 1989. Osmotic pressure of aqueous polyethylene glycols: relationship between molecular weight and vapor pressure deficit. *Plant Physiology* **91,** 766-769.

Monforte AJ, Tanksley SD. 2000. Fine mapping of a quantitative trait locus (QTL) from *Lycopersicon hirsutum* chromosome 1 affecting fruit characteristics and agronomic traits: breaking linkage among QTLs affecting different traits and dissection of heterosis for yield. *Theoretical and Applied Genetics* **100,** 471-479.

Mueller-Cajar O, Whitney SM. 2008. Directing the evolution of Rubisco and Rubisco activase: first impressions of a new tool for photosynthesis research. *Photosynthesis Research* **98:** 667-675.

Murchie EH, Pinto M, Horton P. 2009. Agriculture and the new challenges for photosynthesis research. *New Phytologist* **181,** 532-552.

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. *Theoretical and Applied Genetics* **110:** 778-786.

Næs T, Mevik BH. 2001. Understanding the collinearity problem in regression and discriminant analysis. *Journal of Chemometrics* **15**, 413-426.

Nguyen HT, Chandra BR, Blum A. 1997. Breeding for drought resistance in rice: physiology and molecular genetics considerations. *Crop Science* **37,** 1426–1434.

Niinemets Ü, Díaz-Espejo A, Flexas J, Galmés J, Warren CR. 2009. Role of mesophyll diffusion conductance in constraining potential photosynthetic productivity in the field. *Journal of Experimental Botany* **60,** 2249-2270.

Ögren E, Evans JR. 1993. Photosynthetic light-response curves. *Planta* **189,** 182-190.

Ookawa T, Hobo T, Yano M, Murata K, Ando T, Miura H, Asano K, Ochiai Y, Ikeda M, Nishitani R. 2010. New approach for rice improvement using a pleiotropic QTL gene for lodging resistance and yield. *Nature communications* **1**, 132.

Ott RL, Longnecker M. 2001. An introduction to statistical methods and data analysis (fifth edition). Duxbury, California.

Oxborough K, Baker NR. 1997. Resolving chlorophyll *a* fluorescence image of photosynthetic efficiency into photochemical and non-photochemical component — calculation of qP and $F'_{\mathrm{v}}/F'_{\mathrm{m}}$ without measuring F'_{o}. *Photosynthesis Research* **54,** 135-142.

Panaud O, Chen X, McCouch SR. 1996. Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Molecular and General Gentics* **252,** 597-607.

Parry MAJ, Reynolds M, Salvucci ME, Raines C, Andralojc PJ, Zhu X-G, Price GD, Condon AG, Furbank RT. 2011. Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *Journal of Experimental Botany* **62:** 453-467.

Peng S, Cassman KG, Virmani SS, Sheehy J, Khush GS. 1999. Yield potential trends of tropical rice since the release of IR8 and the challenge of increasing rice yield potential. *Crop Science* **39**, 1552–1559.

Peng S, Khush GS, Virk P, Tang Q, Zou Y. 2008. Progress in ideotype breeding to increase rice yield potential. *Field Crops Research* **108:** 32-38.

Peng S, Laza RC, Khush GS, Sanico AL, Visperas RM, Garcia FV. 1998. Transpiration efficiencies of indica and improved tropical japonica rice grown under irrigated conditions. *Euphytica* **103,** 103–108.

Penning de Vries FWT. 1991. Improving yields: designing and testing VHYVs [very high-yielding varieties]. In: Penning de Vries FWT, Kropff MJ, Teng PS, Kirk GJD, eds. *Systems simulation at IRRI. IRRI Research Paper*, Vol. 151. Los Baños, Philippines: IRRI, 13-19.

Peterhansel C, Niessen M, Kebeish RM. 2008. Metabolic Engineering Towards the Enhancement of Photosynthesis. *Photochemistry and Photobiology* **84:** 1317-1323.

Pinheiro C, Chaves MM. 2011. Photosynthesis and drought: can we make metabolic connections from available data? *Journal of Experimental Botany* **62,** 869-882.

Pons TL, Pearcy RW. 1994. Nitrogen reallocation and photosynthetic acclimation in response to partial shading in soybean plants. *Physiologia Plantarum* **92,** 636-644.

Poorter H, Niinemets Ü, Poorter L, Wright IJ, Villar R. 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytologist* **182,** 565-588.

Price AH, Cairns JE, Horton P, Jones HG, Griffiths H. 2002. Linking drought-resistance mechanisms to drought avoidance in upland rice using a QTL approach: progress and new opportunities to integrate stomatal and mesophyll responses. *Journal of Experimental Botany* **53,** 989-1004.

Price AH, Steele KA, Moore BJ, Barraclough PB, Clark LJ. 2000. A combined RFLP and AFLP linkage map of upland rice (*Oryza sativa* L.) used to identify QTLs for root-penetration ability. *Theoretical and Applied Genetics* **100,** 49-56.

Price GD, Badger MR, Woodger FJ, Long BM. 2008. Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): functional components, C_i transporters, diversity, genetic regulation and prospects for engineering into plants. *Journal of Experimental Botany* **59:** 1441-1461.

Prudent M, Lecomte A, Bouchet JP, Bertin N, Causse M, Genard M. 2011. Combining ecophysiological modelling and quantitative trait locus analysis to identify key elementary processes underlying tomato fruit sugar concentration. *Journal of Experimental Botany* **62:** 907-919.

Quarrie SA, Quarrie SP, Radosevic R, Rancic D, Kaminska A, Barnes JD, Leverington M, Ceoloni C, Dodig D. 2006. Dissecting a wheat QTL for yield present in a range of environments: from the QTL to candidate genes. *Journal of Experimental Botany* **57,** 2627-2637.

Quick WP, Horton P. 1984. Studies on the induction of chlorophyll fluorescence in barley protoplasts. II Resolution of fluorescence quenching by redox state and the transthylakoid pH gradient. *Proceedings or the Royal Society B* **220,** 371–382.

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. *Journal of Experimental Botany* **56**, 3083-3092.

Quilot B, Wu BH, Kervella J, Génard M, Foulongne M, Moreau K. 2004. QTL analysis of quality traits in an advanced backcross between Prunus persica cultivars and the wild relative species P. davidiana. *Theoretical and Applied Genetics* **109**, 884-897.

Rebetzke GJ, A. G. Condon RAR, Farquhar GD. 2002. Selection for reduced carbon isotope discrimination increases aerial biomass and grain yield of rainfed bread wheat. *Crop Science* **42**, 739-745.

Reddy AR. 1996. Fructose 2,6-bisphosphate-modulated photosynthesis in sorghum leaves grown under low water regimes. *Phytochemistry* **43**, 19-22.

Ren ZH, Gao JP, Li LG, Cai XL, Huang W, Chao DY, Zhu MZ, Wang ZY, Luan S, Lin HX. 2005. A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nature Genetics* **37**, 1141-1146.

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

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

Richards RA. 2000. Selectable traits to increase crop photosynthesis and yield of grain crops. *Journal of Experimental Botany* **51:** 447-458.

Robin S, Pathan MS, Courtois B, Lafitte R, Carandang S, Lanceras S, Amante M, Nguyen HT, Li Z. 2003. Mapping osmotic adjustment in an advanced back-cross inbred population of rice. *Theoretical and Applied Genetics* **107**, 1288–1296.

Rogers OS, Bendich AJ. 1985. Extraction of DNA from milligram amounts of fresh, *herbaruium* and mummified plant tissues. *Plant Molecular Biology* **5,** 69-76.

Rott M, Martins NF, Thiele W, Lein W, Bock R, Kramer DM, Schöttler MA. 2011. ATP synthase repression in tobacco restricts photosynthetic electron

transport, CO_2 assimilation, and plant growth by overacidification of the thylakoid lumen. *The Plant Cell* **23:** 304-321.

Salekdeh GH, Reynolds M, Bennett J, Boyer J. 2009. Conceptual framework for drought phenotyping during molecular breeding. *Trends in Plant Science* **14**, 488-496.

Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush G. 2002. A mutant gibberellin-synthesis gene in rice. *Nature* **416**, 701-702.

Sasaki H, Ishii R. 1992. Cultivar differences in leaf photosynthesis of rice bred in Japan. *Photosynthesis Research* **32**, 139-146.

Scafaro AP, von Caemmerer S, Evans JR, Atwell BJ. 2011. Temperature response of mesophyll conductance in cultivated and wild *Oryza* species with contrasting mesophyll cell wall thickness. *Plant, Cell & Environment* **34,** 1999-2008.

Serraj R, Dimayuga D, Gowda V, Guan Y, Hong He, Impa S, Liu DC, Mabesa RC, Sellamuthu R, Torres R. 2008. Drought resistant rice: physiological framework for an integrated research strategy. In: Serraj R, Bennett J, Hardy B, eds. Drought frontiers in rice: crop improvement for increased rainfed production. Singapore: World Scientific Publishing, Los Baňos, Philippines: International Rice Research Institute, 139–170.

Setter T, Conocono E, Egdane J, Kropff M. 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. *Australian Journal of Plant Physiology* **22**, 441-451.

Shappley ZW, Jenkins JN, Zhu J, McCarty JC. 1998. Quantitative trait loci associated with yield and fiver traits of upland cotton. *Journal of Cotton Science* **4,** 153-163.

Sharma PK, De Datta SK. 1994. Rain water utilization efficiency in rain-fed lowland rice. *Advance in Agronomy* **52,** 85–120.

Shen YJ, Jiang H, Jin JP, Zhang ZB, Xi B, He YY, Wang G, Wang C, Qian L, Li X, Yu QB, Liu HJ, Chen DH, Gao JH, Huang H, Shi TL, Yang ZN. 2004. Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. *Plant physiology* **135,** 1198-1205.

Shiratsuchi H, Yamagishi T, Ishii R. 2006. Leaf nitrogen distribution to maximize the canopy photosynthesis in rice. *Field Crops Research* **95,** 291–304.

Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M. 2008. Deletion in a gene associated with grain size increased yields during rice domestication. *Nature Genetics* **40**, 1023-1028.

Shorter R, Lawn RJ, Hammer GL. 1991. Improving genotypic adaptation in crops – a role for breeders, physiologists and modellers. *Experimental Agriculture* **27**, 155-175.

Simko I, Vreugdenhil D, Jung CS, May GD. 1999. Similarity of QTLs detected for in vitro and greenhouse development of potato plants. *Molecular Breeding* **5,** 417–428.

Sinclair TR. 2011. Challenges in breeding for yield increase for drought. *Trends in Plant Science* **16**, 289-293.

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

Singh U, Ladha JK, Castillo EG, Punzalan G, Tirol-Padre A, Duqueza M. 1998. Genotypic variation in nitrogen use efficiency in medium- and long-duration rice. *Field Crops Research* **58:** 35-53.

Song X, Huang W, Shi M, Zhu M, Lin H-X. 2007. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature Genetics* **39**, 623-630.

Stam P. 1998. Crop physiology, QTL analysis and plant breeding. In: Lambers H, Poorter H, van Vuuren MMI, eds. *Inherent variation in plant growth: physiological mechanisms and ecological consequences*. Leiden, Netherlands: Backhuys Publishers, 429-440.

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

Takahashi Y, Shomura A, Sasaki T, Yano M. 2001. Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the α subunit of protein kinase CK2. *Proceedings of the National Academy of Sciences* **98**, 7922-7927.

Takai T, Ohsumi A, San-oh Y, Laza MRC, Kondo M, Yamamoto T, Yano M. 2009. Detection of a quantitative trait locus controlling carbon isotope discrimination and its contribution to stomatal conductance in *japonica* rice. *Theoretical and Applied Genetics* **118,** 1401-1410.

Takeda S, Matsuoka M. 2008. Genetic approaches to crop improvement: responding to environmental and population changes. *Nature Reviews Genetics* **9**, 444-457.

Takeda T, Suwa Y, Suzuki M, Kitano H, Ueguchi-Tanaka M, Ashikari M, Matsuoka M, Ueguchi C. 2003. The OsTB1 gene negatively regulates lateral branching in rice. *The Plant Journal* **33**, 513-520.

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science* **316**, 1033-1036.

References

Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, Fu Y, Cai H, Wang X, Xie D. 2008. Control of a key transition from prostrate to erect growth in rice domestication. *Nature Genetics* **40**, 1360-1364.

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

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

Tardieu F. 2011. Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. *Journal of Experimental Botany* **63**, 25-31.

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

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. *Journal of Experimental Botany* **62,** 4067-4077.

Tazoe Y, von Caemmerer S, Badger MR, Evans JR. 2009. Light and CO_2 do not affect the mesophyll conductance to CO_2 diffusion in wheat leaves. *Journal of Experimental Botany* **60,** 2291-2301.

Tazoe Y, von Caemmerer S, Estavillo GM, Evans JR. 2011.Using tunable diode laser spectroscopy to measure carbon isotope discrimination and mesophyll conductance to CO_2 diffusion dynamically at different CO_2 concentrations. *Plant, Cell & Environment* **34**, 580-591.

Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* **100,** 697-712.

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.

Tezara W, Mitchell VJ, Driscoll SD, Lawlor DW. 1999. Water stress inhibits plant photosynthesis by decreasing coupling factor and ATP. *Nature* **401**, 914-917.

Tuberosa R. 2012. Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3**, 347.

Tuong T, Bouman B. 2003. Rice production in water-scarce environments. In: Kijne J, Barker R, Molden D, eds. *Water productivity in agriculture: Limits and opportunities for improvement*: CABI Publishing, UK, 53-67.

Tuzet A, Perrier A, Leuning R. 2003. A coupled model of stomatal conductance, photosynthesis and transpiration. *Plant, Cell and Environment* **26**, 1097-1116.

Uehlein N, Otto B, Hanson DT, Fischer M, McDowell N, Kaldenhoff R. 2008. Function of Nicotiana tabacum aquaporins as chloroplast gas pores challenges the concept of membrane CO_2 permeability. *The Plant Cell* **20:** 648-657.

Uga Y, Okuno K, Yano M. 2011. *Dro1*, a major QTL involved in deep rooting of rice under upland field conditions. *Journal of Experimental Botany* **62**, 2485-2494.

United Nations. 2011. World population prospects, the 2010 revision. Department of Economic and Social Affairs, Population Division, Population Estimates and Projections Section. Rome, United Nations.

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*. *Molecular Breeding* **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 Biology* **14,** 179-189.

van Berloo R. 2008. GGT2.0: versatile software for visualization and analysis of genetic data. *Journal of Heredity* **99,** 232-236.

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. *Australian Journal of Agricultural Research* **56**, 883-894.

van Ooijen JW. 2009. MapQTL® 6. Software for the mapping of quantitative trait loci in experimental populations of diploid species. Kyazma BV: Wageningen, Netherlands.

Venuprasad R, Lafitte HR, Atlin GN. 2007. Response to direct selection for grain yield under drought stress in rice. *Crop Science* **47,** 285-293.

Venuprasad R, Sta Cruz MT, Amante M, Magbanua R, Kumar A, Atlin GN. 2008. Response to two cycles of divergent selection for grain yield under drought stress in four rice breeding populations. *Field Crops Research* **107,** 232-244.

von Caemmerer S. 2000. Biochemical Models of Leaf Photosynthesis. CSIRO Publishing, Australia.

von Caemmerer S, Quick WP, Furbank RT. 2012. The development of C_4 rice: current progress and future challenges. *Science* **336:** 1671-1672.

Voorrips RE. 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity* **93,** 77-78.

Wallach D, Makowski D, Jones JW. 2006. *Working with dynamic crop models: Evaluation, analysis, parameterization, and applications*. Amsterdam, The Netherlands: Elsevier.

Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H, Ma H, Zhang G, He Z. 2008. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nature Genetics* **40**, 1370-1374.

Wang YP, Leuning R. 1998. A two-leaf model for canopy conductance, photosynthesis and partitioning of available energy I: Model description and comparison with a multi-layered model. *Agricultural and Forest Meteorology* **91:** 89-111.

Warren CR. 2004. The photosynthetic limitation posed by internal conductance to CO_2 movement is increased by nutrient supply. *Journal of Experimental Botany* **55,** 2313-2321.

Warren CR. 2008. Stand aside stomata, another actor deserves centre stage: the forgotten role of the internal conductance to CO_2 transfer. *Journal of Experimental Botany* **59**, 1475–1487.

Warren CR, Ethier GJ, Livingston NJ, Grant NJ, Turpin DH, Harrison DL, Black TA. 2003. Transfer conductance in second growth Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) canopies. *Plant, Cell & Environment* **26,** 1215-1227.

Wei X, Xu J, Guo H, Jiang L, Chen S, Yu C, Zhou Z, Hu P, Zhai H, Wan J. 2010. DTH8 suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiology* **153**, 1747-1758.

Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X, Cheng Z, Guo X, Wang J, Jiang L, Zhai H, Wan J. 2008. Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. *Cell Research* **18**, 1199-1209.

Westoby M, Falster DS, Moles AT, Vesk PA, Wright IJ. 2002. Plant ecological strategies: some leading dimensions of variation between species. *Annual Review of Ecology and Systematics* **33,** 125-159.

Whisler F, Acock B, Baker D, Fye R, Hodges H, Lambert J, Lemmon H, McKinion J, Reddy V. 1986. Crop Simulation Models in Agronomic Systems. *Advances in Agronomy* **40**, 141-208.

White JW, Hoogenboom G. 1996. Simulating effects of genes for physiological traits in a process-oriented crop model. *Agronomy Journal* **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 Science* **48:** 678-687.

Whitney SM, Sharwood RE. 2008. Construction of a tobacco master line to improve Rubisco engineering in chloroplasts. *Journal of Experimental Botany* **59:** 1909-1921.

Wu C, You C, Li C, Long T, Chen G, Byrne ME, Zhang Q. 2008. RID1, encoding a Cys2/His2-type zinc finger transcription factor, acts as a master switch from vegetative to floral development in rice. *Proceedings of the National Academy of Sciences* **105**, 12915-12920.

Wu YQ, Huang Y. 2007. An SSR genetic map of *Sorghum bicolor* (L.) Moench and its comparison to a published genetic map. *Genome* **50,** 84-89.

Xiao J, Li J, Yuan L, Tanksley SD. 1996. Identification of QTLs affecting traits of agronomic importance in a recombinant inbred population derived from a subspecific rice cross. Theoretical and Applied Genetics **92,** 230-244.

Xing Y, Zhang Q. 2010. Genetic and molecular bases of rice yield. *Annual Review of Plant Biology* **61:** 421-442.

Xing Y, Tang W, Xue W, Xu C, Zhang Q. 2008. Fine mapping of a major quantitative trait loci, *qSSP7*, controlling the number of spikelets per panicle as a single Mendelian factor in rice. *Theoretical and Applied Genetics* **116:** 789-796.

Xu JL, Lafitte, HR, Gao YM, Fu BY, Torres R, Li ZK. 2005. QTLs for drought escape and tolerance identified in a set of random introgression lines of rice. *Theoretical Applied Genetics* **111,** 642–1650.

Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ. 2006. Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* **442**, 705-708.

Xu L, Henke M, Zhu J, Kurth W, Buck-Sorlin G. 2011. A functional–structural model of rice linking quantitative genetic information with morphological development and physiological processes. *Annals of Botany* **107:** 817-828.

Xu Y, Li HN, Li GJ, Wang X, Cheng LG, Zhang YM. 2011. Mapping quantitative trait loci for seed size traits in soybean (*Glycine max* L. Merr.). *Theoretical and Applied Genetics* **122,** 581-594.

Xu Y, This D, Pausch RC, Vonhof WM, Coburn JR, Comstock JP, McCouch SR. 2009. Leaf-level water use efficiency determined by carbon isotope discrimination in rice seedlings: genetic variation associated with population structure and QTL mapping. *Theoretical and Applied Genetics* **118,** 1065-1081.

Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X, Zhang Q. 2008. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nature Genetics* **40**, 761-767.

References

Yang J, Zhao X, Cheng K, Du H, Ouyang Y, Chen J, Qiu S, Huang J, Jiang Y, Jiang L. 2012. A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science* **337**, 1336-1340.

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki aT. 2000. Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. *The Plant Cell* **12**, 2473–2483.

Yeo ME, Yeo AR, Flowers TJ. 1994. Photosynthesis and photorespiration in the genus *Oryza*. *Journal of Experimental Botany* **45,** 553–560.

Yin X. 2013. Improving ecophysiological simulation models to predict the impact of elevated atmospheric CO_2 concentration on crop productivity. *Annals of Botany*, in press (DOI: 10.1093/aob/mct016).

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 Phytologist* **179**, 629-642.

Yin X, Struik PC. 2009a. C_3 and C_4 photosynthesis models: An overview from the perspective of crop modelling. *NJAS-Wageningen Journal of Life Science* **57:** 27-38.

Yin X, Struik PC. 2009b. Theoretical reconsiderations when estimating the mesophyll conductance to CO_2 diffusion in leaves of C_3 plants by analysis of combined gas exchange and chlorophyll fluorescence measurements. *Plant, Cell & Environment* **32,** 1513–1524. (with corrigendum in *Plant, Cell & Environment* **33,** 1595).

Yin X, Struik PC. 2010. Modelling the crop: from system dynamics to systems biology. *Journal of Experimental Botany* **61**, 2171-2183.

Yin X, Struik PC. 2011. The role of QTL-based physiological modelling in crop improvement. *In* Proceedings of the Conference on Systems approaches to crop improvement, 14-15 April 2011, Harpenden, UK, Aspects of Applied Biologists, Systems approaches to crop improvement, **107**, 1-8.

Yin X, van Laar HH. 2005. Crop systems dynamics: an ecophysiological simulation model for genotype-by-environment interactions. Wagningen Academic Publishers, Wageningen, The Netherlands.

Yin X, Harbinson J, Struik PC. 2006. Mathematical review of literature to assess alternative electron transports and interphotosystem excitation partitioning of steady-state C_3 photosynthesis under limiting light. *Plant, Cell and Environment* **29,** 1771-1782.

Yin X, Harbinson J, Struik PC. 2009a. A model of the generalized stoichiometry of electron transport-limited C_3 photosynthesis. In: Laisk A, Nedbal L, Govindjee

(ed) Photosynthesis in silico: Understanding Complexity from Molecules to Ecosystems, vol. 29, Book series 'Advances in Photosynthesis and Respiration'. Springer, The Netherlands, pp 247–273.

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

Yin X, Kropff MJ, Goudriaan J, Stam P. 2000a. A model analysis of yield differences among recombinant inbred lines in barley. *Agronomy Journal* **92**, 114-120.

Yin X, Kropff MJ, McLaren G, Visperas RM. 1995. A nonlinear model for crop development as a function of temperature. *Agricultural and Forest Meteorology* **77:** 1-16.

Yin X, Stam P, Dourleijn CJ, Kropff MJ. 1999b. AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. *Theoretical and Applied Genetics* **99**, 244-253.

Yin X, Stam P, Kropff MJ, Schapendonk AHCM. 2003a. Crop modeling, QTL mapping, and their complementary role in plant breeding. *Agronomy Journal* **95**, 90-98.

Yin X, Sun Z, Struik PC, Gu J. 2011. Evaluating a new method to estimate the rate of leaf respiration in the light by analysis of combined gas exchange and chlorophyll fluorescence measurements. *Journal of Experimental Botany* **62,** 3489-3499.

Yin X, Chasalow SD, Dourleijn CJ, Stam P, Kropff MJ. 2000b. 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, Goudriaan J, Lantinga EA, Vos J, Spiertz HJ. 2003b. A flexible sigmoid function of determinate growth. *Annals of Botany* **91:** 361-371.

Yin X, Schapendonk AHCM, Kropff MJ, van Oijen M, Bindraban PS. 2000c. A generic equation for nitrogen-limited leaf area index and its application in crop growth models for predicting leaf senescence. *Annals of Botany* **85:** 579-585.

Yin X, Struik PC, Tang J, Qi C, Liu T. 2005a. Model analysis of flowering phenology in recombinant inbred lines of barley. *Journal of Experimental Botany* **56**, 959-965.

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

References

Yin X, Chasalow SD, 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 Breeding* **121,** 314-319.

Yin X, Struik PC, Romero P, Harbinson J, Evers JB, van der Putten PEL, Vos J. 2009b. Using combined measurements of gas exchange and chlorophyll fluorescence to estimate parameters of a biochemical C_3 photosynthesis model: a critical appraisal and a new integrated approach applied to leaves in a wheat (*Triticum aestivum*) canopy. *Plant, Cell and Environment* **32**, 448-464.

Yu J, Hu S, Wang J, Wong GKS, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X. 2002. A draft sequence of the rice genome (Oryza sativa L. ssp. indica). *Science* **296**, 79-92.

Yuan L, Yang Z, Yang J. 1994. Hybrid rice in China. In: Virmani SS, ed. *Hybrid Rice Technology: New Developments and Future Prospects*. Los Baños, Philippines: International Rice Research Institute, 143–147.

Yue B, Xiong L, Xue W, Xing Y, Luo L, Xu C. 2005. Genetic analysis for drought resistance of rice at reproductive stage in field with different types of soil. *Theoretical and Applied Genetics* **111,** 1127-1136.

Zhang C. 2006. The genetic analysis and the QTL mapping of the major agronomic traits in RIL between lowland and upland ric (*Oryza Sativa* L.). Dissertation. China Agricultural University, Beijing, China.

Zhao X, Xu J, Zhao M, Lafitte R, Zhu L, Fu B, Gao Y, Li Z. 2008. QTLs affecting morph-physiological traits related to drought tolerance detected in overlapping introgression lines of rice (*Oryza sativa* L.). *Plant Science* **174,** 618-625.

Zhu X-G, Long SP, Ort DR. 2010. Improving photosynthetic efficiency for great yield. *Annual Review of Plant Biology* **61,** 235-261.

Zhu X-G, Portis AR, Long SP. 2004a. Would transformation of C_3 crop plants with foreign Rubisco increase productivity? A computational analysis extrapolating from kinetic properties to canopy photosynthesis. *Plant, Cell & Environment* **27:** 155-165.

Zhu X-G, Ort DR, Whitmarsh J, Long SP. 2004b. The slow reversibility of photosystem II thermal energy dissipation on transfer from high to low light may cause large losses in carbon gain by crop canopies: a theoretical analysis. *Journal of Experimental Botany* **55:** 1167-1175.

Zhu X-G, Zhang G, Tholen D, Wang Y, Xin C, Song Q. 2011. The next generation models for crops and agro-ecosystems. *Science China Information Science* 54:589-597.

Zou J, Zhang S, Zhang W, Li G, Chen Z, Zhai W, Zhao X, Pan X, Xie Q, Zhu L. 2006. The rice *HIGH-TILLERING DWARF1* encoding an ortholog of

Arabidopsis MAX3 is required for negative regulation of the outgrowth of axillary buds. *The Plant Journal* **48**, 687-698.

Summary

Improving grain yield of rice (*Oryza sativa* L.) under both favourable and stressful conditions is the main breeding objective for this crop to ensure food security. Crop growth models based on solid crop-physiological knowledge have long been used to support field research in agriculture. But their applications in exploring genetic variation in e.g. rice germplasm, designing ideotypes, and supporting plant breeding are still limited. Recently, attempts to amalgamate physiological and genetic approaches, whereby quantitative trait loci (QTLs) information has been incorporated into crop models, have resulted in a so-called QTL-based modelling approach. This integrated approach has shown the potential to narrow genotype-phenotype gaps and to resolve genotype × environment interactions. This approach has been proven to be robust in predicting genetic differences in comparatively simple traits, such as leaf elongation rate in maize and flowering time of barley or rice in bi-parental crossing populations under different conditions (in terms of vapour pressure deficit, soil moisture content, temperature and photoperiod). Only in a few cases was the QTL-based modelling approach used to predict complex traits such as yield, and it was less successful. To the best of my knowledge, this approach has not been applied to analyse leaf photosynthesis, certainly not to photosynthesis under drought stress. Furthermore, potential contributions of exploiting natural variation in a genetic population to crop productivity have hardly been quantified. As outlined in Chapter 1, I have performed QTL-based modelling for this thesis research, with a focus on photosynthesis, to develop an efficient marker-assisted strategy for improving grain yield of rice under both favourable and stressful conditions.

Because of the importance of photosynthesis and its sensitivity to drought, I first studied the genetic variation in photosynthesis in a backcross introgression line population developed from a cross between lowland cv. Shennong 265 (sensitive) and upland rice cv. Haogelao (tolerant) (Chapter 2). Gas exchange and chlorophyll fluorescence data were collected under a saturating light condition for the two parents and 94 of their introgression lines (ILs) under drought and well-watered conditions at both flowering and grain filling. Photosynthesis was greatly influenced by the microclimate which unavoidably fluctuates under field conditions: the measured light-saturated photosynthesis fluctuated more than three-fold. This hampers QTL analysis of photosynthesis measured under field conditions. In this study, both a statistical covariant model and a physiological approach were used to standardize the observations. Both approaches identified leaf-to-air vapour pressure difference as the most important factor influencing photosynthesis. After correcting for microclimate fluctuations, significant genetic variation was found in this population, and 1-3

quantitative trait loci (QTLs) were detected per photosynthesis-related trait. A major QTL was mapped near marker RM410 (the interval from 57.3 to 68.4 cM on Chromosome 9) and was consistent for phenotyping at flowering and grain filling, and under drought and well-watered conditions. This QTL consistency was also verified in a greenhouse experiment under controlled conditions. These results might imply that photosynthesis at different phenological stages and under different environmental conditions is influenced by the same genetic factors. These results also provided information for selecting genotypes for a detailed physiological modelling study.

Based on the QTLs for leaf photosynthesis detected in Chapter 2, 13 ILs were carefully selected as representatives of the population to study the physiological basis of genetic variation in leaf photosynthesis and resulting QTLs in ILs (Chapter 3). Measurements of gas exchange and chlorophyll fluorescence were simultaneously conducted at various levels of incident irradiance and of CO_2 to assess the CO_2 and light response curves under drought and well-watered conditions, at flowering and grain filling. Through curve fitting, seven parameters of a photosynthesis model were estimated for each IL, which dissected photosynthesis into stomatal conductance (g_s), mesophyll conductance (g_m), electron transport capacity (J_{max}), and Rubisco carboxylation capacity (V_{cmax}). Although drought and leaf age accounted for the larger proportions of the total variation, significant genetic variation was also found in these parameters. Genetic variation in light saturated photosynthesis, transpiration efficiency (TE), and the major QTL of photosynthesis on Chromosome 9 were mainly associated with variation in g_s and g_m. So, g_s and g_m, which were demonstrated in the literature to be responsible for environmental variation in photosynthesis, were found also to be associated with genetic variation in photosynthesis. Furthermore, relationships between these parameters and leaf nitrogen or dry matter per unit area, which were previously found across environmental treatments, were shown valid for variation across genotypes. In view of these results and literature reports, it was argued that variation in photosynthesis due to environmental conditions and to genetic variation shares common physiological mechanisms. Based on these results from ecophysiological photosynthesis modelling and QTL analysis, ideotypes for leaf photosynthesis and TE were designed, showing 17.0% and 25.1% improvement, respectively, when compared with the best genotype investigated. This analysis also highlights possibilities to improve both photosynthesis and TE simultaneously within the same genetic background.

As rice production is not only determined by photosynthesis, but also by other physiological processes. Physiological traits like individual seed dry weight (S_W), seed N concentration (n_{SO}), maximum plant height (H_{max}), minimum days for vegetative growth phase (m_V), minimum days for reproductive phase (m_R), specific leaf area (S_{la}),

and total crop N uptake (N_{max}) were measured (Chapter 4). The mechanistic crop model GECROS (Genotype-by-Environment interaction on CROp growth Simulator) was used to integrate these physiological component traits and to predict crop yield of individual ILs under well-watered and drought conditions. With measured physiological trait inputs, the GECROS model could account for 72% and 57% of the variation in yield under well-watered and drought conditions, respectively. QTL analysis was performed to these physiological model-input parameters, and molecular marker-based estimates of these traits were calculated from estimated additive allele effects and QTL allelic information of each IL. With marker-based estimates of model inputs replacing the original measured inputs, this QTL/marker-based crop model accounted for 52% and 47% of the variation in yield under well-watered and drought conditions, respectively. Simple correlation and multiple regression analyses showed N_{max} had the most significant effect on yield; five other genotype-specific model-input traits also significantly influenced yield, but S_W did not. Using the marker-based estimates of physiological input parameters, GECROS also gave a fair prediction of variation in yield within a population of 251 recombinant inbred lines of the same parents grown under either well-watered or drought-stressed conditions. Model-based sensitivity analysis provided a tool to rank the relative importance of the identified markers in determining yield, and detected more markers than marker selection using multiple regression for yield per se. In our analysis, markers RM8030 on Chromosome 2 and RM338 on Chromosome 3 were found most important for well-watered and drought-stressed environments, respectively, suggesting that the priority markers for selection to improve grain yield should be environment specific. All these suggest that a QTL/marker based modelling approach might improve the efficiency of marker-assisted selection.

Chapter 4 also showed that leaf photosynthesis was not important in determining the differences in crop yield among the ILs observed in the field experiment, either under drought or under well-watered conditions. This lack of persistence of variation across scales is probably due to the complex hierarchy from leaf-level photosynthesis to crop yield and to interaction and feedback mechanisms occurring between physiological components within the individual plant, between plants of the same crop and between the crop and the environment. Moreover, measured leaf photosynthesis and crop yields are both associated with some random experimental error. Together, the complexities and the experimental noise might mask the potential contribution of the small within-population variation in leaf photosynthesis to the variation in final crop yield. To examine the extent to which natural genetic variation in photosynthesis can contribute to increasing biomass production and yield of rice, the GECROS crop model was used again to analyse the impact of genetic variation in leaf photosynthesis

on crop biomass production (Chapter 5). This was performed by fixing each of the other model-input parameters at a constant value (i.e. the population mean) across the ILs. It was shown that a genetic variation in photosynthesis of 25% can be scaled up equally to crop level, resulting in an increase in biomass of 22-29% across different locations and years. This was in contrast to earlier studies where the percentage of improvement decreases when moving up from leaf to crop level. The difference with earlier studies seems related to the fact that variation in both Rubisco-limited and electron transport-limited photosynthesis was observed in our IL population. Rice production could be significantly improved by mining the natural variation in existing germplasm, especially the variation for parameters that determine light-limited photosynthesis.

In Chapter 6, I summarized and discussed the results presented in this thesis. I evaluated how physiological modelling and QTL/marker-based modelling can assist crop breeding and genetics. Finally, I discussed the future prospects of integrating crop physiology, crop modelling, plant genetics and plant breeding, in the context of 'crop systems biology'.

Samenvatting

Het verbeteren van de korrelopbrengst van rijst (*Oryza sativa* L.), zowel onder gunstige als onder stressvolle omstandigheden, is het belangrijkste veredelingsdoel bij dit gewas. Zo kan de voedselzekerheid gegarandeerd worden. Al sinds lange tijd worden gewasgroeimodellen gebruikt ter ondersteuning van landbouwkundig veldonderzoek. Deze modellen zijn gebaseerd op solide gewasfysiologische kennis. Ze worden echter nog slechts mondjesmaat toegepast bij het verkennen van bijvoorbeeld genetische variatie in rijst, het ontwerpen van ideale planten en het ondersteunen van plantenveredeling. Recentelijk werden modellen ontwikkeld waarin getracht werd fysiologische en genetische benaderingen te combineren. Informatie betreffende zogenaamde loci voor kwantitatieve eigenschapen (in het Engels: quantitative trait loci, afgekort QTLs) werd in gewasmodellen ingebouwd en op deze manier ontstond de zogenaamde QTL-gebaseerde modelbenadering. Met deze geïntegreerde aanpak bleek het mogelijk om de kloof tussen genotype en fenotype te verkleinen en de interacties tussen genotype en omgeving mechanistisch te benaderen. Inmiddels is gebleken dat deze aanpak op robuuste wijze genetische verschillen in relatief eenvoudige eigenschappen kan voorspellen. Voorbeelden van dergelijke eigenschappen zijn snelheid van bladstrekking bij maïs en bloeitijd van gerst of rijst in twee-ouder kruisingspopulaties onder verschillende omstandigheden (in termen van dampspanningstekort, bodemvocht, temperatuur en daglengte). Slechts in enkele gevallen werd de QTL-gebaseerde modelbenadering gebruikt om complexe eigenschappen zoals opbrengst te voorspellen; deze modellen waren daarin ook minder succesvol. Bij mijn weten is deze benadering niet eerder toegepast om bladfotosynthese te analyseren, zeker niet de fotosynthese onder droogtestress. Bovendien zijn de potentiële bijdragen van het exploiteren van natuurlijke variatie in een genetische populatie aan het verhogen van de gewasproductie nauwelijks gekwantificeerd. Zoals uiteengezet in Hoofdstuk 1, heb ik in dit promotieonderzoek QTL-gebaseerde modellen ontwikkeld en benut om fotosynthese te onderzoeken en daarmee een efficiënte merker- ondersteunde strategie voor het verbeteren van de korrelopbrengst van rijst te ontwikkelen, zowel onder gunstige als onder stressvolle omstandigheden.

Vanwege het belang van de fotosynthese en zijn gevoeligheid voor droogte, heb ik allereerst onderzocht hoe groot de genetische variatie in fotosynthese was in een terugkruisingspopulatie van introgressielijnen. Deze populatie was ontwikkeld uit een kruising tussen het voor natte teelt geschikte ras Shennong 265 (gevoelig voor droogte) en het voor droge teelt geschikte ras Haogelao (droogtetolerant) (Hoofdstuk 2). Metingen aan gasuitwisseling en chlorofylfluorescentie werden uitgevoerd onder

verzadigende lichtomstandigheden voor de twee ouders en 94 van hun introgressielijnen (ILs) onder omstandigheden van droogte en van voldoende water, zowel tijdens de bloei als tijdens de korrelvulling. Fotosynthese werd sterk beïnvloed door het microklimaat en dat fluctueert onvermijdelijk onder veldomstandigheden: de gemeten lichtverzadigde fotosynthese fluctueerde meer dan drievoudig. Dit belemmerde de QTL-analyse van de fotosynthesemetingen onder veldomstandigheden. In deze studie werden zowel een statistisch model met covariantie als een fysiologische benadering gebruikt om de waarnemingen te standaardiseren. Beide benaderingen gaven aan dat de dampdrukverschillen tussen blad en lucht de grootste invloed op de fotosynthese hadden. Na correctie voor schommelingen in het microklimaat, bleek deze populatie significante genetische variatie te bevatten. We konden één tot drie QTLs detecteren per fotosynthese-gerelateerde eigenschap. Een belangrijk QTL werd gelokaliseerd nabij merker RM410 (het interval 57,3 tot 68,4 cM op Chromosoom 9); deze was consistent voor fenotypering tijdens de bloei en de korrelvulling en voor fenotypering onder omstandigheden van droogte en van voldoende water. Deze consistentie in QTLs werd ook bevestigd in een kasproef onder gecontroleerde omstandigheden. Deze resultaten lijken aan te geven dat fotosynthese in verschillende fenologische stadia en onder verschillende omstandigheden door dezelfde genetische factoren wordt beïnvloed. Deze resultaten gaven ook informatie voor het selecteren van de genotypen voor een gedetailleerde fysiologische modelstudie.

Op basis van de QTLs voor bladfotosynthese die werden gedetecteerd in Hoofdstuk 2 werden 13 ILs geselecteerd die als representatief voor de populatie konden worden beschouwd. Deze 13 ILs werden gebruikt om de fysiologische basis van genetische variatie in bladfotosynthese en de daaruit voortvloeiende QTLs te bestuderen (Hoofdstuk 3). Gasuitwisseling en chlorofylfluorescentie werden gelijktijdig gemeten bij verschillende niveaus van invallende straling en CO_2. Met deze metingen werden de CO_2- en lichtresponscurves onder omstandigheden van droogte en van voldoende water bepaald, zowel tijdens de bloei als tijdens de korrelvulling. Voor elke IL werden via "curve fitting" zeven parameters van het fotosynthesemodel geschat. Op deze manier werd de fotosynthese herleid tot CO_2 geleidbaarheid van de huidmondjes (g_s) en van het mesofyl (g_m), elektronentransportcapaciteit (J_{max}) en Rubisco-carboxyleringscapaciteit (V_{cmax}). Hoewel droogte en bladleeftijd het grootste deel van de totale variatie verklaarden, werd ook een significante genetische variatie voor deze parameters gevonden. Genetische variatie in lichtverzadigde fotosynthese, transpiratie-efficiëntie (TE), en de belangrijke QTL voor fotosynthese op Chromosoom 9 waren voornamelijk gekoppeld aan variatie in g_s en g_m. Dus bleken g_s en g_m, die volgens literatuur verantwoordelijk zijn voor de omgevingsvariatie in fotosynthese, ook

verband te houden met genetische variatie in fotosynthese. Bovendien bleken de relaties tussen deze parameters en bladstikstof of droge stof per bladoppervlakte-eenheid zowel te gelden voor variatie als gevolg van milieueffecten (zoals eerder reeds werd aangetoond) als voor variatie tussen genotypen. In het licht van deze resultaten en van eerdere rapportages in de literatuur werd beargumenteerd dat de variatie in fotosynthese als gevolg van omgevingsfactoren en als gevolg van genetische variatie op de zelfde fysiologische mechanismen gebaseerd zijn. Op basis van deze resultaten van het ecofysiologisch modelleren van fotosynthese en de QTL analyse werden ideotypes voor bladfotosynthese en TE ontworpen, die een verbetering van respectievelijk 17,0% en 25,1% gaven ten opzichte van het beste experimenteel onderzochte genotype. Deze analyse onderstreept dat het mogelijk is om de fotosynthese en de TE gelijktijdig te verbeteren binnen dezelfde genetische achtergrond.

Rijstproductie wordt niet alleen bepaald door fotosynthese, maar ook door andere fysiologische processen. Fysiologische eigenschappen, zoals het drooggewicht van het individuele zaad (S_W), de N-concentratie in het zaad (n_{SO}), de maximale planthoogte (H_{max}), het minimum aantal dagen voor vegetatieve groei (m_V), het minimum aantal dagen voor reproductieve groei (m_R), de specifieke bladoppervlakte (S_{la}), en de totale N opname van het gewas (N_{max}) werden gemeten (Hoofdstuk 4). Het mechanistische gewasmodel GECROS (Genotype-by-Environment interaction on CROp growth Simulator [NL: simulator van gewasgroei voor de interactie tussen genotype en milieu]) werd gebruikt om deze fysiologische deeleigenschappen te integreren en de gewasopbrengst van individuele ILs te voorspellen onder omstandigheden van voldoende water en van droogte. Wanneer het model werd gevoed met invoer van fysiologische eigenschappen op basis van feitelijke metingen kon het GECROS model de variatie in opbrengst voor 72% (nat) en 57% (droog) verklaren. Er werd een QTL analyse uitgevoerd op deze fysiologische model-invoerparameters. Vervolgens werden schattingen van deze eigenschappen berekend op basis van geschatte additieve alleleffecten en QTL allel-informatie van elke IL, gebruikmakend van moleculaire merkers. Wanneer de gemeten invoerparameters werden vervangen door op merkers gebaseerde schattingen verklaarde dit op QTL / merkers gebaseerde gewasmodel 52% (nat) en 47% (droog) van de variatie in opbrengst. Enkelvoudige correlatie-analyses en multiple regressie-analyses toonden aan dat N_{max} de meeste invloed had op de opbrengst; vijf andere genotype-specifieke model-invoer eigenschappen beïnvloedden de opbrengst ook significant, maar S_W deed dat niet. Met behulp van de merker-gebaseerde schattingen van fysiologische invoerparameters gaf GECROS ook een redelijke voorspelling van de variatie in opbrengst binnen een populatie van 251 recombinante inteeltlijnen van dezelfde ouders geteeld onder omstandigheden van

voldoende water of van droogte. Een gevoeligheidsanalyse met behulp van het groeimodel bleek een nuttig middel om het relatieve belang van de geïdentificeerde merkers voor de opbrengst vast te stellen. Bovendien werden zo meer merkers gedetecteerd dan op basis van merkerselectie met multiple regressie voor opbrengst *per se*. In onze analyse bleken de merkers RM8030 op Chromosoom 2 en RM338 op Chromosoom 3 het belangrijkst voor respectievelijk omstandigheden van voldoende water en van droogte. Dit suggereert dat de belangrijkste merkers om te gebruiken in de selectie op hoge korrelopbrengst milieu-specifiek zijn. Dit alles suggereert dat een QTL / merker gebaseerde modelbenadering de efficiëntie van merker-gestuurde selectie kan verbeteren.

Hoofdstuk 4 toonde ook aan dat bladfotosynthese niet bepalend was voor de opbrengstverschillen tussen de ILs zoals die werden waargenomen in de veldproef. Dat gold zowel voor de omstandigheden van voldoende water als voor die van droogte. Dit gebrek aan persistentie van variatie over de schalen heen is waarschijnlijk te wijten aan de complexe hiërarchie, van bladfotosynthese naar gewasopbrengst, en aan interacties en terugkoppelingsmechanismen die zich voordoen tussen fysiologische componenten van de individuele plant, tussen planten van hetzelfde gewas en tussen het gewas en de omgeving. Bovendien zijn de gemeten bladfotosynthese en de opbrengst beide behept met een zekere, willekeurige experimentele fout. Samen kunnen de complexiteit en de experimentele ruis de potentiële bijdrage van de kleine variatie in bladfotosynthese binnen een populatie aan de variatie in de uiteindelijke opbrengst maskeren. Teneinde de mate waarin natuurlijke genetische variatie in fotosynthese kan bijdragen aan het verhogen van de biomassaproductie en de opbrengst van rijst te onderzoeken, werd het GECROS gewasmodel wederom gebruikt om het effect van genetische variatie in bladfotosynthese op biomassaproductie van het gewas te analyseren (Hoofdstuk 5). Dit werd uitgevoerd door telkens één van de model-inputparameters constant te houden, dat wil zeggen te fixeren op het gemiddelde van de populatie voor alle ILs. Er werd aangetoond dat een genetische variatie in bladfotosynthese van 25% gelijkwaardig kan worden opgeschaald naar het gewasniveau, resulterend in een toename in biomassa van 22 tot 29% over verschillende locaties en jaren. Deze uitkomst was in tegenstelling tot eerdere studies waarbij het percentage verbetering afnam bij het opschalen van bladniveau naar gewasniveau. Het verschil met eerdere studies lijkt verband te houden met het feit dat in onze IL populatie variatie in zowel Rubisco-gelimiteerde als elektronentransport-gelimiteerde fotosynthese werd waargenomen. Rijstproductie kon aanzienlijk worden verbeterd door de natuurlijke variatie in het bestaande kiemplasma uit te baten, vooral de variatie van de parameters die de licht-gelimiteerde fotosynthese bepalen.

In Hoofdstuk 6 heb ik de resultaten van dit proefschrift samengevat en bediscussieerd. Ik evalueerde hoe fysiologisch modelleren en QTL / merker-gebaseerd modelleren veredeling van gewassen en genetica kunnen bijstaan. Tot slot besprak ik de vooruitzichten voor de toekomst van de integratie van gewasfysiologie, gewasmodellering, plantengenetica en de plantenveredeling, dat alles in het kader van de zogenaamde gewassysteembiologie.

Acknowledgements

First and foremost, I would like to express my deep and sincere gratitude to my multidisciplinary PhD supervisory team made up of Prof. dr. ir. Paul C. Struik [Centre for Crop Systems Analysis, WUR (CSA)], Prof. Huaqi Wang (Plant Breeding & Genetics, China Agricultural University), Dr. Xinyou Yin (CSA), and Dr. ir. Tjeerd-Jan Stomph (CSA). The team supported me from all the aspects. I could not have wished for a better one. You all have made it possible for me to commence and complete this challenging work.

Prof. Huaqi Wang was my supervisor during my MSc study (2004-2007). He is a famous aerobic rice breeder. He introduced me to the fascinating world of plant genetics. He gave me lots of practical guidance and support in the field experiments. I was inspired by his enthusiasm and hard work in rice breeding. After my MSc study, he encouraged me to apply for a scholarship from the China Scholarship Council, and sent me to Wageningen University to pursue my career abroad. Thanks to him, I got the chance to study in this beautiful, quiet, and world famous agricultural university. During my PhD study, he also gave me lots of support in arranging field experiments.

I am very grateful to Prof. Paul C. Struik for waiving my tuition fee and accepting me as a PhD student. Although I had great difficulties in the first period of my PhD study, he was very supportive and positive with me. He is very patient and his office was always open to students for discussion. He is very busy as the head of CSA, but he is so energetic and efficient that he could deal with every single detail in my PhD work. Paul is the best project leader I could ever have. He is open minded and ready to take responsibility and make decisions. His enthusiasm and love for science is inspiring me to pursue my career in science! I'd like also to express my gratitude to his wife Prof. dr. ir. Edith T. Lammerts van Bueren for her kindness and hospitality.

I am deeply grateful to my daily supervisor Dr. Xinyou Yin for introducing me to the fascinating world of 'crop modelling' and 'crop systems biology', for providing a solid basis in setting the whole framework of this thesis, for sharing with me your incredible knowledge and ideas about crop physiology and the future of crop science. He guided me throughout my study period even when I was in China. He patiently supervised me into the right direction. This project would not have been possible without his inspiring and novel thoughts. He was involved in almost all aspects of

planning, conducting, discussing, and writing. He always guided me through his quick responses, sharp analyses, and great mentorship. In him I see lots of the merits an excellent scientist should have. I hope one day I could reach the same level he has. Special thanks to him and his wife Yarong for their hospitality.

I would like to give special thanks to Dr. Tjeerd-Jan Stomph for his constructive suggestions during our project meetings, for patient discussions in details of the experiments, for his day-to-day support and guidance. He was motivating, encouraging, and enlightening. His carefulness and support ensured my experiments could run well. I am very grateful to Tjeerd-Jan for editing the manuscripts, helpful suggestions, many insightful discussions and suggestions. I really enjoyed the pleasant time with him and his wife Leentje den Boer.

I am thankful to Ing. Peter E.L. van der Putten. He is very dedicated and responsible. Without him, it would not have been possible for me to run the greenhouse experiment smoothly. I am thankful to Youbao Tian for assistance in the field experiments. I am also grateful to the staff members of the Wageningen Plant Sciences Experimental Centre (Unifarm), for supporting me with all the facilities and ensuring the success of the greenhouse experiment. I am thankful to Dr. Ad H. C. M. Schapendonk, Ing. Sander Pot, and Dr. Govert Trouwborst from Plant Dynamics BV for their technical training and support in photosynthesis measurements.

I am grateful to Dr. Jeremy Harbinson (HPC, WUR) for discussing the experimental setup for photosynthesis measurements. I am also grateful to Prof. Jaume Flexas (Universitat de les Illes Balears, Spain) for the fruitful discussion with him on photosynthesis and mesophyll conductance.

I would like to express my gratitude to journal editors, Dr. Mary Traynor, Dr. Diana Hudspith, Dr. Jianhua Zhang, Dr. Raquel González Cuesta, and Dr. Greg Rebetzke from *Journal of Experimental Botany* for their constructive suggestions on the manuscripts and assistance in the publishing processes. I thank anonymous reviewers for the critical and valuable comments which greatly improved the quality of manuscripts.

I would also like to thank all the people who made my life easier and happier in Wageningen. My special thanks to Mrs. Jenny Elwood, Mr. Eugen Goosens, Mrs. Wampie van Schouwenburg and Mrs. Sjanie van Roekel for their help in the logistics and secretarial work, Mrs. Gijsbertje Berkhout and Mr. Alex-Jan de Leeuw for their assistance in finance related matters. Thanks to all colleagues in CSA and Plant Production Systems group. I had a very good time with them during the last five years. I would like to thank Dr. ir. Jan Vos and his wife Ada Wolse, Dr. Wopke van der Werf and his wife Saskia Beverloo for the great hospitality.

Publications of the author

Gu J, Yin X, Stomph TJ, Struik PC. 2013. Can exploiting natural genetic variation in leaf photosynthesis contribute to increasing rice productivity? A simulation analysis. (Submitted)

Gu J, Yin X, Zhang C, Wang H, Struik PC. 2013. Molecular marker-based crop modelling to predict variation in yield among rice (*Oryza sativa* L.) genotypes grown under well-watered and drought-stress conditions. (Submitted).

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

Gu J, Yin X, Struik PC, Stomph TJ, Wang H. 2012. 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. *Journal of Experimental Botany* 63, 455-469.

Yin X, Sun Z, Struik PC, **Gu J**. 2011. Evaluating a new method to estimate the rate of leaf respiration in the light by analysis of combined gas exchange and chlorophyll fluorescence measurements. *Journal of Experimental Botany* 62, 3489-3499.

Wang X, **Gu J**, Song Q, Zhao P, Wang H. 2011. Development and analysis of upland rice introgression lines. *Southwest China Journal of Agricultural Science* 24(2): 406-409. (In Chinese with English abstract)

Zhao P, **Gu J**, Hao J, Hu J, Wang H. 2010. Drought tolerance identification and photosynthesis parameters analysis in booting stage of different generation introgression line. *Chinese Agricultural Science Bulletin* 26(14): 133-137. (In Chinese with English abstract)

Wang X, **Gu J**, La H, Song Q, Zhao P, Wang H. 2009. QTLs mapping on root traits of introgression lines at heading stage. *Chinese Agricultural Science Bulletin* 25(12): 14-19. (In Chinese with English abstract)

PE&RC PhD Training Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (6 ECTS)
- Combining biochemical modelling and genetic mapping to understand the physiological-genetic regulation of photosynthesis in response to drought (2008)

Writing of project proposal (4.5 ECTS)
- QTL-based physiological modelling to unravel the complexity of drought tolerance in rice (*Oryza sativa* L.) (2008)

Post-graduate courses (2.6 ECTS)
- Increasing photosynthesis in plants; PE&RC (2011)
- Mixed linear models; PE&RC (2012)

Laboratory training and working visits (2 ECTS)
- Training on gas exchange and chlorophyll fluorescence measurements; Plant Dynamics B.V. (2008)

Deficiency, refresh, brush-up courses (3 ECTS)
- Modern statistics for the life sciences (2008)

Competence strengthening / skills courses (2.1 ECTS)
- Techniques for writing and presenting scientific papers; CENTA (2011)
- How to write a world-class paper; Library WUR, Wageningen (2010)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.1 ECTS)
- PE&RC Days; WUR, Wageningen (2010, 2011)
- PE&RC Weekend; WUR, Wageningen (2011)

Discussion groups / local seminars / other scientific meetings (7.5 ECTS)
- Discussion group of Centre for Crop Systems Analysis (CSA) (2008-2012)
- Photosynthesis seminar series (2011)
- Seminar series plant sciences (2009-2012)

International symposia, workshops and conferences (3 ECTS)
- International aerobic rice workshop; Beijing (2007)
- Society of Experimental Biology (SEB) main meeting; Glasgow (2011)

Supervision of 2 MSc students (6 ECTS)
- Genetic variation for yield and yield components in a rice introgression line population under drought and well-watered conditions
- Genetic variation for photosynthesis in a rice introgression line population under drought and well-watered conditions

Curriculum vitae

Junfei Gu was born on 7th November, 1981 in Jiangsu Province, China. In 2000, he graduated from Jiangsu Rudong Senior High School, which was founded in 1938, as a state-level exemplary ordinary senior high school. At 2004, he finished his Bachelor of Science degree in Agriculture, majoring in Agronomy, at the College of Agriculture and Biotechnology, China Agricultural University (CAU), Beijing, China. From 2004 to 2007, he studied in the Department of Plant Breeding and Genetics, CAU, for an MSc degree. During his MSc studies, he worked on a thesis titled "Developing of introgression lines of upland rice (*Oryza sativa* L.) and QTL mapping on root traits". With a scholarship from the China Scholarship Council, he moved to the Netherlands and enrolled to the PhD programme of Wageningen University in December 2007. During his PhD study, he worked on the thesis "QTL-based physiological modelling of leaf photosynthesis and crop productivity of rice (*Oryza sativa* L.) under well-watered and drought environments" at the Crop Physiology group of the Centre for Crop Systems Analysis.