SUPPLEMENTARY DATA

Description of crop growth model GECROS

GECROS (Genotype-by-Environment interaction on CROp growth Simulator) is a generic crop model operating in daily time steps (with phenology and photosynthesis related processes simulated in shorter time scales). 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, 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 to allow 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 are listed at the end of this document.)

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)].

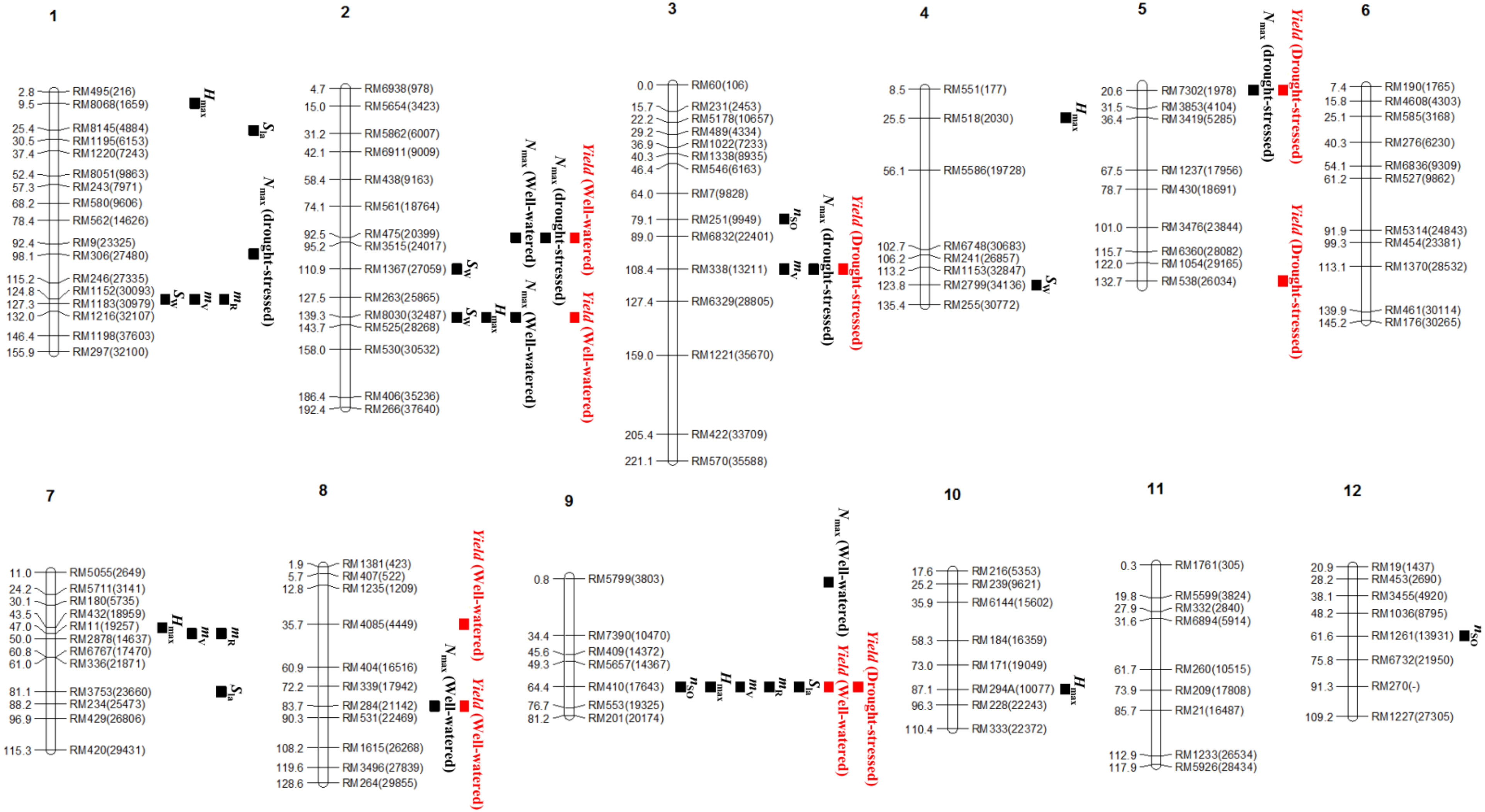

Fig. S2. Correlation between simulated values from marker-based model parameters and those from measured model parameter values for 96 rice genotypes of the IL population for grain yield (A) and biomass (B), for well-watered (open symbols, non-bold *r* value) and drought-stressed environment (filled symbols, bold *r* value).

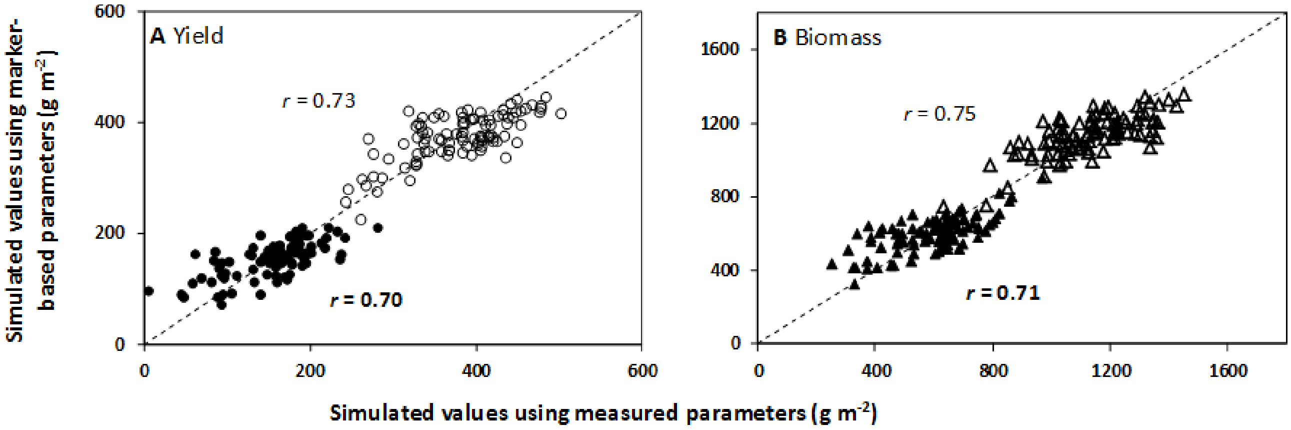

Table S1. Simple correlation coefficients for model input parameters, total crop N uptake (N_{max}) and yields for well-watered (WW) and drought-stressed (DS) environments (*n* = 96; for definition of these traits, see Table 1 in the main text).

	S_W	n_{SO}	H_{max}	m_V	m_R	S_{la}	$N_{max(WW)}$	$N_{max(DS)}$	Yield$_{(WW)}$	Yield$_{(DS)}$
S_W	1									
n_{SO}	0.11	1								
H_{max}	0.50***	0.28**	1							
m_V	-0.26*	0.12	-0.03	1						
m_R	0.26*	-0.13	0.02	-0.95***	1					
S_{la}	0.04	0.10	0.21*	0.44***	-0.48***	1				
$N_{max(WW)}$	0.29**	0.11	0.44***	-0.29**	0.32**	-0.20	1			
$N_{max(DS)}$	0.16	-0.16	0.29**	-0.42***	0.44***	-0.04	0.43***	1		
Yield$_{(WW)}$	0.26*	-0.33***	0.17	-0.51***	0.53***	-0.30**	0.76***	0.52***	1	
Yield$_{(DS)}$	0.09	-0.27**	0.00	-0.64***	0.65***	-0.21*	0.30**	0.77***	0.63***	1

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

References

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* **112,** 465-475.

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.