Field Crops Research 172 (2015) 119–131

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Predicting maize kernel number using QTL information

Agustina Amelong[a,*], Brenda L. Gambín[a], Alan D. Severini[b], Lucas Borrás[a]

[a] Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Campo Experimental Villarino S/N, Zavalla S2125ZAA, Santa Fe, Argentina
[b] Estación Experimental Agropecuaria Pergamino, Instituto Nacional de Tecnología Agropecuaria, Ruta 32 km 4.5, CP 2700 Pergamino, Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 30 May 2014
Received in revised form
17 November 2014
Accepted 26 November 2014
Available online 19 December 2014

Keywords:
Modeling
Crop physiology
Zea mays
Yield
Yield components

ABSTRACT

Most maize yield variations are explained by changes in the number of established kernels. Kernel number is, in turn, highly dependent upon ear biomass accumulation around flowering. Both are quantitative traits highly influenced by the environment. Determining the genetic basis of quantitative traits is complex because of usual genetic × environment interactions (GxE). Crop physiology models are proposed to help overcome this problem, as they are structured to predict consequences of GxE interactions based on dynamic responses. We studied the genetic basis of maize kernel number determination at the plant level by conducting two quantitative trait loci (QTL) analysis: (i) on final traits per se (kernel number per plant, KNP, and ear biomass per plant, EB) and (ii) on specific model parameters of well-documented curves describing KNP and EB response to plant growth around flowering. Quantitative trait loci for KNP, EB and model parameters relating KNP and EB to plant growth were determined for 125 RILs of the IBM Syn4 (B73 × Mo17) at two environments. We later grew several of these RILs and others from the same population not included in the QTL analysis and attempted to predict EB and KNP based on QTL information coming from each analysis. We hypothesized that doing the QTL analysis on crop physiology model parameters that describe the response of KNP and EB to plant growth is better than using direct QTL information.

All traits showed significant variation, and both analyses detected several QTL for the studied traits. Associated QTL for EB and KNP per se did not localize with QTL detected for model parameters. This is the first report describing genomic regions for key physiological traits related to maize biomass partitioning around flowering and kernel set efficiency per unit of accumulated EB 15 days after anthesis. Quantitative trait loci information of model parameters helped to predict accumulated EB and KNP with higher accuracy ($r^2 = 0.13$ and 0.12, $p < 0.001$, for EB and KNP, respectively) than trying to predict EB and KNP based on QTL detected on final traits per se ($r^2 < 0.01$ and <0.01, $p > 0.10$, for EB and KNP, respectively). However, predictions using an average crop physiology model parameter across genotypes and individual RIL plant growth gave the highest accuracy ($r^2 = 0.46$ and 0.37, $p < 0.001$, for EB and KNP, respectively). As such, we identified chromosome areas including potentially relevant genes involved in maize KNP determination, but this information helped to predict KNP at different environments only partially, suggesting other approaches might be needed.

Abbreviations: BLUP, best linear unbiased predictor; BIC, bayesian information criterion; CIM, composite interval mapping; C_{EB}, curvature for the relationship between plant growth rate and ear biomass; C_{KN}, curvature for the relationship between ear biomass and kernel number per plant; EB, ear biomass per plant 15 days after 50% anthesis; EB_b, ear biomass threshold for kernel set; Exp, experiment; GxE, genetic by environment interaction; IBM, intermated B73 × Mo17 population; IS_{EB}, initial slope of the relationship between ear biomass and plant growth rate; IS_{KN}, initial slope for the relationship between kernel number per plant and ear biomass; KNP, kernel number per plant; LOD, logarithm of odds; MIM, multiple interval mapping; PGR, plant growth rate around the flowering period; $PGR_{50\%}$, average plant growth rate; PGR_b, plant growth rate threshold for ear biomass accumulation; PGR_g, plant growth rate for fraction g of the population of plants; PGR_{SD}, standard deviation of plant growth rate; QTL, quantitative trait loci.

* Corresponding author. Tel.: +54 341 4970080x1229.
E-mail addresses: aamelong@unr.edu.ar, agusamelong@gmail.com (A. Amelong).

http://dx.doi.org/10.1016/j.fcr.2014.11.014
0378-4290/

1. Introduction

Maize yield is determined by the harvested kernel number per unit land area and average kernel weight. Both traits are important yield components, but kernel number is responsible for most yield variations (Early et al., 1967; Otegui, 1995; Chapman and Edmeades, 1999). Understanding and predicting the number of kernels per plant or per unit land area for different genotypes and environments is important for guiding maize breeding and crop management for yield improvement.

Kernel number per plant (KNP) is a quantitative trait. Quantitative traits are usually controlled by a large number of genes and are highly influenced by the environment. Chromosomal regions controlling quantitative traits are commonly referred as QTL (quantitative trait loci). It is possible to locate a QTL in the genome and estimate its effect by associating the trait of interest with molecular DNA markers (Tanksley, 1993). Several studies have focused on studying QTL for KNP in maize (e.g., Beavis et al., 1994; Ribaut et al., 1997; Agrama et al., 1999; Frova et al., 1999; Peng et al., 2011). However, most of these QTL studies failed to be consistent when evaluated at different environments. Bernardo (2008) estimates that genetic × environment interactions (GxE) are the major cause for discrepancies in results from QTL studies with complex traits.

A proposed solution to solve problematic GxE interactions is to use physiology-inspired models (Hammer et al., 2006). These models are equations with specific parameters describing the trait response to the environment. The response curves can vary for different genotypes as equation parameters can be genotype specific. These parameters allow linking the environment with the final trait. Reymond et al. (2003) determined the genetic basis (QTL) of the curve parameters and estimated the trait of interest (leaf growth) with the QTL effect of each model parameter (Tardieu, 2003; Hammer et al., 2006; Messina et al., 2009). If the model adequately captures the physiological determinants of the genetic variation it is possible to predict the trait of interest at different environmental scenarios for each genotype. Equations and parameters are the link between the environment and the genotype. Examples of this approach have been reported for leaf growth in rice (*Oryza sativa* L.; Wu et al., 2002), leaf elongation in maize (*Zea mays* L.; Reymond et al., 2003), flowering time in barley (*Hordeum vulgare* L.; Yin et al., 2005) and drought tolerance in maize (*Zea mays* L; Messina et al., 2011). Here, we followed a similar approach to study maize KNP determination.

The model we used is based on the well-documented relationship between KNP and plant growth during flowering (Edmeades and Daynard, 1979; Tollenaar et al., 1992; Otegui and Bonhomme, 1998; Andrade et al., 1999). Differences in KNP are commonly related to the plant growth rate (PGR) during ca. 30 days around flowering (Otegui and Bonhomme, 1998; Andrade et al., 1999). There are a number of studies describing genotype specific parameters relating biomass partitioning to the reproductive tissue bearing kernels in relation to plant growth, and its effect on the number of established kernels (e.g., Echarte et al., 2004; Echarte and Tollenaar, 2006). This is an opportunity to explore the genetic basis of KNP, as there is a response curve for the trait of interest (KNP) to link environmental variations with genotype-dependent parameters.

The response curve of KNP to PGR is usually analyzed as the result of two distinctive processes (Andrade et al., 1999; Echarte et al., 2004; Borrás et al., 2007; Pagano and Maddonni, 2007; D'Andrea et al., 2008): (i) the relation between KNP and the accumulated ear biomass (EB) at the end of the flowering period (15 days after 50% anthesis) and (ii) variations in accumulated EB in relation to PGR during flowering. The former is usually identified as the kernel set efficiency per unit of accumulated reproductive biomass, and the latter represents the biomass partitioning to the reproductive structure bearing kernels (Vega et al., 2001). Model parameters describing both curves are known to be genotype specific.

In the present article we studied the genetic basis of maize KNP determination by two different approaches: (i) QTL analysis on final traits per se (KNP and EB), and (ii) QTL analysis on model parameters of response curves describing KNP and EB to plant growth. The value of the second approach grounded on a similar physiological background has been incorporated recently for breeding simulations studies (Messina et al., 2011). We hypothesized that a QTL analysis on crop-physiology model parameters that describe the response of KNP to plant growth enhances the accuracy of predicting KNP at different growth environments than trying to predict KNP by doing a KNP QTL analysis directly. We used a RIL (IBM Syn4 B73 × Mo17) population and conducted the QTL study at two contrasting growth environments. Finally, we tested the hypothesis using other non-related experiments.

2. Materials and methods

2.1. Plant material

Parental inbred lines B73 and Mo17 and the IBM Syn4 (B73 × Mo17) RIL population were used. B73 and Mo17 are both reference breeding lines (Troyer, 1999), and B73 has been sequenced and thoroughly studied (Schnable et al., 2009). Recombinant inbred lines were derived from eight generations of self-pollination, after being intermated four times at the F_2 generation (Lee et al., 2002). Genotypic marker data (RFLP, SSR, ESTs-SSR, INDELS, SNPs, etc.) for these RILs are publicly available at Maize GDB website (http://www.maizegdb.org, verified 1 March 2014). The population has become a reference population in different mapping studies (http://www.maizegdb.org, verified 1 March 2014). Parental inbred lines differ in KNP (Abertondo, 2007; Severini et al., 2011) and PGR and EB accumulation during flowering (Borrás et al., 2009; Severini et al., 2011). Also, parental lines differ in model parameters describing plant biomass partitioning during flowering (Borrás et al., 2009). The RILs from the IBM Syn4 differ in KNP (Abertondo, 2007). Seeds were increased at Nidera Seeds (Venado Tuerto, Argentina), located less than 100 miles from university trials. Seeds were produced by selfing and bulked over the entire row.

2.2. Field experiments and experimental design

Five field experiments (Exp) were conducted (Table 1). Experiments I, II and III were located at Campo Experimental Villarino, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Zavalla, Argentina (33°01S, 60°52W), Exp IV at the Estación Experimental INTA Pergamino, Argentina (33°56S, 60°33W) and Exp V at the Brunner Farm, Iowa State University, USA (42°2N, 96°39W).

Exps I and II were planted on 14 September 2009 and 4 October 2010, respectively. Genotypes included 167 RILs and IBM Syn4 parents in Exp I, and 238 RILs plus parents in Exp II. A total of 125 RILs were coincident between Exps I and II. Experiments III and IV included some coincident and other RILs not used at Exps I and II. Experiments III and IV were planted on 17 October 2011 and 15 September 2009, respectively. Experiment V was planted on 11 May 2007 and included only the population parents.

Plots were arranged in a randomized complete block design with three replicates in all experiments. Plots were always overplanted and thinned at V2–V3. In Exps I and II stand densities were 5 and 7.5 pl m^{-2}, respectively. In Exps III, IV and V stand density treatments were used for all genotypes (Table 1). Stand density was 4 and 10 pl m^{-2} in Exp III, 3 and 12 pl m^{-2} in Exp IV and 3 and 9 pl m^{-2} in Exp V, arranged in a split plot design with

Table 1
General description of field experiments.

Exp	Location	Planting date	Number of evaluated genotypes	Stand density
				pl m^{-2}
I	Zavalla	14 Sept 2009	167 + parents	5
II	Zavalla	4 Oct 2010	238 + parents	7.5
III	Zavalla	17 Oct 2011	11 + parents	4 and 10
IV	Pergamino	15 Sept 2009	8 + parents	3 and 12
V	Iowa	11 May 2007	parents	3 and 9

stand density as plot and genotypes as sub-plot. Plots consisted of four rows 0.70 m apart, 5.5 m long in Exps I and II. In Exp III rows were 5.5 m long and 0.52 m apart. In Exp IV rows were 7 m long and 0.7 apart. In Exp V each plot consisted of six (high density) or eight (low density) rows, 5.5 m long and 0.76 m apart. Details from Exp V have been described in Severini et al. (2011). Fertilizer in Exps I, II and III was broadcast 4 to 5 d before planting at a rate of 100 kg N ha^{-1} (20-0-0-16, N-P-K-S) and 20 kg N ha^{-1} (10-50-0-0) was incorporated at planting. In Exp IV we applied 200 kg N ha^{-1} (46-0-0-0) divided in two stages (V4 and V10). In Exp V fertilizer was broadcast before planting at a rate of 110 kg N ha^{-1} (46-0-0-0).

The experimental area was always kept free of weeds and pests. Experiments were conducted without visible water limitations. Experiments II and III were watered with a sprinkler irrigation system during flowering and early grain filling with ca. 150 mm in both cases. In the other experiments no irrigation was needed because timely rainfall events occurred and no visible signs of water stress were evident.

2.3. Phenotypic measurements

Plant growth rates were determined on 15 plants per replicate in Exps I to IV, and 30 plants per replicate in Exp V. These plants were tagged approximately 15 days before 50% anthesis and were always consecutive plants in center rows. Nondestructive allometric models (Vega et al., 2000) were used to estimate plant biomass at the pre- and post-flowering stages for all tagged plants. The pre-flowering biomass samples were taken ca. 10 days before 50% anthesis. The allometric approach was used to estimate biomass accumulation, partitioning and kernel set at the individual plant level. This technique has been used successfully to provide an accurate representation of plant biomass corresponding to tagged plants remaining in the field until harvest (Vega et al., 2000; Echarte et al., 2004; Gambín et al., 2006). The pre- and post-flowering allometric models were developed each using 9 to 18 additional tagged plants per genotype (or per genotype × density combination when corresponding). For each model we sampled three plants from border rows at each replicate. The pre-flowering model was based on the linear regression between shoot biomass and stem volume (Vega et al., 2001; Gambín et al., 2006). Stem volume was calculated from plant height (ground level up to the uppermost leaf collar) and stem diameter at the base of the stalk. The r^2 values for this model ranged from 0.69 to 0.98 ($p < 0.001$, $n = 9$ at Exps I to IV and $n = 18$ at Exp V) across genotypes and years. The post-flowering biomass sample was taken 15 days after 50% anthesis, and the allometric model used stem volume and maximum apical ear diameter with husks (Vega et al., 2001; Gambín et al., 2006), and was fitted using a multiple linear regression analysis following Borrás et al. (2009). The r^2 values for this model ranged from 0.51 to 0.99 ($p < 0.001$, $n = 9$ at Exps I to IV and $n = 18$ at Exp V). Shoot biomass was determined after cutting plants and drying them in an air-forced oven at 65 °C for at least 7 days. These measurements were also used to estimate EB 15 days after 50% anthesis. The r^2 values ranged from 0.54 to 0.99 ($p < 0.001$, $n = 9$ at Exps I to IV and $n = 18$ at Exp V).

Plant growth rate around flowering (g pl^{-1} d^{-1}) was calculated as the difference between post-flowering and pre-flowering plant biomass (g pl^{-1}) divided by the number of days between samples. The plant-to-plant variability in PGR around flowering was determined for each plot using the individual tagged plants within each plot (individual PGR standard deviation; PGR_{SD}).

Ears from tagged plants were harvested at physiological maturity (defined as 75% milk line; Hunter et al., 1991). Kernels per plant were counted manually. Ears were dried in an air-forced oven and shelled individually. Plants with less than ten kernels at maturity were considered barren (Tollenaar et al., 1992). Barrenness was determined as the percentage (%) of barren plants. Prolificacy was reported as the average number of ears per plant at each plot (ear pl^{-1}).

In order to test individual trait variability, a general linear model was used for each trait separately. PROC GLM from SAS® (SAS Institute, 1999) was used for this analysis. Coincident genotypes from Exps I and II were analyzed together and the model included environment (years), blocks nested within environments, genotypes and genotype × environment interaction. Environments and blocks within environments were considered as fixed factors while genotypes and the genotype × environment interaction were considered as random factors. The rest of the genotypes from Exps I and II and the other experiments were analyzed separately. The model included blocks, genotypes, and in Exps III, IV and V stand density and genotype × stand density interaction. Individual means were compared by protected least significant difference (LSD).

2.4. Parameters of the crop physiology model

The model used to relate KNP with PGR was based on a modified version of Borrás et al. (2009). Kernel number per plant determination was analyzed as two processes. First, EB response curve to PGR was analyzed fitting a hyperbolic function (Eqs. (1) and (2)):

$$EB = 0 \quad \text{if PGR} \leq \text{PGR}_b \tag{1}$$

$$\text{EB} = \frac{[\text{IS}_{\text{EB}}(\text{PGR} - \text{PGR}_b)]}{[1 + \text{C}_{\text{EB}}(\text{PGR} - \text{PGR}_b)]} \quad \text{if PGR} > \text{PGR}_b \tag{2}$$

where IS_{EB} is the initial slope of the response curve of EB to PGR, PGR_b is the PGR threshold for EB accumulation and C_{EB} defines the curvature of this relationship at high PGR around flowering.

Second, KNP response curve to EB was analyzed by fitting a hyperbolic function similar to (1) and (2) (Eqs. (3) and (4)]:

$$\text{KNP} = 0 \quad \text{if EB} \leq \text{EB}_b \tag{3}$$

$$\text{KNP} = \frac{[\text{IS}_{\text{KN}}(\text{EB} - \text{EB}_b)]}{[1 + \text{C}_{\text{KN}}(\text{EB} - \text{EB}_b)]} \quad \text{if EB} > \text{EB}_b \tag{4}$$

where EB_b is the ear biomass threshold for kernel set, IS_{KN} is the initial slope of KNP vs. EB and C_{KN}, curvature of the relationship.

These relations were built using all tagged plants from each 125 RILs, combining phenotypic data from Exps I and II (90 plants per line, corresponding to 15 plants per replicate, three replicates per trial and two growing seasons). Equations were individually fitted for each genotype, and parameters PGR_b, IS_{EB} and C_{EB} for the relation of EB vs. PGR and EB_b, IS_{KN} and C_{KN} for the relation of KNP vs. EB were obtained. The different stand densities used at Exps I and II provided a range of PGR, allowing us to build solid response curves

for each genotype. The r^2 of these relations was 0.70 or higher and always highly significant ($p < 0.001$).

Broad sense heritability of model parameters was calculated on a mean basis as:

$$H^2 = \frac{\sigma_G^2}{\left[\sigma_G^2 + \sigma_e^2/r\right]} \quad (5)$$

where σ_G^2 is the genotypic variance, σ_e^2 is the plot residual variance and r is the number of replicates (Hallauer et al., 2010). Parameters used for heritability estimates differ from values used for QTL analysis, where curves were built with data combining all replicates and experiments (Exps I and II). For heritability estimates model parameters were calculated combining each replicate from Exps I and II ($n = 30$). Our heritability calculation (Eq. (5)) does not include a GxE term because it could not be estimated.

2.5. Genetic map construction and QTL analysis

The genetic analysis was done on final traits per se (KNP and EB), and on parameters of the crop physiology model (PGR_b, IS_{EB}, C_{EB}, EB_b, IS_{KN} and C_{KN}). For this we used phenotypic data from Exps I and II of 125 common RILs.

For EB and KNP phenotypic analysis we used a multi-trait multi-environment mixed model following Malosetti et al. (2008) and Alvarez Prado et al. (2013). We assumed genotypes to be random. Random genetic effects were assumed to be normally distributed with a variance-covariance structure. The choice of the best model for our data was based on Bayesian information criterion (BIC) (Schwarz, 1978). For our data set, the best variance-covariance model was compound symmetry. Best linear unbiased predictors (BLUP) of each parameter for each genotype were estimated. Estimations were calculated with SAS® PROC MIXED (SAS Institute, 1999).

For model parameters (PGR_b, IS_{EB}, C_{EB}, EB_b, IS_{KN} and C_{KN}), one value per genotype was obtained after fitting the response equations (Eqs. (1)–(4)). Parameters that were not normally distributed were transformed (Lübberstedt et al., 1998; Fiedler et al., 2012). Modified Shapiro and Wilk test (Rahman and Govindarajulu, 1997) was used for this analysis with InfoStat v2013 (DiRienzo et al., 2013). This was the case for IS_{EB}, C_{EB}, IS_{KN} and C_{KN}. IS_{EB} and IS_{KN} were transformed with $\sqrt{1}/(1+x)$ and C_{EB} and C_{KN} were transformed with $(1/1+x)$.

DNA marker data used for building the linkage map were retrieved from the Maize GDB website (http://www.maizegdb.org, verified 1 March 2014). Various types of molecular markers are available. Prior to linkage analysis, the χ^2 (Chi-square) test was used to verify 1:1 mendelian segregation for each molecular marker (Kearsey and Pooni, 1996). Markers showing a distortive segregation were discarded from the analysis. From DNA markers showing the expected mendelian segregation we selected approximately four markers per bin (according to the IBM Syn4 consensus map) to construct a preliminary linkage map. A QTL analysis was carried out in order to detect interval markers associated with each trait. This preliminary linkage map was constructed with 362 molecular markers. Markers were added around the detected QTL. A total of 513 DNA markers were used from the IBM Syn4 consensus map for the final linkage map. Linkage analysis was carried out using MapDisto V1.7 (Lorieux, 2007). Map distances were computed with the Haldane mapping function (Haldane, 1919). A common map was used for all traits.

Quantitative trait loci mapping was done using WinQTL Cartographer V2.5 (Wang et al., 2010). We followed a procedure divided in two steps. First, with the final linkage map a multiple trait mapping procedure was done, which considers trait correlations. Composite interval mapping method of selection was used, with a stepwise forward-backward regression analysis (Model 6 from WinQTL Cartographer V2.5; Wang et al., 2010). We used a threshold of 0.05 to select putative QTL to be used as cofactors for controlling genetic background and a 10 cM window size for removing temporarily the marker effects when scanning the chromosome, with scanning intervals of 1 cM between flanking markers and putative QTL. Quantitative trait loci positions were assigned to relevant regions at the point of maximum LOD. Quantitative trait loci positions detected in the first step were regarded as candidate QTL and constituted the initial model for the second step. In this final step, we constructed a multi-QTL model by using a multi-trait multiple interval mapping (MT-MIM) procedure. By giving an initial model, the procedure estimated the model parameters, refined the estimates of QTL positions within intervals, tested the significance of all parameters for individual traits and searched for more QTL. Stepwise selection procedure was used and the final model was selected from a sequence of nested tests (Basten et al., 2004). We used the likelihood ratio test (LRT) to compare the significance of each model refinement (Kao, 1999). The threshold for declaring the presence of a significant QTL was LOD = 2.5 (Van Ooijen, 1999). Additive effects of transformed traits were estimated with untransformed data. Only correlated traits were grouped and analyzed together.

Despite several authors consider that epistasis has an important contribution in quantitative traits in maize (Blanc et al., 2006; Dudley and Johnson, 2009; Gonzalo et al., 2010), we did not calculate this in our analysis. In our population the power to detect epistasis is rather low, as the number of individuals to test all interactions is too small (Zhang et al., 2012).

2.6. Predicted EB and KNP using QTL information

Ear biomass and KNP were estimated for a set of independent data using genetic information. This includes a set of genotypes from Exps I and II that were not included in the QTL analysis, and genotypes from Exps III, IV and V (Table 1). Estimations were made following two different procedures. The first used QTL information on traits per se (EB and KNP), and the second used QTL information of parameters of the crop physiology model.

The phenotypic value of a trait was estimated using the following genetic model. Assuming two QTL were detected for any trait of interest, the genetic model is:

$$Y_k = \mu + a_i X_{Aik} + a_j X_{Ajk} \quad (6)$$

where γ is the phenotypic value of a trait on the k-th genotype ($k = 1, 2, \ldots, n$); μ is the trait mean of the 125 genotypes; a_i and a_j are the additive effects of the two putative QTL (Q_i and Q_j, respectively); x_{Aik} and x_{Ajk} are coefficients of QTL effects derived according to the genotypes of the flanking markers and the test positions.

For the first prediction, γ_k in model (6) represents the EB or KNP for k-th genotype, and both were determined straightforward based on detected QTL from the genetic analysis on traits per se.

For the second approach, γ_k in model (6) represents each parameter of the crop physiology model for k-th genotype (PGR_b, IS_{EB}, C_{EB}, EB_b, IS_{KN} and C_{KN}). The value of each parameter was determined based on detected QTL from the genetic analysis conducted on model parameters. Once parameter values for each genotype were estimated, it was possible to construct the response curves of EB to PGR and KNP to EB for each specific genotype. Ear biomass and KNP estimates following this approach requires PGR as input. From a corrected mean of $PGR_{50\%}$ and PGR_{SD}, for each experiment × genotype × stand density combination (SAS® PROC GLM, SAS Institute, 1999), estimates were done following these steps:

I. Divide the population into cohorts and calculate a PGR for each cohort (PGR_g, Fig. 1A). The variability in PGR for each experiment × genotype × stand density combination was divided into

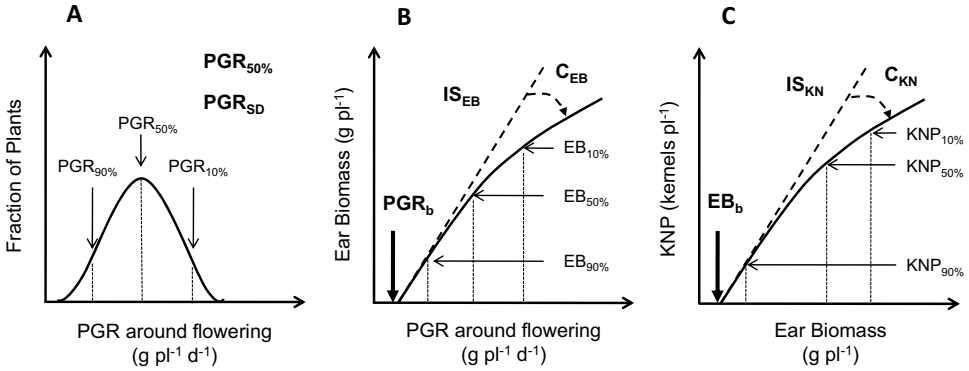

Fig. 1. Schematic diagram describing the crop physiology model for EB and KNP determination (based on Borrás et al., 2007). (A) The population of plants from each treatment combination is parsed into sub fractions. A plant growth rate (PGR_g) is calculated for each fraction of the population based on the average plant growth rate ($PGR_{50\%}$) and the standard deviation of plant growth rate (PGR_{SD}). Arrows indicate fractions of the population having average ($PGR_{50\%}$), 10% fastest ($PGR_{10\%}$), and 10% slowest ($PGR_{90\%}$) plant growth rates. (B) The PGR_g determines the ear biomass (EB) that each fraction of the population achieves 15 d after 50% anthesis (EB_g). The hyperbolic function for EB (Eqs. (1) and (2)) is defined by a plant growth rate threshold for ear growth (PGR_b), initial slope (IS_{EB}) and curvature (C_{EB}), all of which are genotype-dependent. (C) The EB_g determines the kernel number per plant (KNP) that each fraction of the population achieves. The hyperbolic function for KNP (Eqs. (3) and (4)) is defined by EB threshold for kernel set (EB_b), initial slope (IS_{KN}) and curvature (C_{KN}), all of them genotype-dependent. Averaging the EB and KNP determined for each fraction of the population of plants (EB_g and KNP_g) gives the average value for each specific treatment combination.

100 equal cohorts. The value assigned to each cohort of PGR_g was based on the $PGR_{50\%}$ and PGR_{SD} for each treatment combination, assuming a normal distribution of growth rates (when tested PGR was normally distributed in all experiments). Allometric models used to estimate individual PGR provided plant-to-plant variability for each combination.

II. Calculate the expected accumulated EB 15 d after 50% anthesis for each PGR_g. This is based on the EB vs. PGR response curve controlled by parameters PGR_b, IS_{EB}, and C_{EB} (Fig. 1B).

III. Calculate estimated KNP for each canopy cohort. This is based on the KNP vs. EB curve controlled by parameters EB_b, IS_{KN} and C_{KN} (Fig. 1C).

IV. The estimated EB and KNP of each cohort was averaged and finally estimated EB and KNP for each experiment × genotype × stand density combination were determined.

It is important to understand that the need to divide the population of plants into cohorts is based on the concept that although PGR is normally distributed, EB and KNP are not (Borrás et al., 2007). Calculating EB and KNP using a single PGR value per plot do not give the same result than estimating the EB and KNP for each PGR fraction of plants and then averaging the different cohorts (Borrás et al., 2007).

Last, we further tested this second approach using only an average value for each model parameter for all RILs (i.e., not considering QTL information) and specific RIL values for $PGR_{50\%}$ and PGR_{SD}. This was done for testing the effect of having specific genotypic values for model parameters.

Predicted EB and KNP values were compared to observed corrected means of EB and KNP from each experiment × genotype × stand density combination (SAS® PROC GLM, SAS Institute, 1999).

3. Results

3.1. KNP and yield differences from Exps I and II

A total of 125 RILs were phenotyped for final yield per plant, KNP, EB, PGR, PGR_{SD}, barrenness and prolificacy during two growing seasons. Average values across RILs and explored variation for each trait are described in Table 2. Significant differences were found among genotypes, and a genotype × year interaction was significant ($p < 0.001$) for all traits. This interaction indicated a modification in genotype ranking across years. Average differences between Exps I and II for yield and other traits indicated that the two environments (tested as years) helped explore a wide phenotype range for each genotype. Genotype yields ranged from 2.2 to 6.0 Mg ha^{-1} in Exp I and from 1.0 to 5.5 Mg ha^{-1} in Exp II. Kernel number per plant ranged from 201 to 654 kernels pl^{-1} in Exp I and from 69 to 396 kernels pl^{-1} in Exp II. Differences among genotypes in KNP were positively correlated with yield differences ($r^2 = 0.84$; $p < 0.001$; $n = 250$).

Accumulated EB 15 days after anthesis varied from 7.2 to 34.8 g pl^{-1} in Exp I. Experiment II ranged from 2.8 to 18 g pl^{-1} (Table 2).

Plant growth rate around flowering also showed significant variation. It ranged from 2.2 to 5.5 g pl^{-1} d^{-1} and from 1.3 to 3.2 g pl^{-1} d^{-1} for Exps I and II, respectively. The standard deviation for PGR (PGR_{SD}) varied from 0.2 to 2.1 and 0.3 to 1.7 g pl^{-1} d^{-1} in Exps I and II, respectively. Average PGR_{SD} values were 0.7 g pl^{-1} d^{-1} for Exp I and 0.6 g pl^{-1} d^{-1} for Exp II (Table 2).

Values for barrenness ranged from 0 to 17.6% and 0 to 59.5% at Exps I and II, respectively. Prolificacy ranged from 0.82 to 1.84 and 0.40 to 1.40 for Exps I and II, respectively.

In summary, genotypes and environments showed large and significant variation for all phenotyped traits.

3.2. Differences among RILs in model parameters

Parameters of the response curves of KNP to EB (Fig. 1C) and EB to PGR (Fig. 1B) were determined for each of the 125 RILs evaluated in Exps I and II. Response curves and estimated parameters are described in Fig. 2 for one specific RIL (IBM157) as an example. Recall, parameters of interest were: (i) the PGR threshold of the EB response to PGR (PGR_b), (ii) the initial slope of this response (IS_{EB}), (iii) the curvature of this relationship at high PGR around flowering (C_{EB}) (iv) the EB threshold for kernel set (EB_b), (v) the initial slope of KNP response to EB (IS_{KN}) and (vi) the curvature of this relationship (C_{KN}) (Fig. 2).

The variability among the 125 RILs for these parameters was large and the distribution frequency always showed transgressive segregation (Fig. 3). The maximum explored range was 0 to 2.1 g pl^{-1} d^{-1} for PGR_b, 3.9 to 45.8 g g^{-1} d^{-1} for IS_{EB}, 0 to 3.1 g^{-1} pl^{-1} d^{-1} for C_{EB}, −3.3 to 6.9 g pl^{-1} for EB_b, 21 to 152 kernels g^{-1} for IS_{KN} and 0 to 0.34 g^{-1} pl^{-1} for C_{KN}, (Fig. 3).

Table 2
Phenotypic data describing yield per plant, kernel number per plant, average plant growth rate (PGR), individual PGR standard deviation (PGR_{SD}), ear biomass, barrenness and prolificacy of 125 RILs from Exps I and II used for QTL analysis.

Exp		Yield (g pl^{-1})	Kernel number per plant (kernels pl^{-1})	PGR (g pl^{-1} d^{-1})	PGR_{SD} (g pl^{-1} d^{-1})	Ear biomass (g pl^{-1})	Barrenness (%)	Prolificacy (ear pl^{-1})
I	Average	89	401	3.4	0.7	18.0	1.8	1.09
	Min	43	201	2.2	0.2	7.2	0.0	0.82
	Max	120	654	5.5	2.1	34.8	17.6	1.84
II	Average	42	232	2.3	0.6	9.3	12.6	0.88
	Min	13	69	1.3	0.3	2.8	0.0	0.40
	Max	73	396	3.2	1.7	18.0	59.5	1.40
	Gen	***[a]	***	***	***	***	***	***
	Year	***	***	***	***	***	***	***
	*Gen*Year*	***(17)[b]	***(69)	***(0.5)	***(0.2)	***(3.4)	***(12.6)	***(0.15)

[a] ns: not significant ($p > 0.05$); * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.
[b] LSD values ($p < 0.05$).

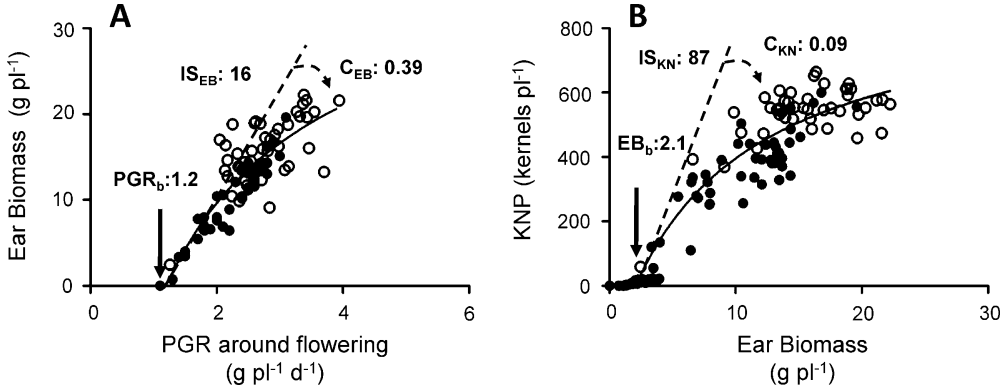

Fig. 2. Example showing values of model parameters for one RIL (IBM157). (A) Describes ear biomass (EB) response to plant growth rate (PGR) ($n = 90$, $r^2 = 0.74$). (B) Describes kernel number per plant (KNP) response to EB ($n = 90$, $r^2 = 0.77$). Model parameters are: PGR threshold for EB accumulation (PGR_b, g pl^{-1} d^{-1}), initial slope (IS_{EB}, g g^{-1} d^{-1}) and curvature (C_{EB}, g^{-1} pl^{-1} d^{-1}) of EB vs. PGR, EB threshold for KNP set (EB_b, g pl^{-1}), initial slope (IS_{KN}, kernels g^{-1}) and curvature (C_{KN}, g^{-1} pl^{-1}) for the relationship of KNP vs. EB.

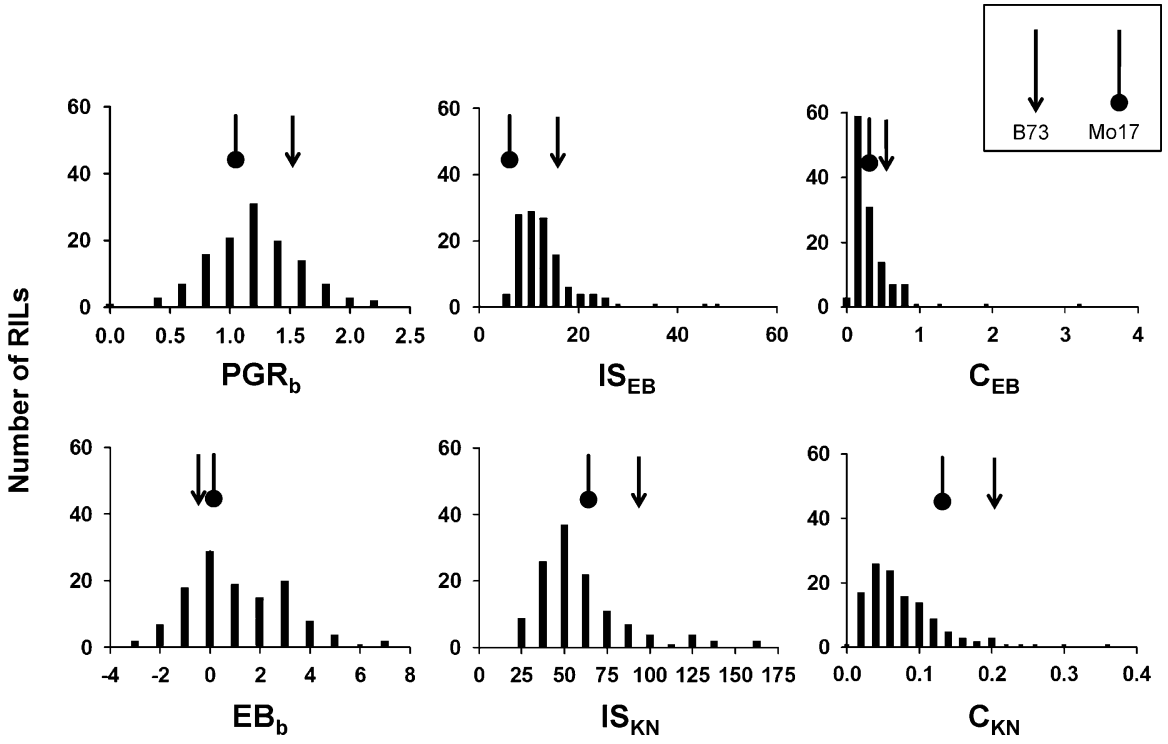

Fig. 3. Frequency distribution of model parameters (PGR_b, IS_{EB}, C_{EB}, EB_b, IS_{KN} and C_{KN},) in the evaluated RILs population. PGR_b (g pl^{-1} d^{-1}) is the PGR threshold for EB accumulation, IS_{EB} (g g^{-1} d^{-1}) is the initial slope and C_{EB} (g^{-1} pl^{-1} d^{-1}) curvature of EB vs. PGR relationship, EB_b is the EB threshold for kernel set (g pl^{-1}), IS_{KN} (kernels g^{-1}) is the initial slope and C_{KN} (g^{-1} pl^{-1}), curvature of the relationship of KNP vs. EB (see Fig. 2 for details).

Table 3
Pearson correlations between parameters of the crop physiology model. Parameter PGR_b ($g\,pl^{-1}\,d^{-1}$) is the PGR threshold for EB accumulation, IS_{EB} ($g\,g^{-1}\,d^{-1}$) is the initial slope and C_{EB} ($g^{-1}\,pl^{-1}\,d^{-1}$) curvature of EB vs. PGR relationship. Parameter EB_b ($g\,pl^{-1}$) is the EB threshold for kernel set, IS_{KN} (kernels g^{-1}) is the initial slope and C_{KN} ($g^{-1}\,pl^{-1}$), curvature of the relationship of KNP vs. EB (see Fig. 2 for details).

	PGR_b	IS_{EB}	C_{EB}	EB_b	IS_{KN}
IS_{EB}	0.51***[a]	–	–	–	–
C_{EB}	0.44***	0.85***	–	–	–
EB_b	n.s.	n.s.	n.s.	–	–
IS_{KN}	n.s.	n.s.	n.s.	0.61***	–
C_{KN}	n.s.	n.s.	n.s.	0.53***	0.95***

[a] ns: not significant ($p > 0.05$); * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Broad sense heritability was 0.55 for PGR_b, 0.43 for IS_{EB}, 0.43 for C_{EB}, 0.58 for EB_b, 0.19 for IS_{KN} and 0.17 for C_{KN}.

As expected, significant correlations were observed among model parameters within each response curve (Table 3). Parameters PGR_b, IS_{EB} and C_{EB} were positively correlated ($p < 0.001$). Similar positive correlations were found for parameters EB_b, IS_{KN} and C_{KN} ($p < 0.001$).

3.3. QTL analysis on EB and KNP

A QTL analysis was done on final traits per se (EB and KNP) summarized at Table 2. A multi-trait mapping using CIM analysis was carried out, which constituted the starting point for the MT-MIM analysis. Total genetic variation explained by the final MT-MIM model was 45 and 34% for EB and KNP respectively (Table 4).

Fig. 4 describes the QTL localization together with the estimated additive effect of B73 alleles for each analyzed trait.

A total of twelve QTL were found, but not all QTL were significant for every individual trait because the maximum marginal peak was lower than the LOD threshold. However, for these QTL the joint analysis was significant (data not shown). Eight significant QTL were detected for EB and five for KNP (Table 4). Consistent with their phenotypic correlation ($r^2 = 0.67$; $p < 0.001$; $n = 250$), positive correlations were found between additive effects of EB and KNP QTL located at the same positions.

Table 4
Description (chromosome, Chrom; position, cM; logarithm of odds, LOD; additive effect) of detected QTL for final traits (EB: ear biomass 15 days after anthesis, and KNP: kernel number per plant), and for parameters of the crop-physiology model (see Fig. 2). Additive effects for model parameters were calculated using untransformed values.

QTL Analysis	Trait	r^2	Chrom	Position (cM)	LOD	Additive effect
Final traits	EB	0.45	3	120.0	5.7	−1.3
			4	120.0	5.0	1.3
			4	140.0	5.7	−1.3
			5	152.8	8.6	1.8
			7	40.0	8.8	−1.8
			7	60.9	7.7	1.5
			8	46.0	5.4	−1.3
			10	0.1	3.9	1.0
	KNP	0.34	3	120.0	5.2	−22
			5	152.8	10.6	33
			7	40.0	6.2	−23
			7	60.9	10.9	34
			8	46.0	6.1	−26
Response curve parameters	PGR_b	0.39	2	113.7	2.8	−0.2
			2	133.8	5.6	−0.3
			2	140.0	3.5	0.2
			3	183.1	4.7	−0.3
			4	0.1	2.6	0.2
	IS_{EB}	0.17	3	183.1	4.5	−3.1
			8	49.5	2.7	2.4
	C_{EB}	0.26	2	133.8	3.6	−0.19
			3	173.6	8.2	0.23
			3	183.1	4.6	−0.19
			7	147.7	6.5	0.02
			8	49.5	3.2	0.13
	EB_b	0.25	1	69.1	4.6	−1.1
			7	91.8	4.8	1.2
	IS_{KN}	0.25	1	69.1	10.3	−14.3
			2	105.0	13.6	8.9
			5	15.0	7.2	−6.5
			5	120.0	10.6	−12.3
			7	91.8	15.8	6.3
			8	101.1	14.3	−11.9
			8	120.0	5.8	6.6
	C_{KN}	0.25	1	69.1	5.3	−0.02
			2	105.0	7.3	0.02
			5	15.0	7.8	−0.02
			5	120.0	19.5	−0.04
			7	91.8	5.2	0.01
			8	101.1	9.5	−0.02
			9	90.0	3.3	−0.02

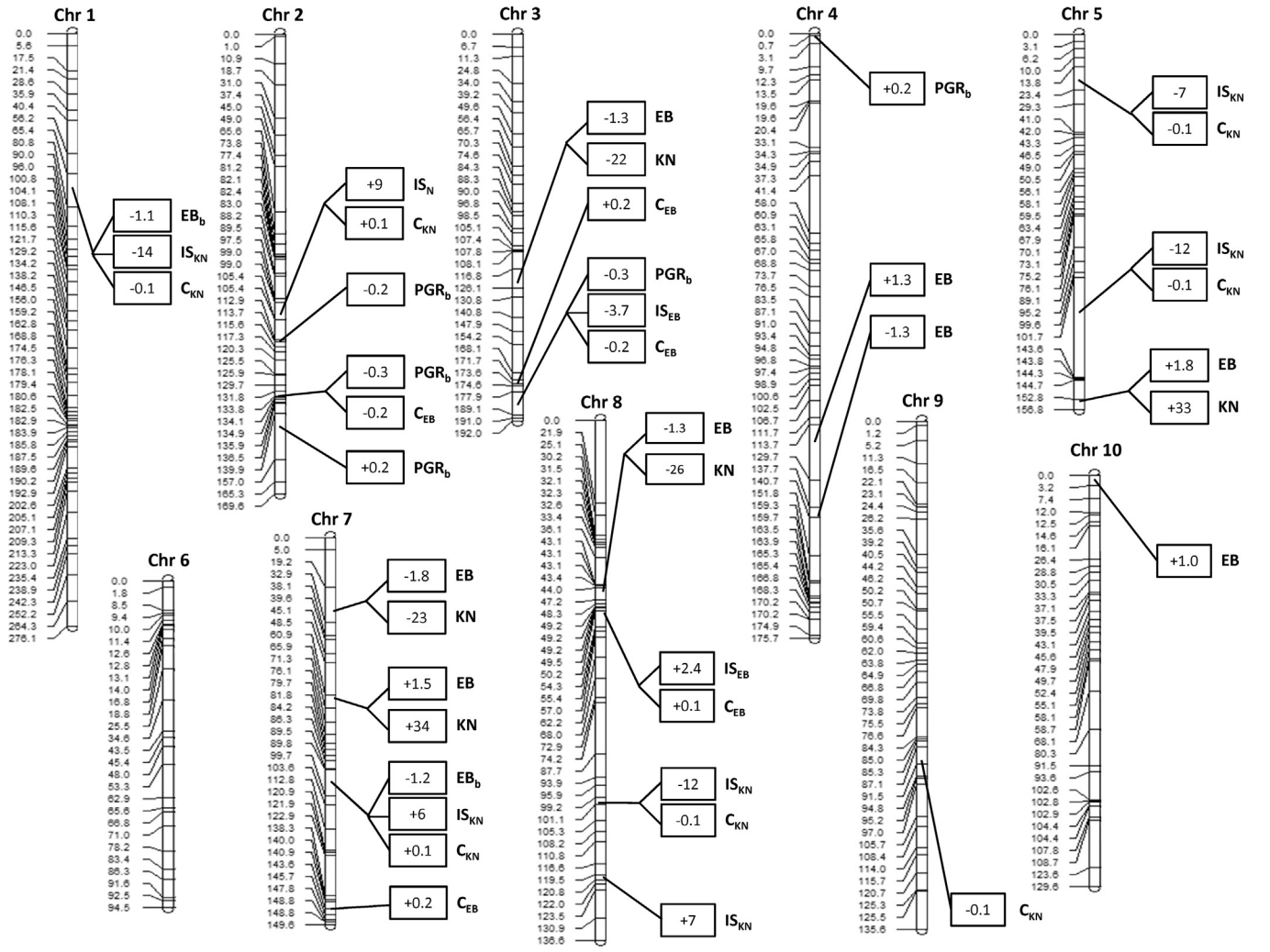

Fig. 4. Chromosomal location of quantitative trait loci (QTL) detected for eight analyzed traits: kernel number per plant (KNP), ear biomass (EB), plant growth rate (PGR) threshold for EB accumulation (PGR_b), initial slope (IS_{EB}) and curvature (C_{EB}) of the relationship of EB vs. PGR, EB threshold for kernel set (EB_b), initial slope (IS_{KN}) and curvature (C_{KN}) of the relationship of KNP vs. EB. QTL are represented by bars with a connector to the corresponding position on the chromosome. Values in each bar show the additive effect of the QTL. Additive effects correspond to B73 allele.

3.4. QTL analysis on response curve parameters of the crop physiology model

The genetic basis of the crop physiology model parameters was studied in terms of QTL. First a multi-trait mapping using CIM and afterwards a MT-MIM analysis were pursued. Only correlated traits were grouped together. For PGR_b, IS_{EB} and C_{EB} analysis, 13 QTL were found. Not all QTL from the final MT-MIM model were significant for individual traits, where maximum marginal peaks were lower than the LOD threshold. However, for these QTL joint analysis was significant (data not shown). Five significant QTL were detected for PGR_b, two for IS_{EB}, and five for C_{EB}. Eleven QTL were detected for EB_b, IS_{KN} and C_{KN}. Again, not all QTL from the final MT-MIM model were significant for individual traits (data not shown). Two significant QTL were detected for EB_b, seven for IS_{KN} and seven for C_{KN} (Fig. 4 and Table 4).

Percentage of explained genetic variation for each trait by the final MT-MIM model ranged from 17 to 39% (Table 4). Importantly, mostly all detected QTL for KNP and EB did not localize at the same positions were model parameter QTL were detected (Fig. 4 and Table 4).

Last, several detected QTL showed close positions together with opposite effects. This is the case for QTL located on chromosomes 2 (position 134 and 140 cM), 3 (position 174 and 183 cM), 4 (position 120 and 140 cM), 7 (position 40 and 61 cM) and 8 (position 101 and 120 cM). Previous evidences using MIM have described similar results (Kao et al., 1999; Hernandez-Valladares et al., 2004; Balint-Kurti et al., 2007). The basis of this is MIM being a powerful and precise method for QTL mapping, detecting closely linked QTL that in other methods remain undetected.

3.5. Predicted vs. observed EB and KNP

The value of two prediction approaches for EB and KNP was tested. The first approach involved using EB and KNP QTL information from an analysis on these final traits per se. Our second approach involved using QTL information for genotype specific parameters of a crop physiology model. The predictive value of both approaches was tested using five independent data sets (Table 5).

First, we used RILs from Exps I and II that were not included in the QTL analysis (44 random genotypes at Exp I and 115 random genotypes at Exp II). Second, we conducted independent experiments (Exps III and IV) at other growing environments (two years and two stand densities) using genotypes that had been, or not, previously used. Last experiment (Exp V) included the parents of the population (B73 and Mo17) grown at two contrasting stand

Table 5
Phenotypic information of independent data used for testing the accuracy of EB and KNP predictions.

Exp		n	PGR (g pl^{-1} d^{-1})	PGR_{SD} (g pl^{-1} d^{-1})	Ear biomass (g pl^{-1})	Kernel number (kernels pl^{-1})
I	Average	44	3.4	0.7	17.9	381
	Min		1.6	0.3	5.5	131
	Max		5.1	1.2	43.7	564
	Gen		***(0.5)[a,b]	****(0.2)*	****(4.1)*	****(64)*
II	Average	115	2.3	0.7	8.0	197
	Min		1.5	0.4	0.8	5
	Max		3.6	1.5	17.5	359
	Gen		****(0.5)*	****(0.2)*	****(2.7)*	****(70)*
III	Average	26	3.7	0.7	12.5	288
	Min		2.6	0.3	2.4	64
	Max		5.1	1.2	24.7	571
	Gen		**	****(0.2)*	*ns*	****(74)*
	Dens		***	**(0.1)*	****(2.1)*	****(29)*
	*Gen*Dens*		**(0.5)*	*ns*	*ns*	*ns*
IV	Average	20	3.3	0.6	17.7	346
	Min		2.1	0.4	8.1	215
	Max		4.8	0.9	31.8	573
	Gen		****(0.3)*	*ns*	****(3.4)*	***
	Dens		****(0.2)*	***(0.1)*	****(1.5)*	***
	*Gen*Dens*		*ns*	*ns*	*ns*	***(87)*
V	Average	4	4.4	0.7	18.9	478
	Min		3.1	0.4	6.9	345
	Max		6.4	1.1	38.0	738
	Gen		**	*ns*	***	***
	Dens		***	*ns*	***	***
	*Gen*Dens*		***(0.4)*	*ns*	***(4.7)*	****(53)*

[a] ns: not significant ($p > 0.05$); * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.
[b] LSD values ($p < 0.05$).

densities. As such, a total of 209 data points provided a wide range of genotypes and environments for testing our two prediction approaches.

Ear biomass and KNP significantly varied among genotypes and between stand densities within each experiment (Table 5). Genotype × stand density interaction was found for EB in Exp V and for KNP in Exps IV and V. Considering the entire data set, EB ranged from 0.8 to 43.7 g pl^{-1} and KNP ranged from 5 to 738 kernels pl^{-1}.

Plant growth rate and PGR_{SD} were inputs for the prediction based on QTL for the response curve parameters of the crop physiology model, and also varied among genotypes and stand densities. A genotype × stand density interaction was significant ($p < 0.01$) for PGR in Exps III and IV. Across experiments, PGR varied from 1.5 to 6.4 g pl^{-1} d^{-1}, and PGR_{SD} varied from 0.3 to 1.5 g pl^{-1} d^{-1}.

Fig. 5 describes the relationship between predicted and observed values for EB and KNP calculated using different approaches. Simulated data were discriminated between genotypes that were coincident with the 125 RILs used for parameter determination but tested at different environments, and the rest of the genotypes. Predicted values of accumulated EB and KNP using genetic information from QTL analysis over traits per se were never able to accurately estimate the same RILs used in the QTL analysis at a different growth environment, or other RILs from the same population under different environments (Fig. 5A and B and Table 6).

Predictions of EB and KNP using genetic information from QTL analysis on response curve parameters of the crop physiology model and using $PGR_{50\%}$ and PGR_{SD} data as inputs improved the estimations overall accuracy (Fig. 5C and D; Table 6). Predictions of EB and KNP, testing either the same or different RILs used in the QTL analysis grown at a range of environments were more accurate and significant ($p < 0.001$) than predictions based on QTL information from final traits per se (Table 6).

Despite predictions involving QTL information for model parameters increased the overall accuracy when compared with predictions using QTL data from traits per se, the use of specific parameter values for each RIL was not the best option. Using a single average value for each parameter for the entire RIL population and considering only the specific RIL $PGR_{50\%}$ and PGR_{SD} data as inputs showed the highest KNP and EB prediction accuracy ($r^2 = 0.46$, $p < 0.001$ and $r^2 = 0.37$, $p < 0.001$, for EB and KNP, respectively; Table 6; Fig. 5E and F). This result showed that the increased prediction accuracy when using QTL information on model parameters was mostly related to incorporating genotypic differences in plant growth.

4. Discussion

We have determined the genetic basis of two traits of interest (EB and KNP) doing a QTL analysis on (i) traits per se and (ii) response curve parameters that describe well-known relations between KNP and EB to plant growth around flowering. These curve parameters have shown significant variation among genotypes in previous studies (Tollenaar et al., 1992; Andrade et al., 1999; Echarte et al., 2004; Borrás et al., 2007; Pagano and Maddonni, 2007; D'Andrea et al., 2008; Borrás et al., 2009). Here, we phenotyped a RIL population and observed large variation for these parameters. Importantly, the genetic basis of traits per se and physiological parameters describing the response of these traits to the environment were different.

Physiological dissection and modeling of traits provide an avenue by which crop growth models could contribute to integrate molecular genetic technologies and crop improvement (Hammer et al., 2006; Chapman et al., 2003). Key components of the process are trait understanding and modeling, followed by determination of the genetic architecture of adaptive traits. Messina et al. (2011) applied this genotype-to-phenotype approach for multiple traits and simulated performance landscapes within an operational maize breeding program. Their crop model includes algorithms to

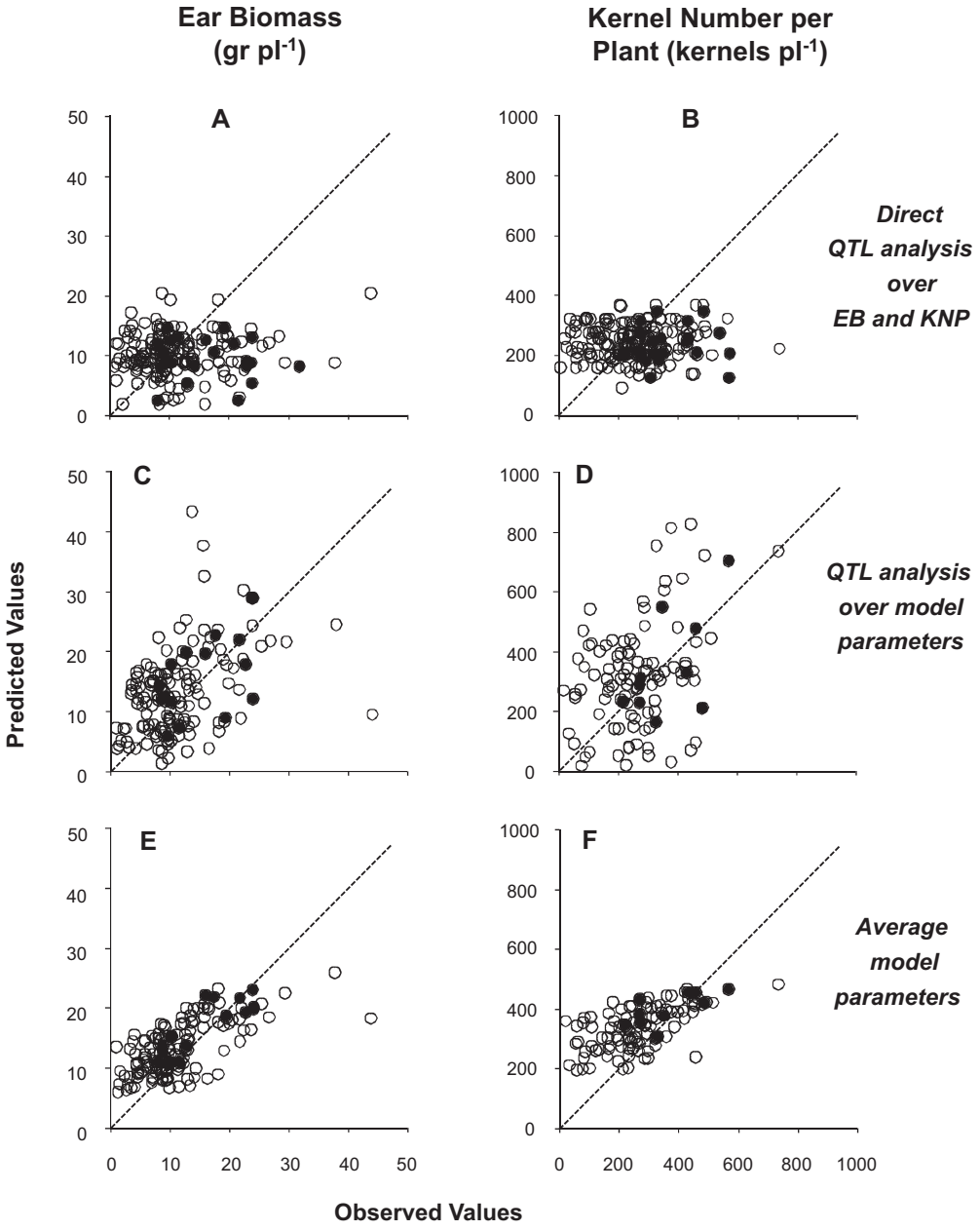

Fig. 5. Relationship between predicted and observed accumulated ear biomass 15 days after anthesis and kernel number per plant. Predictions were done following three different approaches. (A) and (B) show predicted ear biomass (EB; (A)) and kernel number per plant (KNP; (B)) based on QTL information of traits per se. (C) and (D) show predicted EB (C) and KNP (D) with QTL information of parameters of the proposed crop physiology model (see Fig. 1 and Section 2 for further details). (E) and (F) show predicted EB (E) and KNP (F) with average parameters values of the proposed crop physiology model. White circles represent RILs that were not considered in QTL analysis. Black circles represent RILs that were used at the QTL analysis but were grown at other growth environments. The dash line represents the 1:1 relation.

Table 6
Coefficient of determination and significance of linear regression between observed and estimated ear biomass (EB) and kernel number per plant (KNP). Three different predictions are compared: using QTL information of final traits, using average model parameters or using QTL of model parameters. Data are classified in: (i) coincident 125 RILs used in the QTL analysis but grown at other environments, (ii) different RILs from the same population that were not used in the QTL analysis and grown at other environments and (iii) coincident and different RILs used for the QTL analysis all together.

Data	EB QTL	EB Model parameter QTL	EB Average model parameters	KNP QTL	KNP model parameter QTL	KNP Average model parameters
Coincident RILs (n: 22)	r^2: <0.01; ns[a]	r^2:0.27; *	r^2:0.68; ***	r^2: <0.01; ns	r^2:0.28; *	r^2:0.45; ***
Different RILs (n: 185)	r^2: <0.01; ns	r^2:0.10; ***	r^2:0.41; ***	r^2: <0.01; ns	r^2:0.10; ***	r^2:0.33; ***
All RILs together (n: 207)	r^2: <0.01; ns	r^2:0.13; ***	r^2:0.46; ***	r^2: <0.01; ns	r^2:0.12; ***	r^2:0.37; ***

[a] ns: not significant ($p > 0.05$); * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

simulate kernel number grounded on similar physiological perspectives applied here (Messina et al., 2009). We hypothesized that the conceptual model described in Fig. 1 was an opportunity to predict EB and KNP variations.

One of the recognized benefits of studying the genetic basis of a particular trait based on a crop physiology model approach is that GxE interaction effects are implicitly considered (Hunt et al., 1993; Messina et al., 2011). This is because current crop growth models are structured to predict consequences of genotype × environment × management interactions (Hammer et al., 2006). Studying the genetic basis of a quantitative trait per se (in terms of QTL) in different environments generally leads to inconsistent results in most cases. Literature contains several examples of this kind for KNP or similar traits in maize (Agrama et al., 1999; Beavis et al., 1994; Ribaut et al., 1997; Frova et al., 1999; Peng et al., 2011). In our case, environmental or management variables would affect final EB or KNP through modifications on PGR. Genotype × environment interactions are an emergent property of the considered model (Hammer et al., 2005).

We compared doing a QTL study on parameters of a crop physiology model with a traditional approach for studying the genetic basis of KNP, which is QTL analysis on the trait per se. Including genetic variation on response curve parameters (i.e., QTL information) did increase EB and KNP prediction accuracy. The increase was mostly related to incorporating specific genotype plant growth data rather than incorporating genotype specific parameters describing biomass partitioning and kernel set differences (Fig. 5). This result also indicates that, overall, genotype specific plant growth information might be more important than genotype specific parameter values related to biomass partitioning and kernel set efficiency. In relation to this, Hernández et al. (2014) found that kernel set differences at high stand density conditions among current commercial hybrids from Argentina are not only related to differences in biomass partitioning but also to plant growth at stressful conditions.

The reduced accuracy when combining a crop physiology model and QTL information could be related to the accumulation of several sources of error, from experiments, modeling and QTL estimates (Uptmoor et al., 2008). Regarding the latter, average estimates of phenotypic variances associated with identified QTL could be overestimated when low number of individuals is evaluated (i.e., Beavis effect; Beavis, 1994; Xu, 2003). Another possibility for the small accuracy of EB and KNP estimations when doing the QTL analysis over model parameters is the model we are currently using. Determining the genetic basis of model parameters instead of final trait per se is an approach that has been previously used. Reymond et al. (2003), Yin et al. (2005), and Uptmoor et al. (2008) applied this approach for other traits (leaf expansion rate in maize, time to flowering in barley and time to floral induction and flowering in *Brassica oleracea*, respectively). Predictions made with measured parameters were more accurate than predictions made with QTL-based estimates of parameters. However, this was not always the case. Yin et al. (2000) tested the ability of a crop growth model to predict yield variations in a set of RILs in barley. In this study neither yield predictions with QTL-based estimates of parameters nor measured parameters were accurate, suggesting the physiology behind the model was inappropriate. Although we are not able to test this particular comparison here, it is important to identify the possibility that our model is not robust enough for the application we are pursuing.

Yin et al. (2005) proposed that the gain from QTL statistics that removes part of random noise from original model input traits may not be sufficient to compensate for the loss due to residual genetic variance that is not captured by the identified QTL. Besides, the phenotypic variation accounted by the identified QTL in these studies was generally low (as in our case, see Table 4). Wang et al. (2012) compared QTL analysis for measured and mathematically derived traits and concluded that the use of derived traits in QTL mapping may increase the QTL number and decrease the proportion of phenotypic variance than can be explained by component QTL. The increased complexity of the genetic architecture of derived traits reduces QTL detection, and increases the false discovery rate (Wang et al., 2012). However, they also recognize that their detection power was reduced for derived traits, but almost unbiased position estimations were achieved.

Finally, we have described the existence of genotypic variability in most eco-physiological mechanisms involved in maize kernel set within a RIL population. Genotypic variability for such mechanisms was known among inbred lines (Echarte and Tollenaar, 2006), but information regarding parental inbred line to derived hybrid correlation for parameters of interest, and inheritance of the traits we analyzed, is limited. Such information exists for N response (D'Andrea et al., 2013), and is relevant for hybrid development. Any information on parental inbred lines that is indicative of derived hybrid performance is highly desirable. Munaro et al. (2011) found heterosis is important for plant growth and ear biomass accumulation, but at present we do not know the effect over our crop physiology parameter traits.

5. Conclusions

Key issues emerging from this study are:

- Several QTL for ear biomass accumulation and KNP were identified. Positive associations were found between additive effects of EB and KNP QTL when located at similar positions, confirming the dependence of KNP to ear biomass.
- QTL of well-documented response curve parameters that describe the responses of KNP and EB to PGR were identified. The detected QTL for the different model parameters did not localize at the same positions as QTL detected for EB and KNP per se.
- We were not able to predict EB or KNP at independent experiments using detected QTL for these traits directly.
- Using QTL information for model parameters describing the response curve between EB, KNP, and PGR helped predict EB and KNP of independent experiments, indicating the value of the approach.
- Despite including genotype specific parameter values estimated with QTL information helped with KNP and EB prediction accuracy, using an average value for each parameter was a better option. This indicated that the increased accuracy when compared to the QTL per se approach was mostly related to incorporating plant growth data rather than incorporating genotype specific parameters describing biomass partitioning and kernel set differences.
- Our results are an important starting point for studies related to the genetic determination of maize KNP and its physiological determinants. Specific chromosome regions were described that deserve further dissection.

Acknowledgments

Authors thank M Lee for providing the seed used in the present study, S Alvarez Prado for helping with data analysis, D Novoa and Nidera S.A. for helping with seed increases, and many undergraduate students from UNR that helped with field data collection. A Amelong held a graduate scholarship from CONICET, the Scientific Research Council of Argentina. L Borrás and BL Gambín are members of CONICET.

References

Abertondo, V.J., 2007. Phenotypic Analysis of Intermated B73 × Mo17 (IBM) Population. Iowa State University, USA (M.Sc. Thesis).

Agrama, H.A.S., Zakaria, A.G., Said, F.B., Tuinstra, M., 1999. Identification of quantitative trait loci for nitrogen use efficiency in maize. Mol. Breed. 5, 187–195.

Alvarez Prado, S., López, C.G., Gambín, B.L., Abertondo, V.J., Borrás, L., 2013. Dissecting the genetic basis of physiological processes determining maize kernel weight using IBM (B73 × Mo17) Syn4 population. Field Crops Res. 145, 33–43.

Andrade, F.H., Vega, C.R.C., Uhart, S., Cirilo, A., Cantarero, M., Valentinuz, O., 1999. Kernel number determination in maize. Crop Sci. 39, 453–459.

Balint-Kurti, P.J., Zwonitzer, J.C., Wisser, R.J., Carson, M.L., Oropeza-Rosas, M.A., Holland, J.B., Szalma, S.J., 2007. Precise mapping of quantitative trait loci for resistance to southern leaf blight, caused by Cochliobolus heterostrophus race O, and flowering time using advanced intercross maize lines. Genetics 176, 645–657.

Basten, C.J., Weir, B.S., Zeng, Z.B., 2004. QTL Cartographer, Version 1.17. Department of Statistics, North Carolina State University, Raleigh, NC.

Bernardo, R., 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci. 48, 1649–1664.

Beavis, W.D., Smith, O.S., Grant, D., Fincher, R., 1994. Identification of quantitative trait loci using a small sample of topcrossed and F_4 progeny from maize. Crop Sci. 34, 882–896.

Beavis, W.D., 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. In: Proc. Corn Sorghum Ind. Res. Conf., 7–8 Dec. 1994, Chicago, IL. Am. Seed Trade Assoc., Washington, DC, pp. 250–266.

Blanc, G., Charcosset, A., Mangin, B., Gallais, A., Moreau, L., 2006. Connected populations for detecting a quantitative trait loci and testing for epistasis: an application in maize. Theor. Appl. Genet. 113, 206–224.

Borrás, L., Westgate, M.E., Astini, J.P., Echarte, L., 2007. Coupling time to silking with plant growth rate in maize. Field Crops Res. 102, 73–85.

Borrás, L., Astini, J.P., Westgate, M.E., Severini, A., 2009. Modeling anthesis to silking in maize using a plant biomass framework. Crop Sci. 49, 937–948.

Chapman, S.C., Edmeades, G.O., 1999. Selection improves drought tolerance in tropical maize populations II. Direct and correlated responses among secondary traits. Crop Sci. 39, 1315–1324.

Chapman, S., Cooper, M., Podlich, D., Hammer, G., 2003. Evaluating plant breeding strategies by simulating gene action and dryland environment effects. Agron. J. 95, 99–113.

D'Andrea, K.E., Otegui, M.E., Cirilo, A.G., 2008. Kernel number determination differs among maize hybrids in response to nitrogen. Field Crops Res. 105, 228–239.

D'Andrea, K.E., Otegui, M.E., Cirilo, A.G., Eyhérabide, G.H., 2013. Parent–progeny relationships between maize inbreds and hybrids: analysis of grain yield and its determinants for contrasting soil nitrogen conditions. Crop Sci. 53, 2147–2161.

Di Rienzo, J.A., Casanoves, F., Balzarini, M.G., Gonzalez, L., Tablada, M., Robledo, C.W., 2013. InfoStat versión 2013. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina.

Dudley, J.W., Johnson, G.R., 2009. Epistatic models improve prediction of performance in corn. Crop Sci. 49, 763–770.

Early, E.B., McIlrath, W.O., Seif, R.D., 1967. Effects of shade applied at different stages of plant development on corn (*Zea mays* L.) production. Crop Sci. 7, 151–156.

Echarte, L., Andrade, F.H., Vega, C.R.C., Tollenaar, M., 2004. Kernel number determination in Argentinean maize hybrids released between 1965 and 1993. Crop Sci. 44, 1654–1661.

Echarte, L., Tollenaar, M., 2006. Kernel set in maize hybrids and their inbred lines exposed to stress. Crop Sci. 46, 870–878.

Edmeades, G.O., Daynard, T.B., 1979. The relationship between final yield and photosynthesis at flowering in individual maize plants. Can. J. Plant Sci. 59, 585–601.

Fiedler, K., Bekele, W.A., Friedt, W., Snowdon, R., Stützel, H., Zacharias, A., Uptmoor, R., 2012. Genetic dissection of the temperature dependent emergence process in sorghum using a cumulative emergence model and stability parameters. Theor. Appl. Genet. 125, 1647–1661.

Frova, C., Krajewski, P., di Fonzo, N., Villa, M., Sari-Gorla, M., 1999. Genetic analysis of drought tolerance in maize by molecular markers I. Yield components. Theor. Appl. Genet. 99, 280–288.

Gambín, B.L., Borrás, L., Otegui, M., 2006. Source-sink relations and kernel weight differences in maize temperate hybrids. Field Crops Res. 95, 316–326.

Gonzalo, M., Holland, J.B., Vyn, T.J., McIntyre, L.M., 2010. Direct mapping of density response in a population of B73xMo17 recombinant inbred lines of maize (*Zea Mays* L.). Heredity 104, 583–599.

Haldane, J.B.S., 1919. The combination of linkage values, and the calculation of distance between the loci of linked factors. J. Genet. 8, 299–309.

Hallauer, A.R., Carena, M.J., Miranda Filho, J.B., 2010. Quantitative Genetics in Maize Breeding, third ed. Springer, New York, NY.

Hammer, G.L., Chapman, S., van Oosterom, E., Podlich, D.W., 2005. Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems. Aust. J. Agric. Res. 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 Plant Sci. 11, 587–593.

Hernández, F., Amelong, A., Borrás, L., 2014. Genotypic differences among Argentinean maize hybrids in yield response to stand density. Agron. J. 106, 2316–2324.

Hernandez-Valladares, M., Naessens, J., Gibson, J.P., Musoke, A.J., Nagda, S., Rihet, P., ole-MoiYoi, O.K., Iraqui, F.A., 2004. Confirmation and dissection of QTL controlling resistance to malaria in mice. Mamm. Genome 15, 390–398.

Hunt, L.A., Pararajasingham, S., Jones, J.W., Hoogenboom, G., Imamura, D.T., Ogoshi, R.M., 1993. GENECALC. Software to facilitate the use of crop models for analyzing field experiments. Agron. J. 85, 1090–1094.

Hunter, J.L., Tekrony, D.M., Miles, D.F., Egli, D.B., 1991. Corn seed maturity indicators and their relationship to uptake of C-14 assimilate. Crop Sci. 31, 1309–1313.

Kao, C.H., Zeng, Z.B., Teasdale, R.D., 1999. Multiple interval mapping for quantitative trait loci. Genetics 152, 1203–1216.

Kearsey, M.J., Pooni, H.S., 1996. The Genetical Analysis of Quantitative Traits. Chapman and Hall, London.

Lee, M., Sharapova, N., Beavis, W.D., Grant, D., Katt, M., Blair, D., Hallauer, A., 2002. Expanding the genetic map of maize with the intermated B73 × Mo17 (*IBM*) population. Plant Mol. Biol. 48, 453–461.

Lorieux, M., 2007. Map Disto, a free user-friendly program for computing genetic maps. In: Computer Demonstration (P958) Given at the Plant and Animal Genome XV Conference, Jan 13–17 2007, San Diego, CA.

Lübberstedt, T., Klein, D., Melchinger, A.E., 1998. Comparative QTL mapping of resistance to *Ustilago maydis* across four populations of European flint-maize. Theor. Appl. Genet. 97, 1321–1330.

Malosetti, M., Ribaut, J.M., Vargas, M., Crossa, J., van Eeuwijk, F.A., 2008. A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). Euphytica 161, 241–257.

Messina, C.D., Hammer, G., Dong, Z., Podlich, D., Cooper, M., 2009. Modelling crop improvement in a GxExM framework via gene–trait–phenotype relationships. In: Sadras, V., Calderini, D. (Eds.), Crop Physiology: Interfacing with Genetic Improvement and Agronomy. Elsevier, Amsterdam, pp. 235–265.

Messina, C.D., Podlich, D., Dong, Z., Samples, M., Cooper, M., 2011. Yield-trait performance landscapes: from theory to application in breeding maize for drought tolerance. J. Exp. Bot. 62, 855–868.

Munaro, E.M., D'Andrea, K.E., Otegui, M.E., Cirilo, A.G., Eyhérabide, G.H., 2011. Heterotic response for grain yield and ecophysiological related traits to nitrogen availability in maize. Crop Sci. 51, 1172–1187.

Otegui, M.E., 1995. Prolificacy and grain yield components in modern Argentinean maize hybrids. Maydica 40, 371–376.

Otegui, M.E., Bonhomme, R., 1998. Grain yield components in maize I. Ear growth and kernel set. Field Crops Res. 56, 247–256.

Pagano, E., Maddonni, G.A., 2007. Intra-specific competition in maize: early established hierarchies differ in plant growth and biomass partitioning to the ear around silking. Field Crops Res. 101, 306–320.

Peng, B., Li, Y., Wang, Y., Cheng, L., Liu, Z., Tan, W., Zhang, Y., Wang, D., Shi, Y., Sun, B., Song, Y., Wang, T., Li, Y., 2011. QTL analysis for yield components and kernel-related traits in maize across multi-environments. Theor. Appl. Genet. 122, 1305–1320.

Rahman, M.M., Govindarajulu, Z., 1997. A modification of the test of Shapiro and Wilk for normality. J. Appl. Stat. 24, 219–235.

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

Ribaut, J.M., Jiang, C., Gonzales-de-Leon, D., Edmeades, G.O., Hoisington, D.A., 1997. Identification of quantitative trait loci under drought conditions in tropical maize 2. Yield components and marker-assisted selection strategies. Theor. Appl. Genet. 94, 887–896.

SAS Institute, 1999. The SAS Online Doc v.8. SAS Inst., Cary, NC.

Severini, A.D., Borrás, L., Westgate, M.E., Cirilo, A.G., 2011. Kernel number and kernel weight determination in dent and popcorn maize. Field Crops Res. 120, 360–369.

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., et al., 2009. The B73 genome: complexity, diversity, and dynamics. Science 326, 1112–1115.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Tanksley, S.D., 1993. Mapping polygenes. Annu. Rev. Genet. 2, 205–233.

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

Tollenaar, M., Dwyer, L.M., Stewart, D.W., 1992. Ear and kernel formation in maize hybrids representing three decades of grain yield improvement in Ontario. Crop Sci. 32, 432–438.

Troyer, A.F., 1999. Background of U.S. hybrid corn. Crop Sci. 39, 601–626.

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

Van Ooijen, J.W., 1999. LOD significance threshold for QTL analysis in experimental populations of diploid species. Heredity 83, 613–624.

Vega, C.R.C., Sadras, V.O., Andrade, F.H., Uhart, S., 2000. Reproductive allometry in soybean maize and sunflower. Ann. Bot. 85, 461–468.

Vega, C.R.C., Andrade, F.H., Sadras, V.O., 2001. Reproductive partitioning and seed set efficiency in soybean sunflower and maize. Field Crops Res. 72, 163–175.

Wang, S., Basten, C.J., Zeng, Z.B., 2010. Windows QTL Cartographer Version 2.5_004.Statistical Genetics. North Carolina State University, USA.

Wang, Y., Li, H., Zhang, L., Lü, W., Wang, J., 2012. On the use of mathematically-derived traits in QTL mapping. Mol. Breed. 29, 661–673.

Wu, W., Zhou, Y., Li, W., Mao, D., Chen, Q., 2002. Mapping of quantitative trait loci based on growth models. Theor. Appl. Genet. 105, 1043–1049.
Xu, S., 2003. Theoretical basis of the Beavis effect. Genetics 165, 2259–2268.
Yin, X., Chasalow, S.D., Dourleijn, C.J., Stam, P., Kropff, M.J., 2000. 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., Struik, P.C., van Eeujwijk, F.A., Stam, P., Tang, J., 2005. QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley. J. Exp. Bot. 56, 967–976.
Zhang, L., Li, H., Wang, J., 2012. The statistical power of inclusive composite interval mapping in detecting digenic epistasis showing common F2 segregation ratios. J. Integr. Plant Biol. 54, 270–279.