QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley

Xinyou Yin¹,²*, Paul C. Struik², Fred A. van Eeuwijk³, Piet Stam³ and Jianjun Tang¹

¹ College of Agriculture, Jiangxi Agricultural University, Nanchang 330045, PR China
² Crop and Weed Ecology Group, Wageningen University, PO Box 430, 6700 AK Wageningen, The Netherlands
³ Laboratory of Plant Breeding, Wageningen University, PO Box 386, 6700 AJ Wageningen, The Netherlands

Received 24 August 2004; Accepted 22 November 2004

Abstract

Combining ecophysiological modelling and genetic mapping has increasingly received attention from researchers who wish to predict complex plant or crop traits under diverse environmental conditions. The potential for using this combined approach to predict flowering time of individual genotypes in a recombinant inbred line (RIL) population of spring barley (Hordeum vulgare L.) was examined. An ecophysiological phenology model predicts preflowering duration as affected by temperature and photoperiod, based on the following four input traits: f_o (the minimum number of days to flowering at the optimum temperature and photoperiod), θ_1 and θ_2 (the development stages for the start and the end of the photoperiod-sensitive phase, respectively), and δ (the photoperiod sensitivity). The model-input trait values were obtained from a photoperiod-controlled greenhouse experiment. Assuming additivity of QTL effects, a multiple QTL model was fitted for the model-input traits using composite interval mapping. Four to seven QTL were identified for each trait. Each trait had at least one QTL specific to that trait alone. Other QTL were shared by two or all traits. Values of the model-input traits predicted for the RILs from the QTL model were fed back into the ecophysiological model. This QTL-based ecophysiological model was subsequently used to predict preflowering duration (d) for eight field trial environments. The model accounted for 72% of the observed variation among 94 RILs and 94% of the variation among the two parents across the eight environments, when observations in different environments were pooled. However, due to the low percentage (34–41%) of phenotypic variation accounted for by the identified QTL for three model-input traits (θ_1, θ_2 and δ), the QTL-based model accounted for somewhat less variation among the RILs than the model using original phenotypic input trait values. Nevertheless, days to flowering as predicted from the QTL-based ecophysiological model were highly correlated with days to flowering as predicted from QTL-models per environment for days to flowering per se. The ecophysiological phenology model was thus capable of extrapolating (QTL) information from one environment to another.

Key words: Ecophysiological modelling, flowering time, genotype–phenotype relationship, Hordeum vulgare L., model-input traits, photoperiod, quantitative trait loci, temperatures.

Introduction

The advent of various molecular markers has enabled the unravelling of complex crop traits into the effect of individual quantitative trait loci, QTL (Paterson et al., 1988). This has opened up opportunities to enhance the efficiency of plant breeding by the use of marker-assisted selection (Ribaut and Hoisington, 1998) and for concepts like ‘breeding by design’ (Peleman and van der Voort, 2003). However, QTL-by-environment interactions (QTL×E), i.e. when QTL expression is conditional on the environment, can be a major obstacle in the application of marker-assisted selection for manipulating complex traits of agronomic importance (Ribaut and Hoisington, 1998; Slafers, 2003). In the classical model for genotype-by-environment interactions (G×E), the regression on the mean model (Finlay and Wilkinson, 1963), the mean phenotypic value across genotypes per environment is used as a measure of environmental quality, the environmental index (Kearsey...
and Pooni, 1996). This approach does not relate phenotypic values to physical measures of the environment such as temperature. A reason to use the regression on the mean model might be that climatic and edaphic factors are hard to obtain, and therefore the plants’ average performance in a particular environment is taken as an indicator of environmental quality. In the Finlay-Wilkinson model, G X E is represented by genotypic differences in sensitivity to the environmental index. The approach provides a summary of phenotypic data containing G X E and models the phenotypic responses in relation to the environment. However, the Finlay–Wilkinson approach does not allow a straightforward prediction of phenotypic responses for environments outside the set of environments used for calibrating the model.

With the development of automated weather station networks in the early 1980s and the availability of soil maps and satellite data, data for climatic and edaphic variables become now widely available (Weiss, 2003). To use the information of physical environments, the factorial regression was proposed as an ordinary linear model that allows the G X E to be modelled directly as a function of environmental variables (van Eeuwijk et al., 1996). The factorial regression approach has been adopted to analyse QTL X E, whereby environmental characterizations (often average or accumulated values over certain phases) are used as co-variables in models allowing for environment-dependent QTL expression (van Eeuwijk et al., 2002). The success of this approach relies on whether the correct physical environmental factors are included and whether incorporated values of these environmental factors match the relevant growth periods. Obviously, the correct choice of physical environmental factors and their values requires the knowledge of crop physiology for the traits under study. Even when the choice is correctly made, the power of the factorial regression to describe genotypic differences in relation to the environment can be limited because complex crop traits are often the result of non-linear interactions between genetic, environmental and managerial factors on multiple component processes, integrated over ontogenetic stages. The inability of the factorial regression to flexibly incorporate managerial factors and temporal (both seasonal and diurnal) dynamics over the growing season and spatial profiles of environmental variables within a crop canopy is the major limitation to it being used to resolve G X E on a biological basis.

The aim was to bring the link between crop physiology and genetics a step further, focusing on the G X E problem while studying genotype-to-phenotype relationships. This is possible because of the availability of ecophysiology-based crop growth models that relate elementary processes to environmental variables. In the context of QTL analysis, the complex trait was first dissected into component traits by the use of an ecophysiological model and then, instead of searching for QTL for the complex trait itself, QTL models for those component traits were found. The component traits often correspond to model-input parameters, reflecting effects with a genetic component. Another category of inputs for ecophysiological models is formed by weather, soil variables and management options. The ecophysiological model structure provides algorithms following physiological principles to quantify the interactions between ontogenetic component processes and environmental factors. Unlike factorial regression models, ecophysiological models use daily (or if needed, derived hourly) values of environmental variables during the whole growing season and consider spatial variation of these variables within the crop (e.g. the decline of radiation with the depth of the crop canopy) to drive simulation for predicting phenotypes. In addition, the effects of managerial factors (e.g. quantity and timing of fertilization) can be more flexibly considered as input variables in ecophysiological models.

In the first studies exploring the usefulness of this ecophysiological approach, QTL for various input traits of a model that predicts crop yields (Yin et al., 1999) were identified. QTL-based parameters were then fed back into the model to predict yield performance of individuals in the population under study (Yin et al., 2000a). Although the correlation between predicted yields using QTL-based parameters and those using the original measured, phenotypic parameters was high, the ability of current crop growth models is not yet sufficient to predict differences in complex traits like yield among individuals of a segregating population (Yin et al., 2000a, b).

Following a similar physiology-based approach, Raymond et al. (2003) demonstrated the potential value of combining ecophysiological modelling and genetic mapping in predicting G X E interaction on a simple trait, leaf elongation rate in maize (Zea mays L.). The QTL analysis was performed on parameters of a simple linear model for predicting the leaf elongation rate as affected by temperature and water deficit. It was shown that the rates of individuals, including lines not used for QTL analysis, of the mapping population were well predicted by the combined QTL- and ecophysiological model, for any climatic scenario.

In the present analysis, the same concept is applied to flowering time in barley (Hordeum vulgare L.). An ecophysiological model for phenology was used, where the daily rate of progress towards flowering is modelled as an interactive function of daily temperature and photoperiod (Yin et al., 2005). The QTL analysis has been performed for each of the model-input parameters. The aim was to test whether the flowering time of individuals in the mapping population under a range of field conditions can be predicted by combined use of QTL and ecophysiological modelling.
Materials and methods

Plant material and parameters of the phenology model

The plant material used in the present analysis consists of 94 individuals produced by eight generations of single seed descent from a cross of the two-row spring barley cultivars (‘Apex’ and ‘Prisma’). This is the same recombinant inbred line (RIL) population that was used in earlier research (Yin et al., 1999, 2000a, b).

The ecophysiological phenology model used here was described in detail by Yin et al. (2005), so only summary information is presented here. In this model, daily development rate, \( \dot{o}_h \), at a developmental stage (0) is described as:

\[
\dot{o}_h = \frac{g(T)/f_o}{(g(T)/h(P)/f_o)} \quad 0 < \delta < \theta
\]

with

\[
g(T) = \max \{0, \frac{[35 - T]/14}{(T/21)^{1/2}}\}
\]

(2a)

\[
h(P) = \max \{0, 1 - \delta (17 - P)\}
\]

(2b)

where \( T \) and \( P \) are temperature and photoperiod. The model has four input traits \( f_o, \dot{o}_1, \dot{o}_2, \) and \( \delta \). \( f_o \) defines the minimum number of days from sowing to flowering when both temperature and photoperiod are at their optimum, representing the genotypic intrinsic earliness of flowering. \( \dot{o}_1 \) and \( \dot{o}_2 \) are the (dimensionless) development stage for the start and the end of the photoperiod-sensitive phase, respectively; \( \delta \) is the parameter characterizing the photoperiod-sensitivity during the photoperiod-sensitive phase (h\(^{-1}\)). Values of these traits for each RIL were estimated, via curve-fitting to data from a photoperiod-controlled greenhouse experiment, where plants were mutually transferred between long-day and short-day photoperiods at 10 d intervals (Yin et al., 2005). Because the function \( g(T) \) is non-linear and temperature fluctuates diurnally, \( g(T) \) was estimated on an hourly basis and hourly \( g(T) \) values were averaged to obtain the daily value. The hourly temperature was estimated from daily maximum and minimum temperature by a sine function. In this way, daily developmental rate, \( \dot{o}_h \), was calculated. The day when total daily \( \dot{o}_h \) accumulated from sowing is just above or equal to 1.0 is the predicted time when an RIL flowers.

QTL detection

The QTL analysis was performed on the basis of the marker linkage map established by Yin et al. (1999) for the ‘Apex×Prisma’ RIL population, which contains 191 AFLP markers and one morphological marker, covering a total map length of 965 cM (Fig. 1). A mapping software MapQTL\(^4\) (Van Ooijen, 2002) was used in identifying QTL positions in the genome for a given trait. QTL for each of the four traits \( f_o, \dot{o}_1, \dot{o}_2, \) and \( \delta \) were identified, using a composite interval mapping (or called the MQM mapping) method (Jansen, 1995) as implemented in MapQTL\(^4\). In this method, background markers are selected to take over the role of the putative QTL as cofactors to reduce the residual variance. A two-stage MQM analysis was performed. In the first stage, a conventional interval mapping was performed at a 2 cM interval; the LOD profiles from interval mapping were inspected and the marker closest to each LOD peak was selected as the cofactor to perform the MQM mapping analysis further. The inclusion of cofactors may lead to new peaks in the LOD profiles, which suggested further cofactors to be included in the analysis. Several cycles were performed to obtain the potentially maximum number of cofactors for the MQM analysis. These cofactor markers were then subjected to backward elimination, as implemented in MapQTL\(^4\), to select the best model for the second stage MQM analysis. Such a backward elimination procedure leaves out one cofactor at a time to create a subset of cofactors. The likelihood of each of these subset models is compared with the likelihood of the full model with all cofactors, and the subset model which causes the smallest change in likelihood is chosen as the starting set for a subsequent round of elimination. This process continues until the change in likelihood is significant according to the 0.05 P-value for the test. The then retained set of cofactors was used in the second stage of the MQM analysis. In the final LOD profile, QTL were declared according to the threshold LOD scores ranging from 2.8 to 3.0 (genome-wide false-positives rate 5%), depending on chromosome map length and the number of chromosome pairs (Van Ooijen, 1999).

Predicting flowering time using QTL-based parameters of the ecophysiological model

The authors wanted to test whether the flowering time could be accurately predicted on the basis of identified QTL for each of the model-input traits \( f_o, \dot{o}_1, \dot{o}_2, \) and \( \delta \). To this end, QTL-based values for each of the traits had to be estimated for each RIL. Because in MapQTL\(^4\), only additive gene actions are assumed for the RIL populations, no attempt was made to model epistatic QTL in this analysis. The genetic (additive effect) predictor of the \( i \)-th QTL genotype for the \( j \)-th RIL, \( g_{ij} \), is obtained from the conditional allelic probabilities at QTL positions given the information at flanking markers (Jiang and Zeng, 1995, 1997), according to:

\[
g_{ij} = Z_{gi}P(QTL)=G1[M1i, M2i]+Z_{gi}P(QTL)=G2[M1i, M2i]
\]

where \( Z_{gi} \) and \( Z_{g2} \) are indicator variables (defined as 1 and −1 for genotypes corresponding to ‘Prisma’ (G1) and ‘Apex’ (G2), respectively), and \( QTL_p= G1[M1i, M2i] \) and \( P(QTL_p= G2[M1i, M2i]) \) are the probabilities of the \( i \)-th QTL of the \( j \)-th RIL to be of genotype G1 and G2, respectively, conditional on the flanking marker genotypes \( M1i \) and \( M2i \).

The trait QTL loci identified within the MapQTL\(^4\) modelling exercises, served as the basis for a last cycle of modelling in which multiple regression models were fitted to trait responses using the genetic predictors, \( g_{ij} \), corresponding to the earlier identified QTL. From these regression models, predicted trait values were produced. The predicted value for the \( j \)-th RIL, \( y_j \), is then:

\[
y_j = \dot{m} + \sum_{i=1}^n \dot{a}_ig_{ij}
\]

where \( \dot{m} \) is the estimated intercept, and \( \dot{a}_i \) is the estimated additive effect of the \( i \)-th QTL on the trait \((i=1, 2, \ldots, n)\). The coefficient of determination for equation (4), a measure for goodness of fit, is defined by the total percentage of phenotypic variation in the model-input trait that is accounted for by the QTL.

Predicted days to flowering using the QTL-based model input traits were compared with predictions using the observed phenotypic input traits. The comparison included the 94 RILs under eight independent field conditions during two growing seasons, created by using different sowing dates (Yin et al., 2005). Such a comparison was also performed for the two parents, which provided an additional independent test, since the QTL analysis used no information of either marker genotype or trait phenotype of the parents. Finally, predicted days to flowering by the QTL-based ecophysiological model were compared with those predicted by identified QTL per field environment for days to flowering per se.

Results and discussion

The mean, minimum and maximum trait values, and the phenotypic correlation coefficients between the four traits are shown in Table 1. All correlations were statistically
significant \((P < 0.01)\), and those between \(\theta_1\) and \(\delta\), and between \(\theta_2\) and \(\delta\) were particularly so. Values of \(\theta_2\), \(\theta_3\), and \(\delta\) were very similar for the two parents and the frequency distribution of trait values showed transgressive segregation (Fig. 2), which indicates that alleles of increasing and decreasing effect are dispersed over the parents, ‘Apex’ and ‘Prisma’. An ecophysiological model analysis (Yin et al., 2005) showed that all four traits were important, albeit to different extents, for predicting differences of flowering time among the individual genotypes of the population.

**QTL for the model-input traits**

The genetic analysis led to identified QTL for each of the four traits (Table 2). The two-stage MQM analysis successfully detected 4–7 QTL for a trait. Some genome positions were common to all four traits, whereas others were only specific to one or two traits (Fig. 1). Using the same ‘Apex×Prisma’ RIL population, Yin et al. (1999, 2003) have shown that a major gene, the *denso* (or designated as *sdw1*) locus (with a dwarfing allele from ‘Prisma’) located at 126.4 cM on chromosome 3(3H) (Fig. 1) pleiotropically affects a number of traits including the preflowering duration. This gene is shown here to be important for, or closely linked to, all four traits \((f_0, \theta_1, \theta_2, \text{ and } \delta)\), the modelling components of the preflowering duration (Fig. 1). The dwarfing allele had an increasing effect on \(f_0\), \(\theta_1\), and \(\delta\), and a decreasing effect on \(\theta_2\) (Table 2). A second genome position of interest that is associated with all traits was mapped at 94.4–99.2 cM on chromosome 5(1H) (Fig. 1). The direction of this locus on the four traits was just opposite to that of the *denso* locus. Both loci were detected in a previous QTL study for preflowering duration measured over two years (Yin et al., 1999). The two loci common to all traits provide the genetic basis for the significant phenotypic correlations among the traits (Table 1).

The locus close to the upper terminal marker of chromosome 3(3H) and the locus at 33.6–36.7 cM of chromosome 4(4H) are both shared by traits \(\theta_1\) and \(\delta\) (Fig. 1). Both loci had an increasing allele from ‘Prisma’ (Table 2). Another region, which is at 66.9–67.6 cM on chromosome 6(6H), is shared by \(\theta_2\) and \(\delta\) (Fig. 1), with an increasing effect on \(\theta_2\) and a decreasing effect on \(\delta\) from ‘Prisma’ (Table 2). These loci gave an additional reason for the high phenotypic correlations between \(\theta_2\) and \(\delta\), and between \(\theta_2\) and \(\delta\) (Table 1). None of these three loci were previously detected for preflowering duration (Yin et al., 1999).

Other mapped QTL are specific to only one trait (Fig. 1). Three specific loci (126.0 cM on chromosome 2(2H), 71.5 cM on chromosome 3(3H) and 144.1 cM on chromosome 5(1H)) were mapped for \(f_0\). Two specific loci (160.8 cM on chromosome 4(4H) and some 4(4H) are both shared by traits \(\theta_2\) and \(\delta\) (Fig. 1). Both loci had an increasing allele from ‘Prisma’ (Table 2).

---

**Table 1. Trait means and ranges (minimum-maximum) of RILs, and phenotypic correlation coefficients between the four traits**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Unit</th>
<th>Mean</th>
<th>Range</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_0)</td>
<td>Thermal day</td>
<td>66.100</td>
<td>66.545</td>
<td>57.148–77.111</td>
</tr>
<tr>
<td>(\theta_1)</td>
<td>–</td>
<td>0.4169</td>
<td>0.4216</td>
<td>0.1590–0.5720</td>
</tr>
<tr>
<td>(\theta_2)</td>
<td>–</td>
<td>0.7513</td>
<td>0.7527</td>
<td>0.6132–0.9880</td>
</tr>
<tr>
<td>(\delta)</td>
<td>h(^{-1})</td>
<td>0.0724</td>
<td>0.0741</td>
<td>0.0499–0.1052</td>
</tr>
</tbody>
</table>

\(a\) Estimated intercept in equation (4), based on multiple linear regression on QTL genotypes (see text).

---

**Fig. 1.** The AFLP marker linkage map as established by Yin et al. (1999) for the ‘Apex×Prisma’ population (marker positions are given in cM, using the Kosambi function), and QTL position for four input traits of an ecophysiological phenology model (black bar for trait \(f_0\), open bar for \(\theta_1\), line-hatched bar for \(\theta_2\), and cross-hatched bar for \(\delta\)). The figure was drawn using MapChart software (Voorrips, 2002). The QTL-position bars, placed on the right side of the relevant marker linkage group, are shown in length as ±4 cM of the highest LOD score position (cf. Table 2). The major dwarfing gene, *denso* (also designated as *sdw1*), is mapped on chromosome 3(3H) as the only morphological marker. Barley chromosomes 1–7 are homeologous to wheat chromosomes and are sometimes designated 7H, 2H, 3H, 4H, 1H, 6H, and 5H, respectively (Kleinhofs and Kilian, 1994). Both chromosome designations are used here.

**Fig. 2.** Frequency distribution of four model-input traits in the ‘Apex×Prisma’ population consisting of 94 RILs. Arrows show values for the two parents (full arrow for ‘Apex’ and dashed arrow for ‘Prisma’).
were mapped for \( f_0 \), both with the increasing allele from ‘Prisma’. The locus at 54.7 cM on chromosome 5(1H) and the locus at 25.8 cM on chromosome 7(5H) were mapped specifically for \( \theta_1 \) and \( \theta_2 \), respectively. Among these specific QTL, only \( f_0 \) loci on chromosomes 2(2H) and 5(1H) were in close proximity to minor QTL detected previously for preflowering duration (Yin et al., 1999).

This analysis gives information about the magnitude of the combined effect of identified multiple QTL for the trait, based on the additive multiple-QTL model, equation (4). The model accounted for 71.0, 37.5, 33.6, and 41.2% of the phenotypic variation in \( f_0 \), \( \theta_1 \), \( \theta_2 \), and \( \delta \), respectively (Table 2).

**Table 2. List of QTL identified for the four model-input traits**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Output from MapQTL(^a)/G 4.0</th>
<th>Calculated QTL effects using equation (4)</th>
<th>( R^2(%)^d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr(^a)</td>
<td>cM(^b)</td>
<td>LOD</td>
<td>( a_i^c )</td>
</tr>
<tr>
<td>( f_0 )</td>
<td>2(2H)</td>
<td>126.0</td>
<td>12.98</td>
</tr>
<tr>
<td></td>
<td>3(3H)</td>
<td>71.5</td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td>125.2</td>
<td>18.22</td>
<td>2.435335</td>
</tr>
<tr>
<td></td>
<td>5(1H)</td>
<td>94.4</td>
<td>6.91</td>
</tr>
<tr>
<td></td>
<td>144.1</td>
<td>5.36</td>
<td>0.962237</td>
</tr>
<tr>
<td>( \theta_1 )</td>
<td>3(3H)</td>
<td>0.7</td>
<td>6.45</td>
</tr>
<tr>
<td></td>
<td>123.2</td>
<td>4.88</td>
<td>0.0293618</td>
</tr>
<tr>
<td></td>
<td>4(4H)</td>
<td>36.7</td>
<td>3.51</td>
</tr>
<tr>
<td></td>
<td>5(1H)</td>
<td>54.7</td>
<td>3.24</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>3(3H)</td>
<td>123.2</td>
<td>6.93</td>
</tr>
<tr>
<td></td>
<td>4(4H)</td>
<td>67.6</td>
<td>3.81</td>
</tr>
<tr>
<td></td>
<td>5(1H)</td>
<td>99.2</td>
<td>6.38</td>
</tr>
<tr>
<td></td>
<td>7(5H)</td>
<td>25.8</td>
<td>4.37</td>
</tr>
<tr>
<td>( \delta )</td>
<td>3(3H)</td>
<td>2.7</td>
<td>4.98</td>
</tr>
<tr>
<td></td>
<td>123.2</td>
<td>3.00</td>
<td>0.00303462</td>
</tr>
<tr>
<td></td>
<td>160.8</td>
<td>4.23</td>
<td>0.00264912</td>
</tr>
<tr>
<td></td>
<td>4(4H)</td>
<td>33.6</td>
<td>5.58</td>
</tr>
<tr>
<td></td>
<td>66.9</td>
<td>5.10</td>
<td>-0.00367094</td>
</tr>
<tr>
<td></td>
<td>5(1H)</td>
<td>94.4</td>
<td>5.92</td>
</tr>
<tr>
<td></td>
<td>6(6H)</td>
<td>40.6</td>
<td>3.26</td>
</tr>
</tbody>
</table>

\( a \) | Chromosome number.  
\( b \) | Counted from the upper terminal marker of the chromosome.  
\( c \) | Additive effect of a QTL on the trait concerned, as defined by: (mean of the ‘Prisma’ allele genotypes − mean of the ‘Apex’ allele genotypes)/2.  
\( d \) | The percentage of phenotypic variation accounted for by all the QTL for the trait concerned, estimated using multiple regression based on equation (4).

**Table 3. The percentage (%) of variation in field observed days to flowering accounted for by the ecophysiological phenology model with either original phenotypic or QTL-based model parameters**

<table>
<thead>
<tr>
<th>Environment</th>
<th>Data points</th>
<th>Original parameters</th>
<th>QTL-based parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 October 2001</td>
<td>94</td>
<td>57.58</td>
<td>40.39</td>
</tr>
<tr>
<td>5 November 2001</td>
<td>94</td>
<td>66.06</td>
<td>43.08</td>
</tr>
<tr>
<td>16 November 2001</td>
<td>94</td>
<td>59.65</td>
<td>29.22</td>
</tr>
<tr>
<td>25 November 2001</td>
<td>94</td>
<td>66.06</td>
<td>43.08</td>
</tr>
<tr>
<td>5 November 2001</td>
<td>94</td>
<td>66.06</td>
<td>43.08</td>
</tr>
<tr>
<td>16 November 2001</td>
<td>94</td>
<td>59.65</td>
<td>29.22</td>
</tr>
<tr>
<td>25 November 2001</td>
<td>94</td>
<td>66.06</td>
<td>43.08</td>
</tr>
<tr>
<td>Across-RIL mean</td>
<td>94</td>
<td>76.18</td>
<td>45.55</td>
</tr>
</tbody>
</table>

\( ^a \) Within an environment.  
\( ^b \) Each environment is indicated by the sowing date of the field experiment.
independent test of their model for predicting leaf elongation rates in maize, Reymond et al. (2003) used not only two parents but also some RILs that had not been included in the QTL analysis. Relative to theirs, this RIL population is smaller. All RILs of the population were included in the mapping analysis for a maximum power of QTL detection.

Predictions for days to flowering using a QTL-based ecophysiological model were directly compared with predictions coming from using the original, phenotypic parameters (Fig. 5). The QTL-based predictions correlated well with the original predictions \((r=0.92)\). Similar good correlation was also obtained from earlier QTL-based predictions for barley grain yields (Yin et al., 2000a).

**Comparison with earlier QTL-based modelling**

The better performance of the model using the original input trait values than the QTL-based model (Fig. 3; Table 3) is not surprising, because variation accounted for by identified QTL was only about 35–40% for most of model-input traits (Table 2). The result reported here for days to flowering differed from the previous report (Yin et al., 2000a) in predicting grain yields in the same ‘Apex×Prisma’ population. In that report, the QTL-based crop growth model performed somewhat better \((r^2=0.524)\) than the model using originally measured traits did \((r^2=0.376)\). The better performance of the QTL-based model was simply attributed to the fact that random errors in the original traits were properly removed by the QTL statistics (Yin et al., 2000a). The result of the current analysis for days to flowering indicates that the answer may not have been that simple.

In the phenology model for predicting flowering time, equations (1) and (2), four input parameters \((f_0, \theta_1, \theta_2, \text{ and } \delta)\) were used and they were all important, although to different extents, to explain differences in flowering time among the RILs (Yin et al., 2005). In the crop growth model used in the earlier study (Yin et al., 2000a, b) for
predicting grain yield, six model-input traits were examined. Among them, only two were found to be important for predicting grain yield differences among the RILs, because use of the across-RIL mean of the other four input parameters resulted in better model predictions than use of measured RIL-specific parameter values (Yin et al., 2000b). The better performance of the QTL-based model in predicting grain yields could simply be due to the fact that relative to the original, phenotypic values, the QTL-based input traits had a narrower range of values for those four unimportant parameters, as they were being shrunk to their across-RIL means. This reasoning is supported by the result of a new simulation, in which a two-parameter crop growth model was used (fixing the other four, unimportant, parameters at their across-RIL means). When the measured phenotypic inputs for the two parameters were used, 64.8% of the variation in grain yield was accounted for. By contrast, for the crop growth model using the QTL-based inputs of the two parameters, this percentage was slightly lower, 62.1%. Reymond et al. (2003) also showed that the QTL-based predictions of maize leaf elongation rates were somewhat more dispersed than the original predictions. It could be a general phenomenon that the gain from the QTL statistics that removes, at least part of, random noise in the original model-input traits, may not be sufficient to compensate for the loss due to the residual genetic variance that is not captured by the identified QTL.

Comparison of QTL-based model predictions with those using QTL for days to flowering per se

To what extent QTL-based model predictions are close to the predictions using QTL for field flowering dates per se was examined next. To this end, the observed days to flowering per environment were subjected to QTL analysis using the method described earlier. One to four QTL were detected for the flowering time in each environment (Table 4). The genome positions around 120 cM on chromosome 2(2H) and at 126 cM on chromosome 3(3H), that turned out to be important in most environments, were in close proximity to the two major QTL found for the model-input trait $f_o$ (Table 2). Some QTL were detected in one or several, but not in other environments, indicating a ‘QTL×environment interaction’, a common finding when data for a quantitative trait observed in multiple environments are subject to QTL analysis (Jansen, 1995; Jiang and Zeng, 1995). The identified QTL accounted for $79.4, 77.1, 65.7, 69.2, 45.2, 33.8, 15.4, and 53.4\%$ of the phenotypic variation of days to flowering among RILs in the eight field environments, respectively (Table 4).

When all data points for both RILs and parents from the eight environments were pooled, QTL for field observed

![Fig. 5. Correlation between predicted days from sowing to flowering from QTL-based model-input traits and those from the original model-input values for 94 RILs and their parents. Data from eight field environments are pooled.](image)

Table 4. Position in cM and LOD value (the first and second figure of each data pair, respectively) of QTL identified for days to flowering when data from each of eight field environments were subjected to QTL analysis

<table>
<thead>
<tr>
<th>Environment $^a$</th>
<th>Chromosome $^b$</th>
<th>1(7H)</th>
<th>2(2H)</th>
<th>3(3H)</th>
<th>4(4H)</th>
<th>5(1H)</th>
<th>7(5H)</th>
<th>$R^2(%)$ $^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 October 2001</td>
<td>125.0; 6.1</td>
<td>126.4</td>
<td>29.6</td>
<td></td>
<td></td>
<td>89.4</td>
<td>5.2</td>
<td>80.8; 3.3</td>
</tr>
<tr>
<td>5 November 2001</td>
<td>124.0; 6.4</td>
<td>126.4</td>
<td>25.4</td>
<td>64.3</td>
<td>3.7</td>
<td>96.3</td>
<td>3.5</td>
<td>77.1</td>
</tr>
<tr>
<td>16 November 2001</td>
<td>120.0; 3.5</td>
<td>126.4</td>
<td>17.6</td>
<td>41.7</td>
<td>3.1</td>
<td>80.4</td>
<td>4.2</td>
<td>65.7</td>
</tr>
<tr>
<td>25 November 2001</td>
<td>125.0; 5.8</td>
<td>126.4</td>
<td>21.7</td>
<td>89.4</td>
<td>3.9</td>
<td>89.4</td>
<td>69.2</td>
<td></td>
</tr>
<tr>
<td>4 December 2001</td>
<td>23.0; 3.1</td>
<td></td>
<td>126.4</td>
<td>18.7</td>
<td></td>
<td>45.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 October 2002</td>
<td>125.0; 4.3</td>
<td>118.8</td>
<td>6.3</td>
<td>126.4</td>
<td>5.0</td>
<td></td>
<td>33.8</td>
<td></td>
</tr>
<tr>
<td>10 November 2002</td>
<td>125.0; 4.3</td>
<td></td>
<td>58.5</td>
<td>3.8</td>
<td>126.4</td>
<td>12.5</td>
<td>15.4</td>
<td></td>
</tr>
<tr>
<td>25 November 2002</td>
<td>30.5; 4.5</td>
<td>118.8</td>
<td>5.4</td>
<td></td>
<td></td>
<td></td>
<td>53.4</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Each environment is indicated by the sowing date of the field experiment.

$^b$ No QTL was found on chromosome 6(6H).

$^c$ The percentage of phenotypic variation accounted for by all identified QTL for days to flowering in the environment concerned, estimated using multiple regression based on equation (4).
days to flowering \textit{per se} accounted for 85.5\% of overall variation (Fig. 6A), somewhat higher than 72.6\%, the percentage of overall variation accounted for by the QTL-based phenology model (Fig. 6B). The slightly better predictions using QTL for field observed days to flowering itself per environment is not surprising because ecophysiological QTL-based model predictions under field environments were completely independent of the greenhouse experiment from which the phenotypic model-input traits had been derived. Another reason is that random noise may have influenced the curve-fitting of the original phenotypic values of ecophysiological model-input traits. Nevertheless, the correlation between two sets of QTL-based predictions was high ($r=0.89$, Fig. 7), indicating that the simple ecophysiological phenology model captures a large part of the photothermal responses of genotypes in the population under the environmental conditions studied. Therefore, a robust ecophysiological model is capable of extrapolating (QTL) information from one environment to another.

**Concluding remarks**

Following our earlier research (Yin \textit{et al.}, 1999, 2000\textit{a, b}), the present study continued to examine the feasibility of combining ecophysiological modelling and genetic mapping to predict performance of individuals in an RIL population under various environmental conditions. The interest in conducting this across-disciplinary research was to explore the utilization of complementary aspects of these two research areas (Yin \textit{et al.}, 2003). QTL mapping allows the dissection of a phenotype into individual genetic factors (Paterson \textit{et al.}, 1988), but its ability to extrapolate QTL information from one set of environment/management conditions to independent new conditions is limited (Stratton, 1998). Ecophysiological modelling can potentially give an insight into how G×E comes about (Tardieu, 2003), but it does not account for the genetic basis of differences in response to environmental changes. Combining ecophysiological modelling and genetic mapping into a QTL-based crop model could be powerful for resolving complex environment-dependent yield traits on a genetic basis. In view of the weakness of current crop-growth models in predicting differences in grain or seed yields among similar lines of a segregating population (Yin \textit{et al.}, 2000\textit{a, b}), the focus in the current analysis was on a relatively simple trait, days to flowering. This analysis, together with the work of Reymond \textit{et al.} (2003) on another simple trait (leaf elongation rate), highlighted a great potential of this

![Fig. 6.](image1.png)

**Fig. 6.** Comparison between observed days from sowing to flowering and those estimated from QTL for days to flowering \textit{per se} (A), and between observed days to flowering and those predicted from using QTL-based phenology model (B), for 94 RILs and their parents. Data from eight field environments are pooled.

![Fig. 7.](image2.png)

**Fig. 7.** Correlation between days to flowering predicted from QTL-based model and those estimated using QTL for days to flowering \textit{per se} in eight environments.
combined approach in predicting the performance of plants in a mapping population carrying any combination of alleles under a wide range of climatic scenarios. While ecophysiological modelling and genetic mapping have evolved independently so far, modellers, physiologists, geneticists, and molecular biologists may work to reach a synergy, a blending of scientific disciplines, to face the challenges posed by the increasing availability of genomic information to predict plant or crop phenotypes for complex traits.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (contract No. 30060036). We thank Hui Li and Peilian Zhang for their assistance in managing the experiments, and Dr MP Boer for providing his program for the estimation of allelic QTL genotypes.

References


