Field Crops Research xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Field Crops Research

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

Combining genome-wide prediction and a phenology model to simulate heading date in spring barley

Ralf Uptmoor[a,*], Klaus Pillen[b], Claudia Matschegewski[a]

[a] Department of Agronomy, University of Rostock, Justus-von-Liebig-Weg 6, 18055 Rostock, Germany
[b] Chair of Plant Breeding, Martin-Luther-University Halle-Wittenberg, Betty-Heimann-Str. 3, 06120 Halle/Saale, Germany

ARTICLE INFO

Article history:
Received 17 September 2015
Received in revised form 6 July 2016
Accepted 3 August 2016
Available online xxx

Keywords:
Photoperiod
Temperature
Heading date
Genomic prediction
Phenology model

ABSTRACT

Phenotype by genotype prediction based on ecophysiological models, which account for allelic gene, QTL, or marker effects, have many possible applications in plant breeding programs. The goal of the present study was to predict heading date of individual lines of a *Hordeum vulgare* x *H. vulgare ssp. spontaneum* BC_2DH-population using a phenology model parameterized with marker effects derived from ridge regression best linear unbiased prediction. The genetic linkage map included SSR markers and flowering-time genes. Effects of photoperiod and temperature on heading date were measured under controlled conditions on a subset of the population comprising the recurrent parent and 36 BC_2DH candidate introgression lines covering the *H. spontaneum* genome. Marker effects, which were subsequently used for model parameterization, were estimated. Model evaluation was carried out on already published field trial data comprising the 36 BC_2DH lines and 266 independent BC_2DH lines from the same cross. Applying the model on the lines used for model parameterization explained 33–51% of heading-date variation in three of the four evaluation environments but only 20% of the variation in the fourth environment. Heading dates of the 266 independent lines were predicted with less accuracy. Between 20 and 25% of phenotypic variation was explained by the model in three environments and only 8% of heading date variation in the fourth environment. The root mean squared error (*RMSE*) was slightly higher for independent lines than for the lines used for model parameterization. Dissecting *RMSE* into its components revealed that *RMSE* was largely influenced by a systematic bias in most environments and by the missing ability of the model to describe the observed variation within the set of genotypes in all environments. Comparing the combined genome-wide prediction (GWP) and phenology model with a conventional GWP model gave similar prediction accuracies if the training set had the same size.

1. Introduction

As a possibility to accelerate breeding efforts, the combination of ecophysiological or phenological modeling and QTL analysis has been suggested (Yin et al., 2014; Chapman, 2008; Hammer et al., 2010). For the prediction of environmental effects on plants with different allele combinations of relevant genes and for a better identification of genetic factors, which underlie complex traits and are influenced by environmental factors, modeling helps to dissect these traits into underlying physiological processes (Tardieu, 2003). Thus, combining QTL and crop modeling could be used for defining ideotypes, *in-silico* testing of possible allelic QTL or gene combinations (Cooper et al., 2002) and for an efficient selection of candidate genes (Boote et al., 2001; Slafer, 2003). Ma et al. (2002) found that QTL mapping approaches using repeated measurements on growth curves provide maximum information about QTL effects and positions and are advantageous if small populations are sampled or medium dense genetic maps are used. Crop models and QTL analysis were combined, e.g., for describing leaf elongation rate in maize (Reymond et al., 2003) and fruit quality in peach (Quilot et al., 2005). Phenological models were used to model temperature effects in *Brassica oleracea* (Uptmoor et al., 2008; Uptmoor et al., 2012), to identify QTL for flowering time in response to photoperiod and temperature in rice (Nakagawa et al., 2005), and to detect flowering time QTL in response to photoperiod in barley, which were subsequently used for model parameterization (Yin et al., 2005a,b). The impact of QTL controlling leaf and silk elongation on crop growth, water use and grain yield was simulated by Chenu et al. (2009). A neural network model with underlying genetic information was developed for simulating flowering time

* Corresponding author.
E-mail address: ralf.uptmoor@uni-rostock.de (R. Uptmoor).

http://dx.doi.org/10.1016/j.fcr.2016.08.006
0378-4290/

in *Arabidopsis thaliana* (Welch et al., 2003), and a genetically parameterized photo-thermal model was used to predict bolting of *A. thaliana* genotypes in natural environments (Wilczek et al., 2009, 2010).

A major problem of classical QTL studies carried out on the progenies of bi-parental crosses is that they ignore small-effect loci, which limits QTL based marker-assisted selection (MAS) strategies. The importance of small effect QTL for accuracies in predicting phenotypes by genotypes is due to the polygenic nature of quantitative traits, in which many genes contribute to the trait but the contributive effect of each single gene is small. GWP models estimate marker effects of the whole genome and, thus, include also small effect QTL (Desta and Ortiz, 2014). GWP was introduced by Meuwissen et al. (2001); first used in animal breeding (Hayes et al., 2009) and later adopted in plant breeding (Lorenz et al., 2011). The integration of crop modeling and GWP has already been discussed (Heslot et al., 2014; Technow et al., 2015).

The present study aims at combining a GWP and a phenology model in order to simulate heading date of any possible progeny of a barley BC_2DH population. Specific objectives were (1) to estimate the effects of daylength and temperature on heading date of the recurrent parent and 36 BC_2DH lines of the *H. vulgare* 'Scarlett' x *H. vulgare* spp. *spontaneum* 'ISR42-8' cross (von Korff et al., 2004), (2) to parameterize a heading date model with genome-wide marker effects, and (3) to compare the combined GWP and phenology model with results of a conventional field-data based GWP model. We hypothesized that heading date can be predicted by the genotype in any environment if estimated marker effects linked to genetic loci, which control development towards heading in response to photoperiod and temperature, are known. GWP models are useful to analyze introgression lines (ILs) or backcross lines carrying more than one donor segment since these statistical models do not suffer from overparameterization if the number of markers is larger than the number of individuals, and effects can be assigned to individual donor segments even if ILs carry multiple donor alleles (Hofheinz and Frisch, 2014; Falke et al., 2014). The number of estimates of allelic marker effects may exceed the number of genotypes since all markers have an effect sampled from the same normal distribution, variances are the same for all effects, and, thus, only one variance is estimated (Meuwissen et al., 2001; Hayes et al., 2009). Ridge regression best linear unbiased prediction (rrBLUP) is the most widely used GWP model, if analysis is restricted to a biparental population (Jannink et al., 2010).

2. Materials and methods

2.1. Plant material and genotype information

We used the malting-barley cultivar 'Scarlett' and 36 *H. vulgare* cv. 'Scarlett' x *H. vulgare* ssp. *spontaneum* 'ISR42-8′ BC_2DH lines with 'ISR42-8′ as donor for model parameterization and field heading-date data of 302 BC_2DH lines of the cross comprising the 36 lines for model parameterization (training set) and 266 independent lines for model evaluation (validation set). The 36 lines were selected from the BC_2DH population according to genome coverage. The whole population was mapped with 97 SSR markers (von Korff et al., 2004). The main flowering time regulators *PpdH1, VrnH1, VrnH2, VrnH3, HvCO1, HvCO2, HvGI, HvFT2, HvFT3*, and *HvFT4* were integrated into the SSR map (Wang et al., 2010).

2.2. Experiments

Three experiments were conducted in order to define model parameters for photoperiod and temperature responses. For the characterization of temperature effects, 36 BC_2DH lines and the recurrent parent were sown in rows in polyvinyl boxes with four temperature treatments, two replications, and seven plants per replication. Boxes were kept in climate chambers at 15 °C for the first ten days after sowing. Temperatures of one treatment were then changed to 10 °C and those of two other treatments to 20 °C. Plants were then grown at constant day/night temperatures of 10, 15, 20, and 25 °C. Temperature of the latter treatment was increased from 20 to 25 °C 13 days after sowing. Plants were cultivated under long day (LD) conditions (16 h) with 14 h full light plus 1 h twilight in the morning and evening. Since barley genotypes may flower before heading, BBCH stage 51 (Lancashire et al., 1991) was recorded daily and defined as time to heading. The tip of the ear becomes visible at the top of the shoot or from the side of the leaf sheath at BBCH 51.

Genotypic variation in response to photoperiod was evaluated in two reciprocal transfer experiments. The LD variable was realized by a daylength of 15 h including 1 h twilight in the morning and evening, the short day (SD) variable by a daylength of 10 h full light. The first reciprocal transfer experiment included all 36 BC_2DH lines and the recurrent parent. Plants were grown under LD and SD conditions in two replications and with two plants per replication. Transfers from LD to SD and *vice versa* were carried out in 14 d intervals. The second reciprocal transfer experiment comprised only the recurrent parent and BC_2DH lines carrying introgressions on the flowering time genes *PpdH1* and *VrnH1*. All other BC_2DH lines of the training set showed no significant difference from 'Scarlett' in the first reciprocal transfer experiment (data not shown). Plants were cultivated in growth chambers with two replications and seven plants per replication at a constant temperature of 22 °C. Seeds were sown in polyvinyl boxes. Initially, half of the plants of each DH-line were placed in climate chambers with LD or SD conditions, respectively. Starting nine days after sowing, one set of ILs was transferred from SD to LD and from LD to SD in seven-day intervals until 80 days after emergence. Once a plant was transferred, it was grown in the new environment until heading. One treatment grew from sowing to heading continuously in LD or SD, respectively.

2.3. Model framework

A multiplicative model was used for simulating time to heading. Daily development rates (ω_i) were calculated as follows:

$$\begin{aligned} \omega_i &= f(T)/f_o && \text{if } \theta \le \theta_1 \text{ or } \theta \ge \theta_2 \\ \omega_i &= f(T) \times f(P)/f_o && \text{if } \theta_1 < \theta < \theta_2 \end{aligned} \quad (1)$$

where $f(T)$ is the temperature response function, $f(P)$ is the photoperiod response function, f_o is the minimum number of days from sowing to heading, θ is the actual photoperiod, θ_1 is the start of the photoperiod sensitive phase and θ_2 is the end of the photoperiod sensitive phase (Yin et al., 2005a).

The temperature response function was computed by using cardinal temperatures (Wang and Engel, 1998):

$$\begin{aligned} f(T) &= (2(T - T_{min})^a (T_{opt} - T_{min})^a - (T - T_{min})^{2a}) / (T_{opt} - T_{min})^{2a} \\ & \qquad \text{if } T_{min} \le T \ge T_{max} \\ f(T) &= 0 \\ & \qquad \text{if } T < T_{min} \text{ or } T > T_{max}, \end{aligned} \quad (2)$$

where T is the actual daily mean temperature, T_{opt} is the temperature optimum for maximum development rates under optimum

Please cite this article in press as: Uptmoor, R., et al., Combining genome-wide prediction and a phenology model to simulate heading date in spring barley. Field Crops Res. (2016), http://dx.doi.org/10.1016/j.fcr.2016.08.006