



Genome-Based Prediction of Time to Curd Induction in Cauliflower

Arne Rosen^{1*}, Yaser Hasan², William Briggs³ and Ralf Uptmoor¹

¹ Faculty of Agriculture and Environmental Science, University of Rostock, Rostock, Germany, ² Institute of Horticultural Production Systems, Leibniz Universität Hannover, Hannover, Germany, ³ Syngenta Seeds B.V., Enkhuizen, Netherlands

The development of cauliflower (*Brassica oleracea* var. *botrytis*) is highly dependent on temperature due to vernalization requirements, which often causes delay and unevenness in maturity during months with warm temperatures. Integrating quantitative genetic analyses with phenology modeling was suggested to accelerate breeding strategies toward wide-adaptation cauliflower. The present study aims at establishing a genome-based model simulating the development of doubled haploid (DH) cauliflower lines to predict time to curd induction of DH lines not used for model parameterization and test hybrids derived from the bi-parental cross. Leaf appearance rate and the relation between temperature and thermal time to curd induction were examined in greenhouse trials on 180 DH lines at seven temperatures. Quantitative trait loci (QTL) analyses carried out on model parameters revealed ten QTL for leaf appearance rate (LAR), five for the slope and two for the intercept of linear temperature-response functions. Results of the QTL-based phenology model were compared to a genomic selection (GS) model. Model validation was carried out on data comprising four field trials with 72 independent DH lines, 160 hybrids derived from the parameterization set, and 34 hybrids derived from independent lines of the population. The QTL model resulted in a moderately accurate prediction of time to curd induction ($R^2 = 0.42\text{--}0.51$) while the GS model generated slightly better results ($R^2 = 0.52\text{--}0.61$). Predictions of time to curd induction of test hybrids from independent DH lines were less precise with $R^2 = 0.40$ for the QTL and $R^2 = 0.48$ for the GS model. Implementation of juvenile-to-adult phase transition is proposed for model improvement.

Keywords: quantitative trait loci, genomic selection, phenology model, cauliflower, vernalization, curd induction

OPEN ACCESS

Edited by:

Daniel Pinero,
Universidad Nacional Autónoma
de México, Mexico

Reviewed by:

Humira Sonah,
Laval University, Canada
Robert VanBuren,
Michigan State University,
United States

*Correspondence:

Arne Rosen
arne.rosen@uni-rostock.de

Specialty section:

This article was submitted to
Plant Genetics and Genomics,
a section of the journal
Frontiers in Plant Science

Received: 24 January 2017

Accepted: 15 January 2018

Published: 05 February 2018

Citation:

Rosen A, Hasan Y, Briggs W and
Uptmoor R (2018) Genome-Based
Prediction of Time to Curd Induction
in Cauliflower. *Front. Plant Sci.* 9:78.
doi: 10.3389/fpls.2018.00078

INTRODUCTION

Floral or curd induction is a temperature-mediated process in cauliflower. While temperature optima for curd induction are around 15°C for many summer cultivars, temperatures above 20°C delay or inhibit curd induction. Wide harvest windows resulting from high ambient temperatures lead to several selective harvests on the same field causing high workloads for growers (Wiebe, 1980). Phenology models are useful tools to cope with uncertainties in time to curd induction and harvest (Jensen and Grevsen, 2005). Several modeling approaches have been made to predict harvest time in cauliflower: The earliest models used linear relations between curd diameter and temperature sum to predict time to maturity, which required curd diameter measurements in the field (Salter, 1969; Hand, 1988; Wurr, 1989; Wurr et al., 1990a,b). Later models integrated

earlier development stages beginning at transplanting (Pearson et al., 1994; Grevsen and Olesen, 1994). Several models divided development before curd induction into the juvenile phase, ending after a certain number of leaf primordia is initiated and during which plants are insensitive to vernalization, and the adult vegetative or vernalization phase, characterized by temperature sensitivity (Wiebe, 1972a,b,c; Wurr, 1989; Kage and Stützel, 1999). Olesen and Grevsen (2000) established a model simulating time to maturity and enabling curd quality prognosis based on weather data. Jensen and Grevsen (2005) presented a web-based management tool for growers and wholesale traders predicting harvest time by applying the temperature-sum rule. In addition to crop phenology, Kage et al. (2001a,b) modeled the dry matter production and partitioning by means of absorbed photoactive radiation and light use efficiency.

The mentioned models are valid for single or few cultivars. Using crop growth or phenology models in plant breeding was first proposed by Yin et al. (2003) via integrating QTL and ecophysiological models. QTL-based phenology models have already been established for crops like wheat (Bogard et al., 2014), barley (Yin et al., 2005b), soy (Messina et al., 2006), and *B. oleracea* (Uptmoor et al., 2012). The studies revealed that QTL can be detected for model parameters describing the response curve of time to flowering, heading or floral induction in relation to the environmental factors photoperiod and/or temperature and that flowering time of progenies within bi-parental crosses can be predicted based on QTL effects. However, most of the studies used no or only few independent lines from the respective populations for model validation, which is apart from QTL detection a prerequisite for possible applications in plant breeding. GS was first introduced in animal breeding (Meuwissen et al., 2001) and has emerged as a new tool in plant breeding by using genome-wide markers and their corresponding effects to predict phenotypes of untested genotypes (Desta and Ortiz, 2014). While most GS studies were conducted in populations with a broad genetic basis, also bi-parental crosses were used (Lorenza and Bernardo, 2009). Combined GS and phenology models have already been tested and are likely to replace QTL-based modeling approaches since predictions can be made across bi-parental populations (Heslot et al., 2014; Technow et al., 2015; Onogi et al., 2016; Uptmoor et al., 2016). Both, QTL-based modeling approaches and combined GS and crop models may dissect complex traits into underlying physiological factors and have in common that QTL or marker effects are estimated for parameters of the response curves of the trait or underlying physiological factor to environmental influences like temperature. While QTL-based modeling approaches consider only main effect loci, GS models estimate effects for all markers and, thus, capture small effect loci as well.

Since *B. oleracea* is closely related to *Arabidopsis thaliana* (Lagercrantz et al., 1996), for which the main regulators of flowering time are known (Jack, 2004; Blázquez, 2005), QTL and flowering-time regulator co-localizations gave hint that these genes may account for flowering-time variation in *B. oleracea*. One important flowering time integrator is *FLOWERING LOCUS C (FLC)*, an inhibitor of flowering, which is downregulated

by vernalization and has a significant impact on flowering-time variation in *A. thaliana* (Koornneef et al., 1994). The *B. oleracea* genome carries several *FLC* homologs (Lagercrantz et al., 1996; Okazaki et al., 2007; Axelsson et al., 2001) and co-localizations with flowering-time QTL indicated that allelic variation of *FLC* might play an important role in flowering-time regulation in *B. oleracea* (Schranz et al., 2002; Brown et al., 2007). However, Okazaki et al. (2007) and Razi et al. (2008) suggested that several tested *FLC* homologs unlikely induce flowering-time variation in *B. oleracea*. While Ridge et al. (2015) found evidence for functional relevance of *BoFLC2* in floral induction, Matschegewski et al. (2015) found no consistent *BoFLC2* transcription patterns in cauliflower breeding lines for different climates and seasons. QTL for curd induction were not only identified in regions harboring *FLC* and *FRI* homologs but also on O6, where *BoAPI-a* and *BoAPI-c* are located (Hasan et al., 2016).

Despite the efforts spent in elucidating genetic mechanisms underlying floral induction in cauliflower, it is still widely unknown, how temperature-dependent development toward curd induction is regulated, which, in turn, hampers the development of new cultivars with stable time to maturity in a broad range of different environments. Identification of candidate genes underlying flowering-time QTL may benefit from integrated modeling approaches since model parameters allow to distinguish vernalization from flowering-time *per se* QTL. To date, most studies integrating genetic and ecophysiological models used QTL rather than genome-wide marker effects, even though state of the art marker-assisted plant breeding strategies rely on GS. Therefore, the present study aims at (1) the identification of QTL for parameters of a phenology model, assuming a linear relation between temperature and thermal-time to curd induction, (2) model parameterization based on both QTL and genome-wide marker effects, (3) model evaluation on multi-environment field data from independent genotypes including DH-lines *per se* and lines crossed with a tester, and (4) comparison of both parameterization principles.

MATERIALS AND METHODS

Plant Material

The present study was carried out on a population consisting of 265 DH lines derived from the F1 of a cross between homozygous parental lines P1 with good harvest-time reliability at higher temperatures and P2 that is less resistant to elevated temperatures but produces high-quality curds. More details about reliability in time to curd induction and harvest time of the PLs are provided by Hasan et al. (2016). A subset of 180 lines was used for model parameterization, while the validation sets included three subsets with 72 additional DH lines *per se*, 160 test hybrids from the parameterization set, and 34 test hybrids obtained from the 72 additional DH lines. Plant material was obtained from SYNGENTA Seeds B.V., Enkhuizen, Netherlands. The PLs were genotyped with a 20k *B. oleracea* Illumina Infinium iSelect single-nucleotide polymorphism (SNP) array (Matschegewski et al., 2015). A subset of SNPs of the array was selected based on

polymorphisms between PLs and genome coverage to genotype the DH-population with 176 SNP markers. The genetic map was constructed using MapMaker 3 and the Haldane function. The total map length was 891.2 cM spanning over nine chromosomes. Average distance between SNP markers was 5.3 cM. Physical marker positions are available from Hasan et al. (2016).

Model Parameterization Trials

Model parameterization is based on controlled temperature treatments described in Hasan et al. (2016). Six trials were carried out in greenhouses with mean temperatures of 11.8, 15.5, 17.3, 19.0, 21.4, and 27.0°C and one in climate chambers with a mean air temperature of 26.5°C. Seeds of all DH lines of the parameterization set were sown in seedling trays and raised in the greenhouse at 22°C until all plants had two to three visible leaves. Afterward, seedlings were planted into 3 l pots, transferred to the temperature treatments in greenhouses or climate chambers, and arranged in randomized complete block designs. Plants were grown at a photoperiod of 16 h, providing 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ additional light by Phillips SON-T Agro lamps when day length was below 16 h or natural radiation during the day below 5 klux. Plants were fertigated daily with 0.5 g l^{-1} Scotts Universal solutions. All trials were terminated after 120 days.

Time to visible curd induction was recorded daily. A curd diameter of 1 cm was used as threshold for considering visible curd induction to have been occurred. Number of leaves larger than 1 cm was counted twice a week. Leaf appearance rate (LAR) was computed as the slope of the regression of leaf numbers on growing degree-days assuming a base temperature of zero.

Model Validation Trials

For model validation, a field trial was carried out with 72 lines in 2013 in Rostock, Germany. Twelve seeds per line were sown on May 6 into seedling trays and raised in the greenhouse. Eight plants of each line were planted into the field on June 5 when most plants had two to three leaves. Field trials were carried out in a randomized complete block design with two replications and four plants per replication. Seedlings were planted in rows with distances of 50 cm between and within rows. A border row was planted to provide similar conditions for all plants. Irrigation was applied as required and the plot was covered with a net to protect plants during the first weeks from cabbage fly (*Delia radicum*)

and other pests. The trial ended on August 16. Time to curd induction was scored twice a week. Beginning of curd formation was defined as curd diameter ≥ 1 cm.

Additional validation trials with similar growing conditions were carried out in 2013 in Ruthe, Germany and in 2011 and 2012 in Zeewolde, Netherlands. In Ruthe, seeds of the validation set were sown on May 14 and eight plants per line were transplanted to the field on June 13 when they had two to three visible leaves. Zeewolde field trials were carried out on the whole DH-population and transplanting took place on June 14, 2011 and May 24, 2012. Mean air temperatures during the experiments were 17.9°C in Rostock, 19.0°C in Ruthe, and 17.4 and 17.2°C in Zeewolde 2011 and 2012, respectively. Only 44 DH lines were used in Zeewolde 2011, whilst all other trials comprised the whole DH *per se* validation set.

Further trials were conducted on the 160 F1 test-hybrids derived from the parameterization set and the 34 hybrids in Ócsa, Hungary, where plants were transplanted on July 7, 2011 and July 6, 2012 and in Zeewolde, where transplanting took place on June 13, 2012. Mean air temperatures in Ócsa were 21.4 and 22°C, respectively; mean air temperature in Zeewolde was 17.4°C. Harvest time was recorded in all Ócsa and Zeewolde trials. Sowing was assumed to have taken place 30 days before transplanting and duration from visible curd induction to harvest was assumed to be 30 days (Table 3).

Data Analysis and Model Description

The cauliflower phenology model was divided into a juvenile and a vernalization phase. During the juvenile phase, plants were assumed to be insensitive to temperature, i.e., development followed the temperature-sum rule. Although it is known that genotypic variation in the number of initiated or visible leaves at juvenile-to-adult transition exists (Hand and Atherton, 1987), juvenility was assumed to end for all genotypes after 7 leaves were visible. In contrast, LAR and phyllochron (P , d), the reciprocal of LAR, was assumed to be genotype specific. The genotype specific temperature sum for juvenile-to-adult transition (TS_f) was calculated as follows:

$$TS_f = 7 \times P$$

The model assumed genetic variation in temperature response from end of juvenility to visible curd induction (curd

TABLE 1 | Population means and standard deviations (SD) of temperature sums (TS) required for visible curd induction at different temperatures.

Temperature [°C]	TS parameterization set [°Cd]		TS parental lines [°C]	
	Mean	SD	P1	P2
11.75	1533	63.6	1629.5	1556.7
15.51	1462	65.7	1640.8	1519.8
17.33	1557	85.6	1923.8	1516.5
19.03	1570	85.4	1868.5	1607.8
21.44	1930	242.3	–	1888.4
26.49	1862	192.5	–	1911.1
27.00	2370	384.2	–	2346.0

Shown are means across DH-lines used for model parametrization TS of parental lines.

diameter > 1 cm). Accelerating or retarding temperature effects on phenological development may end at floral transition, which, however, was not observed. Instead, the relation between thermal time to visible curd induction in seven different treatments and temperature was estimated by linear regression, where the intercept T_0 (°Cd) describes the extrapolated (theoretical) minimum thermal time to curd induction and the slope S (°Cd °C⁻¹) the temperature sensitivity, i.e., a genotype with $S = 0$ is insensitive while a genotype with a large positive S shows strong vernalization response. The daily development rate k_{ij} (d⁻¹) for the i^{th} day and j^{th} line was calculated based on daily mean air temperatures T_i as follows:

$$k_{ij} = \frac{T_i}{(T_i \times S_j + T_{0j})}$$

k_{ij} was added up until the cumulative k_j became ≥ 1 , i.e., when the adult vegetative phase was completed and curd development started.

A linkage map was constructed with MapMaker 3 by simultaneous multipoint-analysis using the Haldane function (Lander et al., 1987; Lincoln et al., 1993). On basis of the results of parameterization trials, QTL analyses were conducted on LAR , S , and T_0 using PlabQTL 1.2 (Utz and Melchinger, 1996). Composite interval mapping (CIM) based on multiple regression with co-factors was applied (Haley and Knott, 1992; Jansen and Stam, 1994; Zeng, 1994; Miller, 2002). The LOD threshold to

define a significant QTL was 2.5. The linkage map was drawn with MapChart 2.3 (Voorrips, 2002).

Additive QTL effects were assigned to all detected loci and QTL based parameters for T_0 , S , and LAR were calculated using the following equation:

$$y_j = m + \sum e_i g_{ij}$$

where y_j is the estimated value of parameter y of the j^{th} line, m is the population mean for the parameter, e_i is the additive effect of the i^{th} QTL, and g_{ij} is the allelic state of the i^{th} QTL in line j . g_{ij} can be -1 or $+1$.

In addition to linkage mapping, marker effects of the three parameters were estimated by genomic selection using the ridge-regression best linear unbiased prediction (rrBLUP) package for R 3.2.4. with the following restricted mixed model (REML) solution:

$$y = X\beta + Zv + \varepsilon$$

where y is the vector of phenotypic observations, X is the matrix of fixed effects β , Z is the matrix of random effects v , and ε is the residual effects matrix (Endelman, 2011). rrBLUP was carried out on the three model parameters LAR , S , and T_0 . Effects were assigned to all markers and the sum of marker effects was used to calculate additive genomic estimated breeding values (GEBV) of the individuals of the validation sets. GEBVs for LAR , S , and T_0 were implemented into the phenology model as parameter input values to simulate heading date of the individuals in field trials.

TABLE 2 | Population mean, standard deviation, and maximum and minimum estimates for leaf appearance rate (LAR), slope (S), and intercept (T_0) of DH lines used for model parameterization and estimates of the parental lines P1 and P2.

Parameter	Mean	SD	Max.	Min.	P1	P2
LAR	0.0216	0.0021	0.0293	0.0161	0.0161	0.0214
S	50.0	20.5	106.0	4.17	39.8	46.7
T_0	631.3	296.5	1374.6	-194.6	1133.2	840.3

TABLE 3 | Number of lines of the parameterization and validation sets used in field trials at different locations in 2011, 2012, and 2013; mean, minimum and maximum temperatures during field trials and means with standard deviations (SD), minima and maxima for time to curd induction.

Trial	Zeewolde 2011	Zeewolde 2012	Ruthe 2013	Rostock 2012	Ócsa 2011	Ócsa 2012
Number of lines						
Parameterization set (DH lines <i>per se</i>)	140	154	-	-	-	-
Validation set (DH lines <i>per se</i>)	44	69	68	67	-	-
Parameterization set (F1 hybrids)	-	156	-	-	134	146
Validation set (F1 hybrids)	-	31	-	-	20	24
Temperature [°C]						
Mean	17.4	17.2	19.0	17.9	21.4	22.0
Min.	13.0	9.1	12.1	11.7	14.9	13.6
Max.	25.4	21.9	26.9	27.4	29.6	29.4
Thermal time to curd induction [°Cd]						
Mean	1276.3	1261.7	1417.0	1494.7	1427.4	1597.7
SD	102.6	73.5	139.4	99.5	82.9	135.2
Min.	1075.3	1006.8	1178.2	1237.9	1291.3	1364.0
Max.	1477.2	1509.6	1810.6	1750.4	1621.3	1868.7

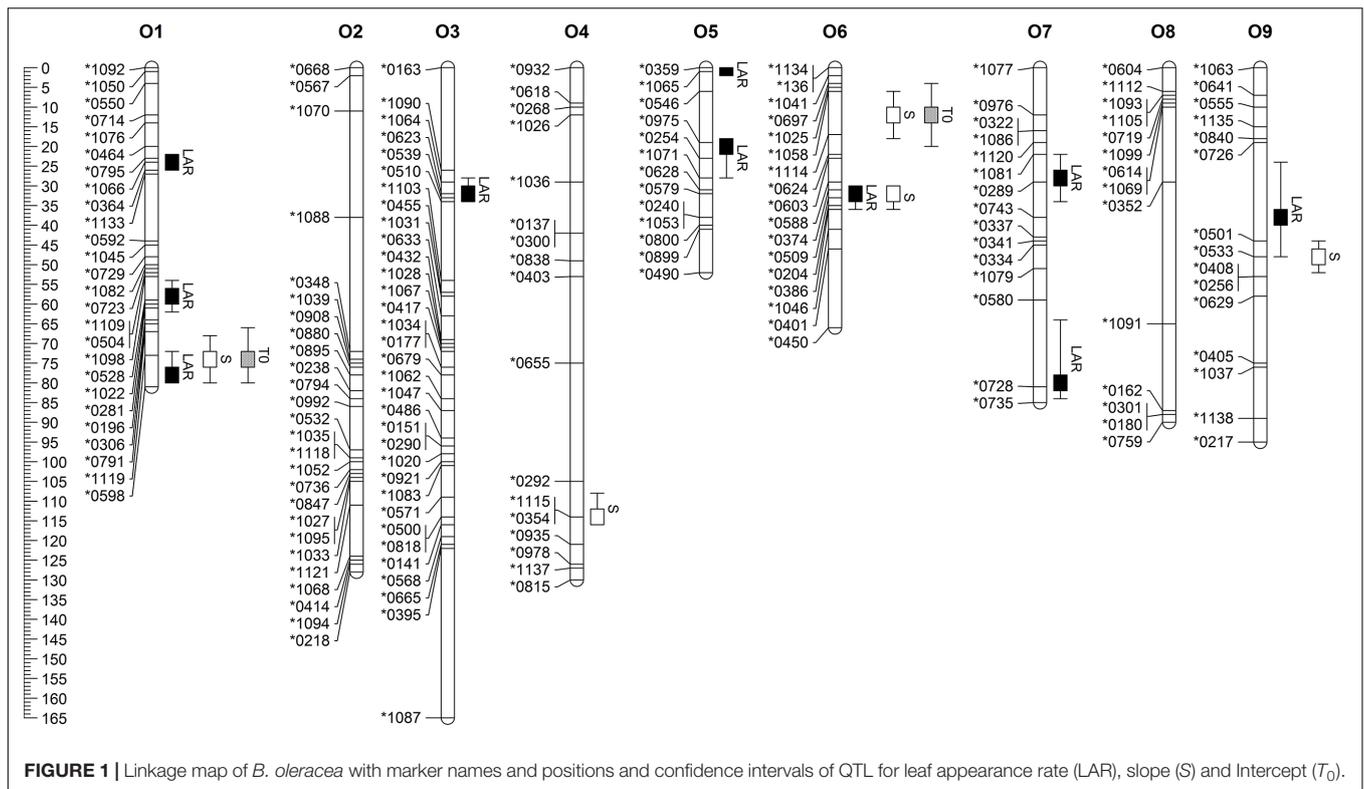


TABLE 4 | List of QTL for leaf appearance rate (LAR), slope (S), and intercept (T_0) including confidence intervals, LOD-scores, % variation explained (R^2) and effects of the P1 allele.

Trait	Chromosome	Position (cM)	Confidence interval (cM)	Nearest marker	LOD-score	R^2 (%)	Additive effect
LAR	1	24	22–26	S1066	4.68	12.5	−0.504
	1	58	54–62	S0528	6.62	17.2	0.795
	1	78	72–80	S0598	2.70	7.5	−0.508
	3	32	28–34	S0623	3.42	9.3	−0.388
	5	0	0–2	S0359	2.70	7.6	−0.425
	5	20	18–28	S0975	4.61	12.4	0.541
	6	32	30–36	S0509	2.99	8.2	0.593
	7	28	22–34	S0289	4.09	11.0	0.504
	7	80	64–84	S0728	2.96	8.1	−0.398
S	9	38	24–48	S0501	7.04	18.3	0.623
	1	74	68–80	S1119	4.12	11.7	−4.972
	4	114	108–116	S1115	4.42	12.4	−5.217
	6	12	46–18	S1114	8.00	21.3	−8.885
	6	32	30–36	S0509	4.55	12.7	−6.891
T_0	9	48	44–52	S0501	3.98	11.2	−4.797
	1	74	66–80	S1119	2.91	8.5	64.62
	6	12	4–20	S1114	4.49	12.6	100.54

RESULTS

Phenotypic Variation for Traits and Model Parameters

Thermal time from transplanting to curd induction under different constant temperatures was measured in the parameterization trials. In general, the higher the temperature,

the higher was the temperature sum required for curd induction. Population means ranged from 753 to 1590°Cd (**Table 1**).

Mean LAR of individual genotypes averaged across all parameterization trials ranged from 0.016 to 0.029. The population mean was 0.022 with a standard deviation of ± 0.002 . Mean S was 50.0 (± 20.5) and ranged from 4.2 to 106.0. Population mean of T_0 was 631°Cd ± 296 °Cd (**Table 2**). Mean R^2

of the relation between temperature and thermal time to curd induction was $0.65 (\pm 0.13)$, maximum R^2 was 0.87 and minimum was 0.05.

In the validation trials the thermal time to curd induction was measured for the parameterization set, the validation set and F1 hybrids of both sets. Mean thermal time to curd induction ranged from 1276.3 to 1597.7°Cd across all trials and the different subsets (Table 3).

QTL and Genome-Wide Marker Effects for Model Parameters

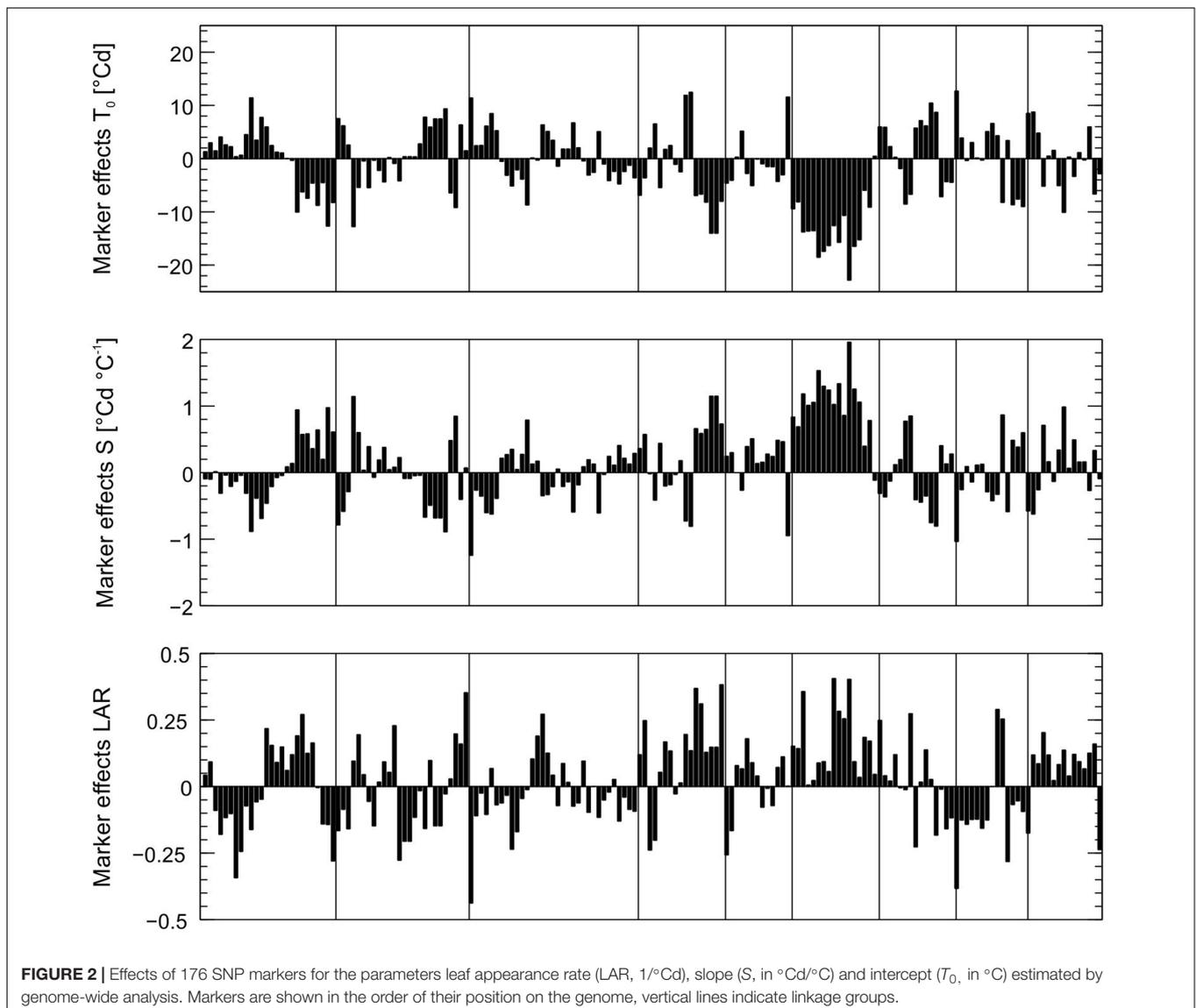
QTL analyses were carried out on model parameters LAR, S , and T_0 . In total, ten QTL were detected for LAR. Three of them were located on chromosome O1 and two on O5 (Figure 1). Five QTL were identified for S . Both S QTL on O6 co-localized with T_0 QTL, while one of them was also found to co-localize with an LAR QTL. The third QTL for S on O9 was located close to a LAR QTL.

Genetic positions, LOD scores, and additive effects of all QTL are summarized in Table 4. LAR QTL on O9 and O5 revealed highest LOD scores of 7.04 and 6.62, respectively. LOD scores of the two QTL for S on O6 were found to be highest with 8.0 and 4.6.

Genome-wide marker effects for the model parameters LAR, S , and T_0 were estimated by rrBLUP and are displayed in Figure 2. Large marker effects on O6 support linkage-mapping results, which indicated a promising hotspot region on O6.

Genome-Based Prediction of Curding Time

To get an idea about the power of the phenological model, curd induction of the parameterization set was simulated with both QTL-effects and original parameters. A very high prediction accuracy of $R^2 = 0.98$ was observed applying original parameters to predict time to curd induction of the parameterization trials. Figures 3A,C show means across seven temperature treatments.



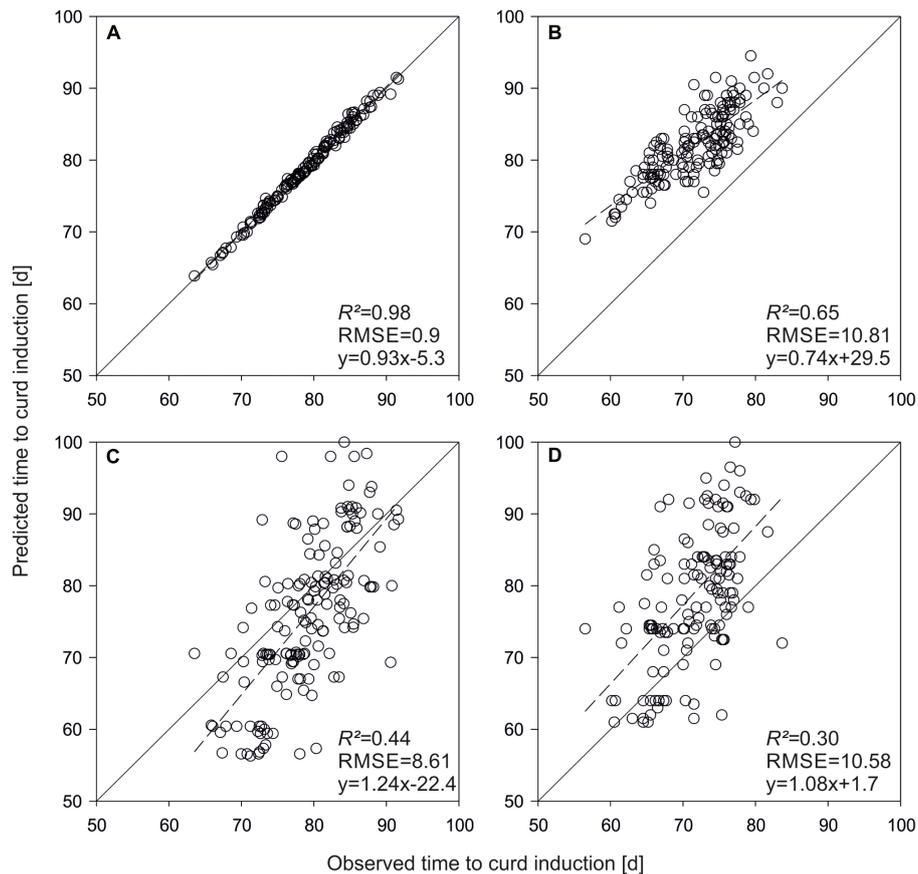


FIGURE 3 | Predicted vs. observed mean time to curd induction of the DH lines used for model parameterization. Shown are means of seven greenhouse trials (**A,C**) and means of field trials in Zeewolde 2011 and 2012 (**B,D**). Predictions are based on original parameters (**A,B**) and QTL effects (**C,D**).

R^2 decreased to 0.44 if original parameters were replaced by QTL-based estimates. A similar situation with an overall reduced R^2 was observed if mean time to curd induction of field data from Zeewolde 2011 and 2012 was simulated (**Figures 3B,C**). R^2 diminished to 0.47 if mean curd induction of F1 hybrids derived from the parameterization set in field trials in Ócsa 2011, 2012 and Zeewolde 2012 was modeled using original parameters (data not shown).

For model evaluation, QTL effects were used as parameter inputs to predict time to curd induction of the validation set. **Figure 4** shows predicted vs. observed time to curd induction of different field trials. In Rostock, mean observed time to curd induction was 83.5 days (± 11.5), while predicted time to curd induction was 79.1 days (± 4.9). The coefficient of determination was $R^2 = 0.45$ (**Figure 4A**). With $R^2 = 0.47$, prediction accuracy was slightly higher in Ruthe (**Figure 4B**), while the model performed best in the Zeewolde trial in 2011 ($R^2 = 0.51$, **Figure 4C**). Interestingly, prediction of independent lines resulted in higher R^2 values if compared to the parameterization set. However, the slope was 1.08 for the parameterization set (**Figure 3D**) but ranged between 1.22 and 1.54 when curding time of independent lines was modeled (**Figure 4**).

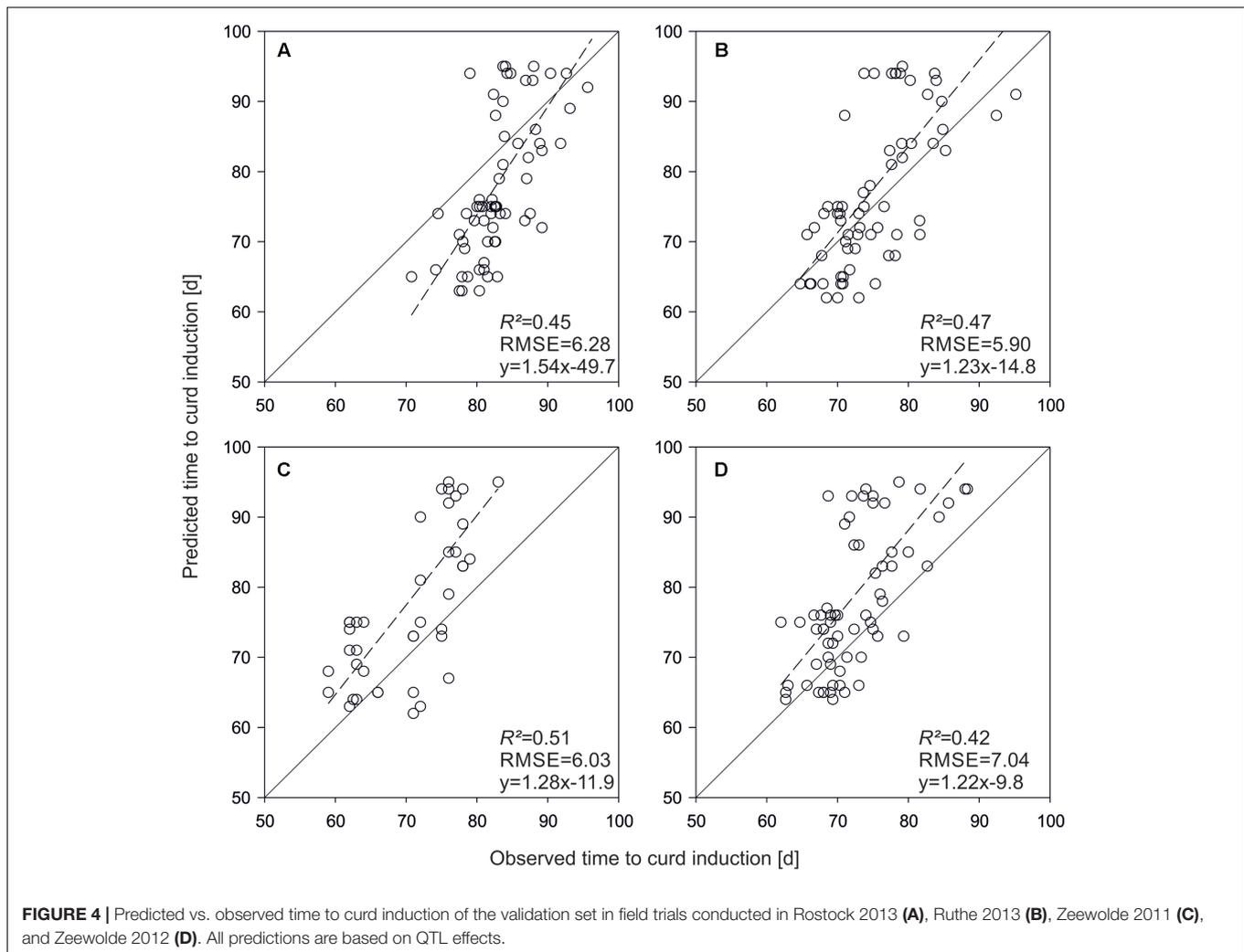
As an alternative method, genome-wide marker effects were estimated. The GS model led to higher prediction accuracies in all field trials if applied on the validation set (**Figure 5**). R^2 increased on average by 10.3%. However, a relatively strong bias was observed in the Zeewolde trials (**Figures 5C,D**).

Applying the QTL-based model on test hybrids developed from the parameterization set resulted in moderate prediction accuracies ($R^2 = 0.34$, **Figure 6A**). If QTL effects were replaced by genome-wide marker effects, R^2 increased to 0.50 (**Figure 6C**). The QTL based simulation of time to curd induction of F1-hybrids derived from the validation set revealed also moderately good results ($R^2 = 0.41$, **Figure 6B**), which were outperformed by the GS based model ($R^2 = 0.48$; **Figure 6D**).

DISCUSSION

QTL Analyses and Genomic Selection

QTL were found for all examined parameters. QTL for S and T_0 were almost congruent in their positions, which was not surprising, as S and T_0 were highly correlated ($r = -0.95$) and lines with a steep slope do have a low intercept and *vice versa*. Similar to the present study, Uptmoor et al. (2008) found QTL

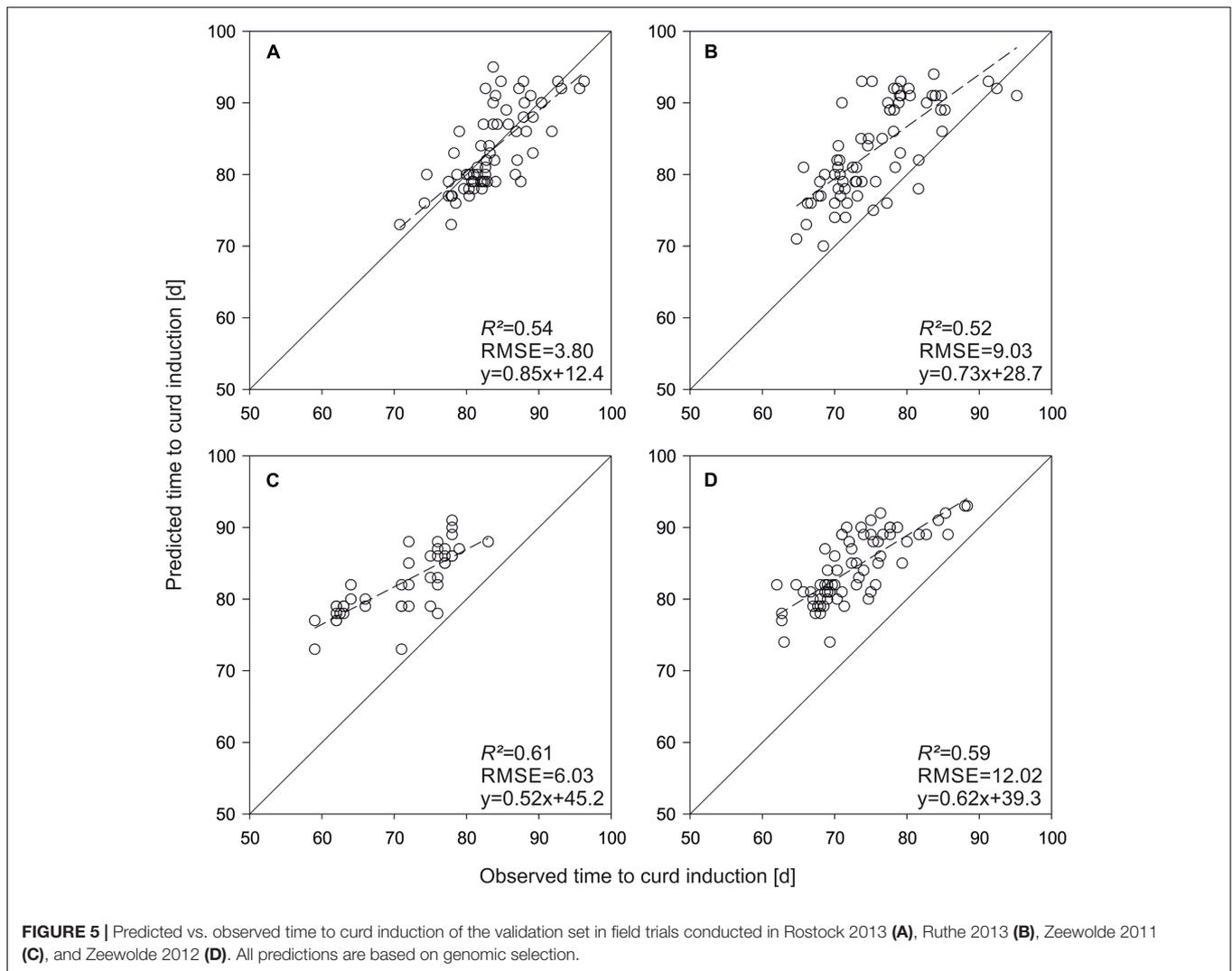


for ‘slope’ on chromosomes O9 and O6 in a *B. oleracea* var. *alboglabra* × *B. oleracea* var. *italica* cross. QTL for ‘slope’ were not detected on chromosomes O4 or O1 but an additional QTL was identified on O3, which was not confirmed in the present study. Uptmoor et al. (2008) detected four QTL for the parameter ‘intercept.’ One was also located O6 but three further QTL on O3, O4, and O5 were not verified in this study. Even if positions are not completely comparable since both studies used different markers and linkage maps, clear parallels between the detected QTL are evident.

At the positions of three of the five QTL for slope, across environment QTL with significant QTL × environment interactions were detected in an earlier study using the same population and phenotype data (Hasan et al., 2016), indicating that input parameters of ecophysiological models are suitable traits for QTL detection. In contrast to the statistical interaction of a QTL with its environments (i.e., different effect sizes in different environments), model parameters display the trait’s response to an environmental factor allowing precise estimations of a new genotype in untested environments (Onogi et al., 2016). The QTL regions on O6 harbor *BoAPI-c* and *BoAPI-a*, the latter

of which was suggested to interact with temperature (Labate et al., 2006). The QTL region on O9 harbors an *FLC* paralog and a *FRIGIDA LIKE (FRL)* ortholog (Hasan et al., 2016).

Usually, a relatively large sample size is required for QTL analyses and several hundreds of lines are recommended for good results (Hyne et al., 1995; Beavis, 1998; Lynch and Walsh, 1998). In general, the higher the number of genotypes used for analyses, the more accurate are the results. QTL with relatively small effects, so called minor QTL, are detected only in large populations, while major QTL can still be found with fewer genotypes (Vales et al., 2005; Slate, 2005), which, however, leads to an underestimation of the total number of QTL, while QTL effects are overestimated since a part of the variation is due to undiscovered minor QTL (Beavis, 1994; Xu, 2003). All significant QTL were tested for QTL×QTL interactions by pairwise linear regression analysis. The only significant interaction was detected for the parameter slope (QTL at 114 cM on O4 and at 12 cM on O6). However, including the interaction effect into the model improved estimations for the slope only slightly and had no positive impact on the simulation studies (data not shown). Hill et al. (2008) found that additive variance accounts often for close



to 100% of the total genetic variance, which might be the reason for the missing positive impact of QTLxQTL interactions on the QTL-based simulation model of the present study.

Genomic selection is a useful tool to overcome these limitations as effects are assigned to all markers. GS was used as an alternative method for model parameterization and regions with large marker effects on O4 and O6 reflect findings from linkage mapping quite well; but in contrast to the QTL model, GS still accounts for genome regions with minor effects.

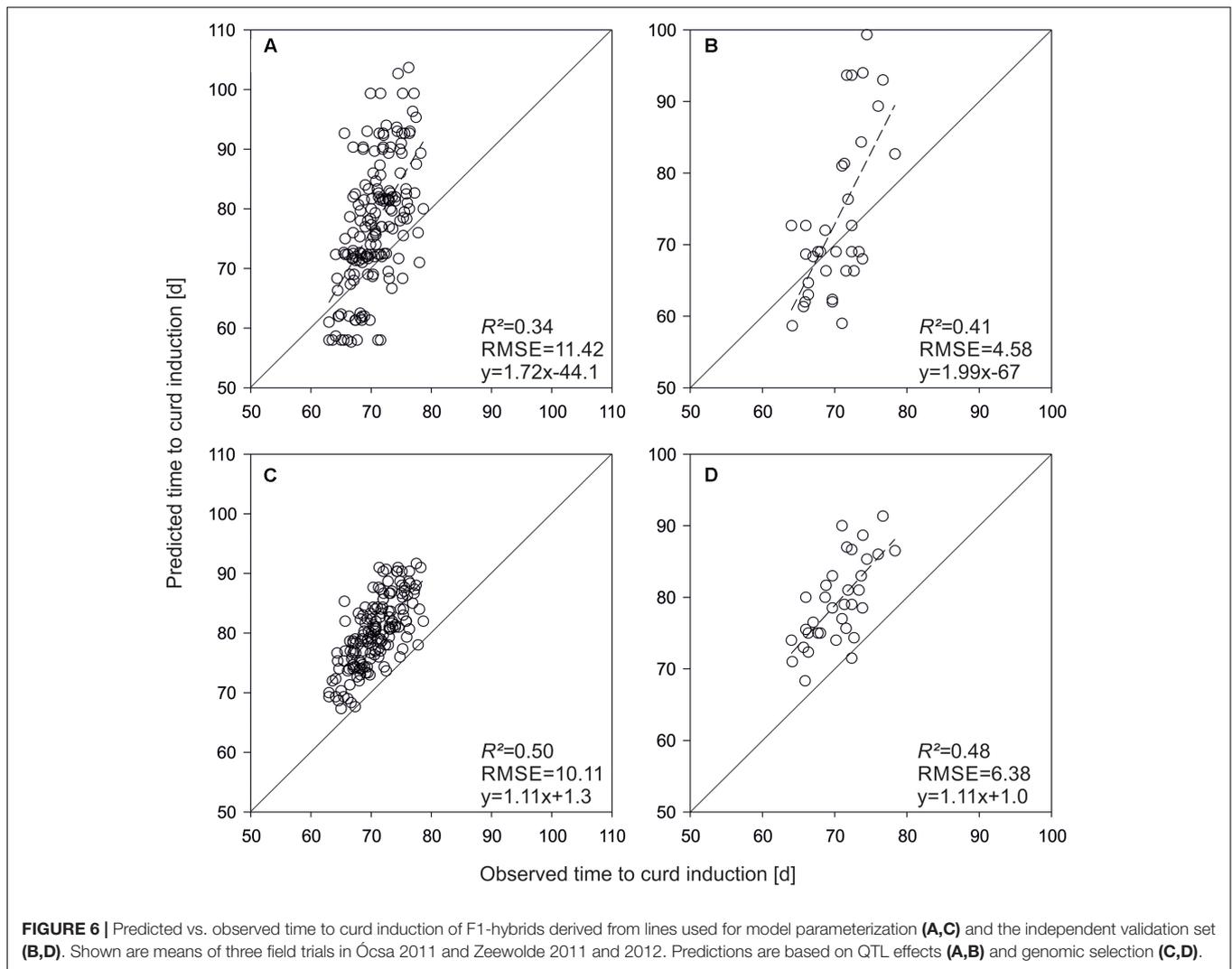
Genome-Based Simulation Models

The original model was able to predict time to curd induction by means of chosen parameters in a satisfactory manner. The possibility to simulate cauliflower development from temperature data only has already been shown (Olesen and Grevsen, 2000). The QTL-based approach was rather successful as well, although predictions based on original parameters were more precise. Assumptions made by Reymond et al. (2003) and Uptmoor et al. (2008) that replacement of model parameters by QTL effects would lead to higher prediction accuracies due to the reduction

of random errors were not confirmed in the present study. While the prediction of hybrids derived from the DH population was possible, prediction accuracies were not as high as those of DH lines *per se*.

Considering simulation studies carried out on independent genotypes of the cross, results adequately reflected expectations if compared to similar studies without independent genotypes (Quilot et al., 2005; Yin et al., 2005a; White et al., 2008; Uptmoor et al., 2012). However, Bogard et al. (2014) attained relatively high prediction accuracies (mean $R^2 = 0.58$) for predicting heading date of independent bread wheat genotypes using a common phenology model and genome-wide association studies for QTL detection. Less precise predictions of the present study may result from existing variation in juvenile-to-adult transition, for which the curd-induction model did not account.

Unexpectedly, the genome-based prediction of independent DH lines exceeded the accuracy of the simulations of lines used for parameterization and even the prediction of F1 hybrids resulted in a slightly higher R^2 . If subsamples of a population are large enough, different random subsamples will provide similar



results regarding QTL positions and effects, hence, prediction accuracies for independent lines are likely to be on the same level as for lines used as parameterization set. Therefore, higher accuracies for independent lines in this study can only be explained by chance. In fact only R^2 values were higher, while slopes > 1 illustrate that the variability in curding time within the validation set was overestimated.

Simulations using genome-wide marker effects led to more precise predictions in comparison to QTL-based simulations, which held true for both independent DH lines of the validation set and test hybrids. The major reason might be explained by missing effects of minor loci in the QTL model. Comparing different models to predict hybrid performance in maize, Guo et al. (2013) found the genome-wide prediction approach being more accurate than predictions based on QTL effects.

In the present study, a simple approach with only three model parameters was used to predict time to curd induction of cauliflower lines and hybrids. Combining GS and ecophysiological models could reduce efforts spent in field trials to test germplasm across wide ranges of different environments

in silico since GS models are able to predict phenotypes only in relatively similar environments while integrated models account for both genotypic and environmental variations in equal measure. Even though model parameterization requires large phenotyping efforts under controlled conditions, more precise predictions may be attained from more complex models.

The uncertainty about juvenile-to-adult transition might be considered as main error source impairing prediction accuracies. Juvenility is assumed to be independent from temperature and to end after a certain number of leaves are initiated (Salter and James, 1974; Hand and Atherton, 1987). Since large variations in juvenile-to-adult transition are present in cauliflower cultivars (Booij and Struik, 1990), certain variation may exist in the examined population as well. Genotype specific estimations of leaf numbers at the end of juvenility should be made in further experiments via reciprocal transfers to improve the model's accuracy by integrating the onset of sensitivity to temperature.

In the present model, LAR was assumed to be constant throughout the vegetative phase. However, it is known that LAR changes with time and more sophisticated models already took

this into account (Kage and Stützel, 1999). Similar approaches are conceivable to further improve simulation accuracies of integrated modeling approaches.

While the model of the present study simulates the development of cauliflower until visible curd induction, other phenology models already implemented curd development to predict harvest time (Pearson et al., 1994; Olesen and Grevsen, 2000). The practical use of a model predicting harvest time instead of time of curd induction is undoubtedly higher since accurate harvest-time predictabilities in diverse environments are a great challenge not only for growers but also for breeders.

CONCLUSION

We conclude from the present study that parameters describing the response of a genotype to an environmental factor are suitable traits for QTL detection. Identified loci reflect across environment QTL with significant QTL \times environment interactions, while having beneficial effects since the response to the influencing factor is quantifiable. The latter allows predicting untested genotypes in a broad range of untested environments. Simulation studies carried out on independent genotypes

REFERENCES

- Axelsson, T., Shavorskaya, O., and Lagercrantz, U. (2001). Multiple flowering time QTLs within several Brassica species could be the result of duplicated copies of one ancestral gene. *Genome* 44, 856–864. doi: 10.1139/g01-082
- Beavis, W. D. (1994). “The power and deceit of QTL experiments: lessons from comparative QTL studies” in *Proceedings of the Forty-Ninth Annual Corn and Sorghum Industry Research Conference*, Washington, DC, 250–266.
- Beavis, W. D. (1998). “QTL analyses: power, precision, and accuracy,” in *Molecular Dissection of Complex Traits*, ed. A. H. Paterson (Boca Raton, FL: CRC Press), 145–162.
- Blázquez, M. A. (2005). Plant science. The right time and place for making flowers. *Science* 309, 1024–1025. doi: 10.1126/science.1117203
- Bogard, M., Ravel, C., Paux, E., Bordes, J., Balfourier, F., Chapman, S., et al. (2014). Predictions of heading date in bread wheat (*Triticum aestivum* L.) using QTL-based parameters of an ecophysiological model. *J. Exp. Bot.* 65, 5849–5865. doi: 10.1093/jxb/eru328
- Booij, R., and Struik, P. C. (1990). Effects of temperature on leaf and curd initiation in relation to juvenility of cauliflower. *Sci. Hortic.* 44, 201–214. doi: 10.1016/0304-4238(90)90120-4
- Brown, A. F., Jeffery, E. H., and Juvik, J. A. (2007). A polymerase chain reaction-based link-age map of broccoli and identification of quantitative trait loci associated with harvest date and head weight. *J. Hortic. Sci.* 132, 507–513.
- Destá, Z. A., and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19, 1360–1385. doi: 10.1016/j.tplants.2014.05.006
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R pack-age rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024
- Grevsen, K., and Olesen, J. E. (1994). Modelling cauliflower development from transplanting to curd initiation. *J. Hortic. Sci.* 69, 755–766. doi: 10.1080/14620316.1994.11516510
- Guo, T., Li, H., Yan, J., Tang, J., Li, J., Zhang, Z., et al. (2013). Performance prediction of F1 hybrids between recombinant inbred lines derived from two elite maize inbred lines. *Theor. Appl. Genet.* 126, 189–201. doi: 10.1007/s00122-012-1973-9
- Haley, C. S., and Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69, 315–324. doi: 10.1038/hdy.1992.131
- Hand, D. J. (1988). *Regulation of Curd Initiation in the Summer Cauliflower*. Nottingham: University of Nottingham.
- Hand, D. J., and Atherton, J. G. (1987). Curd initiation in the cauliflower. *J. Exp. Bot.* 38, 2050–2058. doi: 10.1093/jxb/38.12.2050
- Hasan, Y., Briggs, W., Matschegewski, C., Ordon, F., Stützel, H., Zetsche, H., et al. (2016). Quantitative trait loci controlling leaf appearance and curd initiation of cauliflower in relation to temperature. *Theor. Appl. Genet.* 129, 1273–1288. doi: 10.1007/s00122-016-2702-6
- Heslot, N., Akdemir, D., Sorrells, M. E., and Jannink, J. L. (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* 127, 463–480. doi: 10.1007/s00122-013-2231-5
- Hill, W. G., Goddard, M. E., and Visscher, P. M. (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLOS Genet.* 4:e1000008. doi: 10.1371/journal.pgen.1000008
- Hyne, V., Kearsey, M. J., Pike, D. J., and Snape, J. W. (1995). QTL analysis: unreliability and bias in estimation procedures. *Mol. Breed.* 1, 273–282. doi: 10.1007/BF02277427
- Jack, T. (2004). Molecular and genetic mechanisms of floral control. *Plant Cell* 16(Suppl.), S1–S17. doi: 10.1105/tpc.017038
- Jansen, R. C., and Stam, P. (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* 136, 1447–1455.
- Jensen, A. L., and Grevsen, K. (2005). Web-based harvest forecast of cauliflower for the agri-food-chain. *Acta Hortic.* 674, 329–335. doi: 10.17660/ActaHortic.2005.674.39
- Kage, H., Alt, C., and Stützel, H. (2001a). Predicting dry matter production of cauliflower (*Brassica oleracea* L. *botrytis*) under unstressed conditions: I. Photosynthetic parameters of cauliflower leaves and their implications for calculations of dry matter production. *Sci. Hortic.* 87, 155–170. doi: 10.1016/S0304-4238(00)00177-1
- Kage, H., Stützel, H., and Alt, C. (2001b). Predicting dry matter production of cauliflower (*Brassica oleracea* L. *botrytis*) under unstressed conditions: Part II. Comparison of light use efficiency and photosynthesis-respiration based modules. *Sci. Hortic.* 87, 171–190. doi: 10.1016/S0304-4238(00)00180-1
- Kage, H., and Stützel, H. (1999). A simple empirical model for predicting development and dry matter partitioning in cauliflower (*Brassica oleracea* L. *botrytis*). *Sci. Hortic.* 80, 19–38. doi: 10.1016/S0304-4238(98)00226-X

AUTHOR CONTRIBUTIONS

The trials were carried out by AR, YH, and WB. Experiments were designed by RU and the data was analyzed by AR and YH. The manuscript was written by AR, revised by RU and approved by WB.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the German Federal Ministry of Education and Research (BMBF). The project is part of the AgroCluster WeGa Kompetenznetz Gartenbau (project no. 0315542A). We acknowledge financial support by Deutsche Forschungsgemeinschaft and Universität Rostock within the funding program Open Access Publishing.

- Koornneef, M., Blankestijn-de Vries, H., Hanhart, C., Soppe, W., and Peeters, T. (1994). The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg erecta wild-type. *Plant J.* 6, 911–919. doi: 10.1046/j.1365-313X.1994.6060911.x
- Labate, J. A., Robertson, L. D., Baldo, A. M., and Bjorkman, T. (2006). Inflorescence identity genes alleles are poor predictors of inflorescence type in broccoli and cauliflower. *J. Am. Soc. Hortic. Sci.* 131, 667–673.
- Lagercrantz, U., Putterill, J., Coupland, G., and Lydiate, D. (1996). Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. *Plant J.* 9, 13–20. doi: 10.1046/j.1365-313X.1996.09010013.x
- Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E., et al. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1, 174–181. doi: 10.1016/0888-7543(87)90010-3
- Lincoln, S. E., Daly, M. J., and Lander, E. S. (1993). *Constructing Genetic Linkage Maps with MAPMAKER/EXP Version 3.0: A Tutorial and Reference Manual*. A Whitehead Institute for Biomedical Research Technical Report. Cambridge, MA: Whitehead Institute, 78–79.
- Lorenza, R. E., and Bernardo, R. (2009). Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theor. Appl. Genet.* 120, 151–161. doi: 10.1007/s00122-009-1166-3
- Lynch, M., and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland: Sinauer.
- Matschegewski, C., Zetsche, H., Hasan, Y., Leibeguth, L., Briggs, W., Ordon, F., et al. (2015). Genetic variation of temperature-regulated curd induction in cauliflower: elucidation of floral transition by genome-wide association mapping and gene expression analysis. *Front. Plant Sci.* 6:720. doi: 10.3389/fpls.2015.00720
- Messina, C. D., Jones, J. W., Boote, K. J., and Vallejos, C. E. (2006). A gene-based model to simulate soybean development and yield responses to environment. *Crop Sci.* 46, 456–466. doi: 10.2135/cropsci2005.04-0372
- Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.
- Miller, A. J. (2002). *Subset Selection in Regression*, 2nd Edn. Boca Raton, FL: Chapman & Hall. doi: 10.1201/9781420035933
- Okazaki, K., Sakamoto, K., Kikuchi, R., Saito, A., Togashi, E., Kuginuki, Y., et al. (2007). Mapping and characterization of FLC homologs and QTL analysis of flowering time in *Brassica oleracea*. *Theor. Appl. Genet.* 114, 595–608. doi: 10.1007/s00122-006-0460-6
- Olesen, J. E., and Grevsen, K. (2000). A simulation model of climate effects on plant productivity and variability in cauliflower (*Brassica oleracea* L. *botrytis*). *Sci. Hortic.* 83, 83–107. doi: 10.1016/S0304-4238(99)00068-0
- Onogi, A., Watanabe, M., Mochizuki, T., Hayashi, T., Nakagawa, H., Hasegawa, T., et al. (2016). Toward integration of genomic selection with crop modelling: the development of an integrated approach to predicting rice heading dates. *Theor. Appl. Genet.* 129, 805–817. doi: 10.1007/s00122-016-2667-5
- Pearson, S., Hadley, P., and Wheldon, A. E. (1994). A model of the effects of temperature on the growth and development of cauliflower (*Brassica oleracea* L. *botrytis*). *Sci. Hortic.* 59, 91–106. doi: 10.1016/0304-4238(94)90076-0
- Quilot, B., Kervella, J., Génard, M., and Lescourret, F. (2005). Analysing the genetic control of peach fruit quality through an ecophysiological model combined with a QTL approach. *J. Exp. Bot.* 56, 3083–3092. doi: 10.1093/jxb/eri305
- Razi, H., Howell, E. C., Newbury, H. J., and Kearsley, M. J. (2008). Does sequence polymorphism of FLC paralogues underlie flowering time QTL in *Brassica oleracea*? *Theor. Appl. Genet.* 116, 179–192. doi: 10.1007/s00122-007-0657-3
- Reymond, M., Muller, B., Leonardi, A., Charcosset, A., and 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. doi: 10.1104/pp.013839
- Ridge, S., Brown, P. H., Hecht, V., Driessen, R. G., and Weller, J. L. (2015). The role of BoFLC2 in cauliflower (*Brassica oleracea* var. *botrytis* L.) reproductive development. *J. Exp. Bot.* 66, 125–135. doi: 10.1093/jxb/eru408
- Salter, P. J. (1969). Studies on crop maturity in cauliflower. I. Relationship between the times of curd initiation and curd maturity of plants within a cauliflower crop. *J. Hortic. Sci.* 44, 129–140. doi: 10.1080/00221589.1969.11514301
- Salter, P. J., and James, J. M. (1974). Further studies on the effects of cold treatment of transplants on crop maturity characteristics of cauliflower. *J. Hortic. Sci.* 49, 329–342. doi: 10.1080/00221589.1974.11514587
- Schranz, M. E., Quijada, P., Sung, S.-B., Lukens, L., Amasino, R., and Osborn, T. C. (2002). Characterization and effects of the replicated flowering time Gene FLC in *Brassica rapa*. *Genetics* 162, 1457–1468.
- Slate, J. (2005). Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Mol. Ecol.* 14, 363–379. doi: 10.1111/j.1365-294X.2004.02378.x
- Technow, F., Messina, C. D., Totir, L. R., and Cooper, M. (2015). Integrating crop growth models with whole genome prediction through approximate Bayesian computation. *PLOS ONE* 10:e0130855. doi: 10.1371/journal.pone.0130855
- Uptmoor, R., Li, J., Schrag, T., and Stützel, H. (2012). Prediction of flowering time in *Brassica oleracea* using a quantitative trait loci-based phenology model. *Plant Biol.* 14, 179–189. doi: 10.1111/j.1438-8677.2011.00478.x
- Uptmoor, R., Pillen, K., and Matschegewski, C. (2016). Combining genome-wide prediction and a phenology model to simulate heading date in spring barley. *Field Crops Res.* 202, 84–93. doi: 10.1016/j.fcr.2016.08.006
- Uptmoor, R., Schrag, T., Stützel, H., and 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. doi: 10.1007/s11032-007-9121-y
- Utz, H. F., and Melchinger, A. E. (1996). PLABQTL: a program for composite interval mapping of QTL. *J. Quant. Trait Loci* 2, 1–5. doi: 10.1094/PHYTO.2004.94.8.862
- Vales, M. I., Schön, C. C., Capettini, F., Chen, X. M., Corey, A. E., Mather, D. E., et al. (2005). Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. *Theor. Appl. Genet.* 111, 1260–1270. doi: 10.1007/s00122-005-0043-y
- Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78. doi: 10.1093/jhered/93.1.77
- White, J. W., Herndl, M., Hunt, L. A., Payne, T. S., and Hoogenboom, G. (2008). Simulation-based analysis of effects of Vrn and Ppd loci on flowering in wheat. *Crop Sci.* 48, 678–687. doi: 10.2135/cropsci2007.06.0318
- Wiebe, H. J. (1972a). Wirkung von Temperatur und Licht auf Wachstum und Entwicklung von Blumenkohl. I.: Dauer der Jugendphase für die Vernalisation. *Gartenbauwissenschaft* 37, 165–178.
- Wiebe, H. J. (1972b). Wirkung von temperatur und licht auf Wachstum und Entwicklung von Blumenkohl. II.: optimale Vernalisationstemperatur und Vernalisationsdauer. *Gartenbauwissenschaft* 37, 293–303.
- Wiebe, H. J. (1972c). Wirkung von temperatur und licht auf Wachstum und Entwicklung von Blumenkohl. III.: vegetative phase. *Gartenbauwissenschaft* 37, 455–469.
- Wiebe, H. J. (1980). Production of cauliflower for continuous supply of the market during the harvest season. *Gartenbauwissenschaft* 45, 282–288.
- Wurr, D. C. (1989). Prediction of the time of maturity in cauliflowers: VI symposium on the timing of field production of vegetables. *Acta Hortic.* 267, 387–394.
- Wurr, D. C., Fellows, J. R., and Hiron, R. W. (1990a). The influence of field environmental conditions on the growth and development of four cauliflower cultivars. *J. Hortic. Sci.* 65, 565–572. doi: 10.1080/00221589.1990.11516094
- Wurr, D. C., Fellows, J. R., Sutherland, R. A., and Elphinstone, E. D. (1990b). A model of cauliflower curd growth to predict when curds reach a specified size. *J. Hortic. Sci.* 65, 555–564. doi: 10.1080/00221589.1990.11516093
- Xu, S. (2003). Theoretical basis of the Beavis effect. *Genetics* 165, 2259–2268.

- Yin, X., Stam, P., Kropff, M. J., and Schapendonk, A. H. (2003). Crop modeling, QTL map-ping, and their complementary role in plant breeding. *Agron. J.* 95, 90–98. doi: 10.2134/agronj2003.9000
- Yin, X., Struik, P. C., Tang, J., Qi, C., and Liu, T. (2005a). Model analysis of flowering phenology in recombinant inbred lines of barley. *J. Exp. Bot.* 56, 959–965.
- Yin, X., Struik, P. C., van Eeuwijk Fred, A., Stam, P., and Tang, J. (2005b). QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley. *J. Exp. Bot.* 56, 967–976.
- Zeng, Z.-B. (1994). Precision mapping of quantitative trait loci. *Genet. Soc. Am.* 136, 1457–1468.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Rosen, Hasan, Briggs and Uptmoor. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.