Theor Appl Genet (2016) 129:805–817
DOI 10.1007/s00122-016-2667-5

ORIGINAL ARTICLE

Toward integration of genomic selection with crop modelling: the development of an integrated approach to predicting rice heading dates

Akio Onogi[1] · Maya Watanabe[1] · Toshihiro Mochizuki[2] · Takeshi Hayashi[3] · Hiroshi Nakagawa[3] · Toshihiro Hasegawa[4] · Hiroyoshi Iwata[1]

Received: 19 August 2015 / Accepted: 9 January 2016 / Published online: 20 January 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract

Key message **It is suggested that accuracy in predicting plant phenotypes can be improved by integrating genomic prediction with crop modelling in a single hierarchical model.**

Abstract Accurate prediction of phenotypes is important for plant breeding and management. Although genomic prediction/selection aims to predict phenotypes on the basis of whole-genome marker information, it is often difficult to predict phenotypes of complex traits in diverse environments, because plant phenotypes are often influenced by genotype–environment interaction. A possible remedy is to integrate genomic prediction with crop/ecophysiological modelling, which enables us to predict plant phenotypes using environmental and management information. To this end, in the present study, we developed a novel method for integrating genomic prediction with phenological modelling of Asian rice (*Oryza sativa*, L.), allowing the heading date of untested genotypes in untested environments to be predicted. The method simultaneously infers the phenological model parameters and whole-genome marker effects on the parameters in a Bayesian framework. By cultivating backcross inbred lines of Koshihikari × Kasalath in nine environments, we evaluated the potential of the proposed method in comparison with conventional genomic prediction, phenological modelling, and two-step methods that applied genomic prediction to phenological model parameters inferred from Nelder–Mead or Markov chain Monte Carlo algorithms. In predicting heading dates of untested lines in untested environments, the proposed and two-step methods tended to provide more accurate predictions than the conventional genomic prediction methods, particularly in environments where phenotypes from environments similar to the target environment were unavailable for training genomic prediction. The proposed method showed greater accuracy in prediction than the two-step methods in all cross-validation schemes tested, suggesting the potential of the integrated approach in the prediction of phenotypes of plants.

Communicated by M. J. Sillanpaa.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-016-2667-5) contains supplementary material, which is available to authorized users.

✉ Hiroyoshi Iwata
aiwata@mail.ecc.u-tokyo.ac.jp

[1] Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

[2] Faculty of Agriculture, Kyushu University, Fukuoka 812-8581, Japan

[3] National Agriculture and Food Research Organization Agricultural Research Center, Tsukuba, Ibaraki 305-8666, Japan

[4] National Institute for Agro-Environmental Sciences, Tsukuba, Ibaraki 305-8604, Japan

Abbreviations

BIL	Backcross inbred line
C-Bay	Crop (DVR) model based on an MCMC algorithm
C-Nel	Crop (DVR) model based on a Nelder–Mead algorithm
DTH	Days to heading
DVR	Developmental rate model for rice heading date prediction
DVS	Developmental stage
EBL	Extended Bayesian LASSO

GBLUP	Genomic best linear unbiased prediction
IM	A hierarchical model integrating EBL with the crop (DVR) model
LOEO	Leave-one-environment-out cross-validation
LOELO	Leave-one-'combination of an environment and a line'-out cross-validation
LOLO	Leave-one-line-out cross-validation
MCMC	Markov chain Monte Carlo
QTL	Quantitative trait locus
RMSE	Root mean squared errors
T-Bay	Two-step approach based on C-Bay and EBL
T-Nel	Two-step approach based on C-Nel and EBL

Introduction

Accurate prediction of phenotypes of plants is important for breeding and management because it facilitates the design and selection of new genotypes, the introduction of new genetic resources and decision-making in management practices. It is desirable for a number of traits to be predicted in cultivation, including yield, phenology, stress resistance and morphological architectures. These traits are often complex and influenced by a number of factors such as interactions among genes, environments and management methods. A key statistical technique recently developed for the prediction of complex traits is genomic prediction or genome-wide prediction. Genomic prediction uses genome-wide DNA markers simultaneously as predictor variables and plays an essential role in genomic selection. Although this approach was proposed for predicting genetic merit in animal breeding (Meuwissen et al. 2001), it has been suggested that accurate genomic prediction is of practical use for phenotypic prediction in plants such as crops (Crossa et al. 2010), trees (Resende et al. 2012) and fruits (Iwata et al. 2013). However, a drawback of genomic prediction is that it is often difficult to predict the performance of plants in diverse environments (Resende et al. 2012; Ly et al. 2013), because phenotypes of complex traits are often influenced by genotype–environment interaction. As climate change and variability are serious concerns for crop and livestock production (Thornton et al. 2014), statistical models that enable prediction of the performance of plants cultivated in various environmental conditions will play a crucial role in breeding for environmental adaptability and optimization of management methods.

Another approach for predicting phenotypes of plants is a mathematical modelling approach known as crop modelling or eco-physiological modelling (Yin et al. 2004; Hammer et al. 2006). Crop models generally consist of multiple functions that depict physiological processes (Tardieu 2003). Using environment and management information, including temperature, photoperiod, precipitation and sowing date, the models predict the responses of plants/organs to environment stimuli (Tardieu 2003). Because it is recognized that the parameters of crop models are genetically controlled, parameters are often referred to as genetic coefficients or model input traits (Yin et al. 2003). Thus, genetic dissection of crop model parameters has been conducted to reveal the genetic control of responses to the environment. Models include those for growth and flowering time in barley (Yin et al. 2000, 2005), leaf growth in maize (Reymond et al. 2003), fruit quality in peach (Quilot et al. 2005), flowering time in rice (Nakagawa et al. 2005) and wheat (Bogard et al. 2014), yield under drought stress in rice (Gu et al. 2014) and flowering time in *Brassica oleracea* (Uptmoor et al. 2008). Based on estimated QTL effects, phenotypes of untested lines in untested environments can be predicted via the prediction of crop model parameter values. For example, Gu et al. (2014) predicted yields in rice and reported that prediction accuracy, measured as the coefficient of determination between the observed and predicted values, was 0.20–0.21. Reymond et al. (2003) predicted leaf elongation rates in maize and reported that the coefficient of determination was 0.74. Days to heading (DTH) in wheat was predicted by Bogard et al. (2014) with a root mean squared error (RMSE) between the observed and predicted of 6.3 days. These studies used only significant QTLs (markers) to predict the parameter values of crop models. Because detected QTLs usually capture only a part of the genetic variance and QTL effect sizes tend to be overestimated, using whole-genome markers (i.e. genomic prediction) is preferable for prediction.

A straightforward way to use genomic prediction with crop models is to predict the model input traits (parameters of crop models) using genomic prediction methods. This approach (the two-step approach) is similar to the studies based on detected QTLs described above. In this approach, the model input traits are first statistically optimized. Then the genomic prediction methods are trained with the optimized model input traits. A major advantage of this approach will be the feasibility of statistical implementation. However, when the model input traits are statistically optimized, an issue for this approach will be that uncertainty in the optimization of model input traits is not taken into account in building genomic prediction models because these processes (optimization of model input traits and training of genomic prediction) are performed separately. Another issue is that, because crop model parameters are often optimized for each line/cultivar independently (Nakagawa et al. 2005; Bogard et al. 2014), genetic relatedness among lines/cultivars cannot be taken into account in the optimization. A possible way of addressing these issues would be to adopt an approach more comprehensive than the two-step approach, i.e. a single hierarchical model that combines a crop model with a genomic prediction model.

With this approach (the integrated approach), the issues of the two-step approach can be avoided through the simultaneous inference of crop model parameters and marker effects and consideration of genetic relatedness among lines/cultivars via whole-genome markers. Thus, the predictions of the integrated approach would be expected to be more accurate and robust than those of the two-step approach.

Technow et al. (2015) proposed a statistical approach to combine genomic prediction with crop modelling for predicting the yield of maize. Using genome-wide markers, the authors defined prior distributions of four measurable traits that are inputs for a maize growth model and predicted the input traits and grain yield using an approximated Bayesian computation technique. During simulation analysis, the authors showed that based on environmental information, the integrated approach predicted grain yield more accurately than genomic prediction. However, to date, there has been no comparison between the integrated and two-step approaches; therefore, it has not been revealed whether the integrated approach, which is statistically more complex than the two-step approach is superior to the two-step approach pertaining to their prediction accuracy. Moreover, the integrated approach has not been applied to real data.

To achieve the integration of genomic selection with crop modelling, in this study, we developed an integrated approach to predict the heading date of rice (*Oryza sativa*, L.) and compared the predictive ability of the integrated approach with that of the two-step approaches using real data. Modelling of phenological events including the heading date (i.e. flowering time) and maturity time is important when simulating crop growth, and is thus an indispensable part of a complete crop model (Soltani and Sinclair 2012). We also compared the integrated approach with conventional genomic prediction and phenological (crop) modelling. As a phenological model for the trait, we used a developmental rate (DVR) phenological model (Yin et al. 1997; Nakagawa et al. 2005). We compared four approaches for predicting the heading date. The first approach is conventional genomic prediction. In this approach, heading date was directly predicted using whole-genome markers with the extended Bayesian LASSO (EBL) (Mutshinda and Sillanpaa 2010). The second is conventional phenological modelling. In this approach, heading date was predicted based on daily mean air temperature and photoperiod using the DVR model. Parameters of the DVR model were inferred in two ways using Nelder–Mead (i.e. non-Bayesian) and Markov chain Monte Carlo (MCMC) (i.e. Bayesian) algorithms. The third approach is the two-step approach. In this, the DVR model parameters were first inferred by the Nelder–Mead or MCMC algorithms, and then the genomic prediction model for the parameters was built using the EBL. The fourth approach is the integrated method, wherein the DVR model was combined with a genomic prediction model in a Bayesian two-level hierarchical model. The first level was the DVR model and the second level was the EBL, which inferred whole-genome marker effects from the DVR model parameters. The predictive performance of these methods was evaluated via field experiments using 174 backcross inbred lines (BILs) derived from japonica and indica varieties, Koshihikari and Kasalath, respectively. These lines were cultivated across 3 years in six locations that represented a wide range of day lengths.

Materials and methods

Plant materials and field evaluation

We used 174 Koshihikari/Kasalath//Koshihikari BILs together with the parental lines (Koshihikari and Kasalath). These lines were developed by the National Institute of Agrobiological Sciences Rice Genome Resource Center (Ma et al. 2002). Genotypes and a linkage map of 162 restriction fragment length polymorphism markers are available at http://www.rgrc.dna.affrc.go.jp/ineKKBIL182.html. All these markers were bi-allelic. The BILs were cultivated in six experimental fields across 3 years, resulting in field trials in a total of nine environments (Table S1). At each site, seeds were pre-germinated in water and sown to seedling trays filled with soil. Emergence was recorded for each genotype when about 80 % of plants emerged from the soil surface. Seedlings were grown either in a puddled field or in a poly-house, depending on the environment, and air temperature was measured in the nursery environment. Nursery duration varied from 15 days to 25 days depending on the environments (Table S1), but we commonly transplanted seedlings with 3–4 leaves. Each environment included two field replicates, in each of which 10 seedlings of each BIL were transplanted in a single row, with 30-cm spacing between the rows. The distance between the plants in a row was 15 cm. All fields received an equal amount of fertilizer at a rate of 50 kg/ha of N, 50 kg/ha of P_2O_5 and 50 kg/ha of K_2O, applied prior to transplanting. The field was kept flooded at least until flowering for all lines. Heading dates were recorded for five central plants for each line and were averaged. Although two field replications were available at each environment, missing records were occasionally observed. Averages of two replications and records from one replication had different proportions of environmental (non-genetic) variance. Thus, for each environment, we used the phenotypic records (i.e. DTH) of the replication with the fewest missing records. Air temperatures were measured onsite and the daily photoperiod (P_d)

was calculated astronomically, as per Horie and Nakagawa (1990), from:

$$\cos\left(\varpi P_d/2\right) = -\tan\phi\tan\delta,$$

where ω is the angular rate of daily rotation, ϕ represents the latitude of the observation point and δ is the declination of the sun.

Prediction methods

The prediction methods used were categorized into four groups (Table 1). Genomic prediction based on EBL (the GP method) used DTH and whole-genome marker genotypes as response and predictor variables, respectively. The phenological (DVR) models based on the Nelder–Mead and MCMC algorithms (C-Nel and C-Bay, respectively) predicted DTH based on mean air temperature and photoperiod in the growth period of the rice plants. The two-step methods were the combinations of C-Nel and C-Bay with EBL (T-Nel and T-Bay, respectively). In the EBL regression of the two-step methods, the DVR model parameters derived from the respective algorithms and the marker genotypes were considered to be the response and predictor variables, respectively. In the integrated model method (IM), the DVR model parameters were regressed on genome-wide markers using EBL and the model parameters and marker effects on them were inferred simultaneously. In the Bayesian methods (C-Bay, T-Bay and IM), the DVR model parameters were log-transformed. In T-Bay and IM, regression analysis was also conducted in the log-transformed scale. The prediction methods are explained in the subsequent sections.

EBL

The EBL was developed from the Bayesian LASSO (Mutshinda and Sillanpaa 2010). Whereas the Bayesian LASSO controls the level of shrinkage of predictor variables with a single hyperparameter (Park and Casella 2008), the EBL controls the level with two hyperparameters: one at the global level, the other at the marker-specific level (Mutshinda and Sillanpaa 2010). EBL has been shown to be robust across factors, e.g. the number of QTLs or heritability, which influence the prediction accuracy of genomic prediction (Onogi et al. 2015). We initially tried the genomic best linear unbiased prediction (GBLUP) method, which is more commonly used for genomic prediction. However, because EBL proved more accurate than GBLUP for the genomic prediction of DTHs using the rice population of the present study (data not shown), we focused on EBL. Because the primary purpose of this study was to compare the integrated approach with the two-step approach and genomic prediction, it was important to use the same regression method in all the approaches compared (IM, T-Nel, T-Bay and GP). We had not intended to compare the different regression methods (EBL vs. GBLUP) in this study.

We used the variational Bayesian algorithm introduced by Li and Sillanpaa (2012) and modified by Onogi et al.

Table 1 Prediction methods used in this study

Method[a]				Input[b]		Cross-validation[c]		
Type	Name	Description	Log	*E*	*G*	LOEO	LOLO	LOELO
Genomic prediction	GP	Extended Bayesian LASSO (EBL)			✓		✓[d]	✓[e]
Phenological (crop) model	C-Nel	DVR model with Nelder–Mead optimization		✓		✓		
	C-Bay	DVR model with MCMC	✓	✓		✓		
Two-step approach	T-Nel	C-Nel → EBL		✓	✓	✓	✓	✓
	T-Bay	C-Bay → EBL	✓	✓	✓	✓	✓	✓
Integrated model	IM	C-Bay + EBL	✓	✓	✓	✓	✓	✓

[a] GP is a regression analysis that uses days to heading and genome-wide markers as a response and predictor variables, respectively. We used the extended Bayesian LASSO (EBL) as the regression method. C-Nel is based on the developmental rate (DVR) model, which predicts heading dates on the basis of only environmental information and optimizes the model parameters with a Nelder–Mead algorithm. C-Bay is also based on the DVR model and infers the parameters in a Bayesian framework. T-Nel and T-Bay perform regression analysis using the DVR model parameters estimated by C-Nel and C-Bay, respectively, as response variables, and using the marker genotypes as predictor variables. We used the EBL for regression. IM integrates C-Bay with the EBL in a Bayesian framework and infers the DVR model parameters and whole-genome marker effects on the parameters simultaneously. Log indicates that the parameters of the DVR model were log-transformed and inferred

[b] Input information used for prediction. *E* and *G* indicate environmental and genomic information, respectively

[c] Cross-validation schemes used for evaluation of the prediction methods; LOEO, leave-one-environment-out cross-validation; LOLO, leave-one-line-out cross-validation; LOELO, leave-one-'combination of an environment and a line'-out cross-validation

[d] LOLO was conducted in each environment separately

[e] LOELO was conducted in the environment closest to the target (untested) environment with regard to the root mean squared difference in days to heading, which is presented in Table S2

(2015) for parameter estimation. The algorithm was executed using an R package, VIGoR (Onogi 2015). We provide the model structure of EBL and the calculation procedure performed by the package in the Supplementary Methods. The EBL was used in GP, T-Nel, T-Bay and IM. In each method, the predictor variables were marker genotypes and regression coefficients were marker effects. In GP, the response variable was DTH. In T-Nel and T-Bay, the response variables were the DVR model parameters derived from C-Nel and C-Bay, respectively. In IM, the response variables were the DVR model parameters derived simultaneously with regression coefficients for predictor variables (i.e. marker effects on the DVR model parameters). In these methods (GP, T-Nel, T-Bay and IM), the response variable was standardized and the hyperparameters were set to the values common across the methods (see Supplementary Methods).

DVR model

We used the DVR model developed by Nakagawa et al. (2005) which was originally proposed by Yin et al. (1997) as a 'three-stage Beta model'. The model assumes three developmental phases before flowering: (1) the 'juvenile phase', when plants are insensitive to photo stimulus; (2) the 'photosensitive phase', when photo stimulus is influential on DVR; and (3) the 'post-photosensitive phase' after the photosensitive phase for flowering. DVRs accumulate from the emergence day, and growth of developmental stages (DVSs) is given by:

$$\mathrm{DVS} = \sum_{d=1}^{D} \mathrm{DVR}_d,$$

where D is the number of days since the emergence day. Flowering occurs when DVS reaches 1. The DVR at day d is defined as follows:

$$\mathrm{DVR}_d = \begin{cases} f(T_d)/G & \text{if } \mathrm{DVS}_d < \mathrm{DVS}_1 \text{ or } \mathrm{DVS}_d > \mathrm{DVS}_2 \\ f(T_d)g(P_d)/G & \text{if } \mathrm{DVS}_1 \le \mathrm{DVS}_d \le \mathrm{DVS}_2 \end{cases},$$

where T_d is the mean air temperature (°C) at day d, P_d is the photoperiod (h), f and g denote the functions that relate T_d and P_d with DVR_d, respectively, G (≥ 0) represents earliness of the heading date given optimal temperature and photoperiod, and DVS_1 and DVS_2 denote the ends of the juvenile and photosensitive phases, respectively. The functions f and g are defined as follows:

$$f(T_d) = \begin{cases} \left[\left(\frac{T_d - T_b}{T_o - T_b}\right)\left(\frac{T_c - T_d}{T_c - T_o}\right)^{(T_c - T_o)/(T_o - T_b)}\right]^{\alpha} & \text{if } T_b \le T_d \le T_c \\ 0 & \text{otherwise} \end{cases},$$

and

$$g(P_d) = \begin{cases} \left[\left(\frac{P_d - P_b}{P_o - P_b}\right)\left(\frac{P_c - P_d}{P_c - P_o}\right)^{(P_c - P_o)/(P_o - P_b)}\right]^{\beta} & \text{if } P_o \le P_d \\ 1 & \text{if } P_o > P_d \end{cases},$$

where T_b, T_o and T_c denote the base, optimum and ceiling temperatures, respectively, and P_b, P_o and P_c denote the base, optimum and ceiling photoperiod, respectively. According to Nakagawa et al. (2005), T_b, T_o, T_c, P_b, P_o and P_c were fixed to be 8, 30, 42 °C, 0, 10 and 24 h, respectively. Thus, G represents DTH under a 10-h photoperiod at 30 °C in the present study as well as in Nakagawa et al. (2005). Parameters α ($\ge$0) and β ($\ge$0) represent the temperature and photoperiod sensitivity, respectively. Sensitivity increases as these parameters increase. The change of sensitivity due to these parameters is illustrated in Fig. 1 in Nakagawa et al. (2005). DVS_1 and DVS_2 are defined as

$$\mathrm{DVS}_1 = 0.145 + 0.005G$$
$$\mathrm{DVS}_2 = 0.345 + 0.005G$$

according to Nakagawa et al. (2005). Although some prefixed parameters such as T_o may show genetic variation, we assumed genetic variations existed only in G, α and β in this study. Thus, we use G_i, α_i and β_i to denote the DVR model parameters of line i. Note that all these parameters are constrained to be $\ge$0.

Nelder–Mead optimization of the DVR model

The objective function is

$$\mathrm{argmin}_{G_i,\alpha_i,\beta_i} \sum_{j=1} \left| H_{ij} - h\left(G_i, \alpha_i, \beta_i, \mathbf{T}_j, \mathbf{P}_j\right) \right|,$$

where H_{ij} is the DTH of line i in environment j, h is a function that calculates the DTH for given values of G_i, α_i and β_i and mean daily temperature ($\mathbf{T}_j$) and photoperiod ($\mathbf{P}_j$) in environment j. | | indicates absolute difference. Although we also tried optimization based on squared loss in the preliminary study, prediction accuracy was lower than that based on absolute loss (data not shown). This may indicate the presence of outliers, because the absolute loss function is more robust than the squared loss function in handling outliers. However, we did not observe any obvious outliers in the distribution of DTH (Fig. S1) or in the local outlier factor analysis (Breunig et al. 2000, data not shown). Optimization was conducted for each line. As initial values for the optimization of G, we tried 40, 55 and 80. For α and β, we tried 0.01, 5 and 10. These values were chosen to cover the ranges of estimates in the BIL population derived from Nipponbare and Kasalath (Nakagawa et al. 2005). Consequently, we conducted optimization a total of 27 (3 × 3 × 3) times with different initial value combinations and obtained 27 sets of estimates of G, α and β. From these, we chose the set that minimized the objective

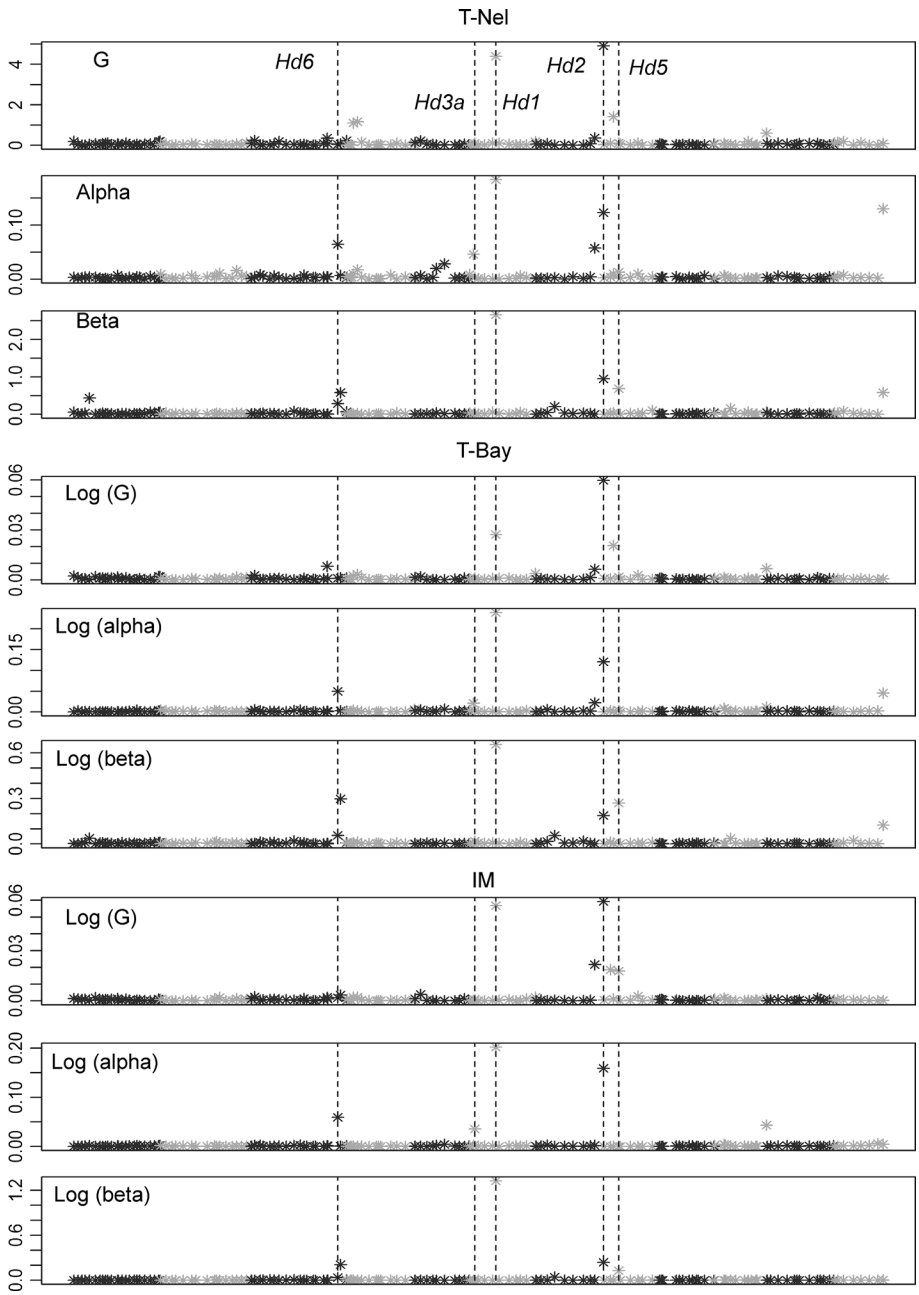

Fig. 1 Marker effects on the developmental rate model parameters derived from the two-step approaches (T-Nel and T-Bay) and the integrated model (IM). *Different shades* indicate different chromosomes (*left to right*, chromosome 1 to chromosome 12). *Dashed lines* indicate the chromosomal positions of known major genes illustrated in the figure. Because T-Bay and IM estimated the marker effects in log-transformed scales, scales of *y axes* differ from those of T-Nel for all the parameters. G, α and β represent earliness of heading date under optimal conditions, temperature sensitivity and photoperiod sensitivity, respectively

function. When multiple sets of estimates minimized the objective function, we calculated the average of the estimates over the sets. Optimization was conducted by using the optim function in R (R Development Core Team 2011).

MCMC inference of the DVR model

In C-Bay, the DVR model was regarded as a regression model where the response variable was fixed to be 1:

$$1 = \sum_{d=1}^{H_{ij}} \mathrm{DVR}_{ij,d} + e_{ij}, \tag{1}$$

where $\mathrm{DVR}_{ij,d}$ is the DVR of line i in environment j at day d and e_{ij} is the residual. The residual was assumed to follow a normal distribution, $N\left(0,\ \sigma_e^2\right)$ where $\sigma_e^2 \sim 1/\sigma_e^2$. Log-transformed values of G_i, α_i and β_i were assumed to follow normal distributions:

$$\tilde{G}_i = \log(G_i) \sim \mathrm{N}\left(\mu_G,\ \sigma_G^2\right)$$

$$\tilde{\alpha}_i = \log(\alpha_i) \sim \mathrm{N}\left(\mu_\alpha,\ \sigma_\alpha^2\right),$$

$$\tilde{\beta}_i = \log(\beta_i) \sim \mathrm{N}\left(\mu_\beta,\ \sigma_\beta^2\right),$$

where μ_G, μ_α and μ_β are the means, and σ_G^2, σ_α^2 and σ_β^2 are the variances. Non-informative prior distributions were assigned to the means and variances. μ_G, μ_α and μ_β were assumed to be proportional to a constant value. σ_G^2, σ_α^2 and σ_β^2 were assumed to follow scaled inverse Chi-squared distributions with the degree of freedom being -2 and the scaling parameter being 0. Parameter estimation was conducted by using a Metropolis algorithm (Metropolis et al. 1953) (see Supplementary Methods). The number of iterations was 60,000, and the sampling interval was 50. The first 200 samples were discarded as burn-in, and the posterior distributions were inferred from 1000 samples. T-Bay used the posterior means of $\tilde{G}$, $\tilde{\alpha}$ and $\tilde{\beta}$.

Integrated model

Together with C-Bay, the DVR model in IM was regarded as a regression model wherein the response variable was fixed to be 1 (Eq. 1). The residual was assumed to follow a normal distribution $N\left(0,\ \sigma_e^2\right)$, where $\sigma_e^2 \sim 1/\sigma_e^2$. $\tilde{G}$, $\tilde{\alpha}$ and $\tilde{\beta}$ were regressed on marker genotypes. For line i, we assume

$$\tilde{G}_i \sim \mathrm{N}\left(\mu_G + \sum_{p=1}^{P} g_p x_{ip},\ 1/\tau_{0,G}^2\right)$$

$$\tilde{\alpha}_i \sim \mathrm{N}\left(\mu_\alpha + \sum_{p=1}^{P} a_p x_{ip},\ 1/\tau_{0,\alpha}^2\right)$$

$$\tilde{\beta}_i \sim \mathrm{N}\left(\mu_\beta + \sum_{p=1}^{P} b_p x_{ip},\ 1/\tau_{0,\beta}^2\right),$$

where g_p, a_p and b_p are effects of marker p, and x_{ip} is the genotype at marker p coded as -1 (AA), 0 (AB) and 1 (BB). IM infers the DVR model parameters ($\tilde{G}$, $\tilde{\alpha}$ and $\tilde{\beta}$) and the parameters included in these regressions simultaneously. The intercepts (μ_G, μ_α and μ_β) were assumed to be proportional to a constant value. As described above, we used the EBL to model marker effects, g_p, a_p and b_p. The prior distributions of the residual precisions, $\tau_{0,G}^2$,$\tau_{0,\alpha}^2$ and $\tau_{0,\beta}^2$ were $\tau_{0,G}^2 \sim 1/\tau_{0,G}^2$, $\tau_{0,\alpha}^2 \sim 1/\tau_{0,\alpha}^2$ and $\tau_{0,\beta}^2 \sim 1/\tau_{0,\beta}^2$, respectively. To infer these parameters, we developed a variational Bayesian method using MCMC simulation to derive the expectations of the DVR model parameters (see Supplementary Methods). In this method, the means and variances of the DVR model parameters (e.g. $\mathrm{E}\left[\tilde{G}_i\right]$ or $\mathrm{V}\left[\tilde{G}_i\right]$) that are required for updating the EBL parameters (e.g. g_p) are obtained via MCMC sampling. Because the variational Bayesian inference is related to the EM algorithm (Ghahramani and Beal 2001; Bishop 2006), our algorithm for IM can be seen as analogous to the Monte Carlo EM algorithm (Wei and Tanner 1990). MCMC samples of $\tilde{G}$, $\tilde{\alpha}$ and $\tilde{\beta}$ at the last iteration of the combined procedure were used to derive the posterior distributions of DTH. The number of MCMC samples at the last iteration was 300.

Full data analysis

To compare the behaviours of the prediction methods, analysis was conducted using full data: GP was fitted to each environment and C-Nel was fitted to each line separately. T-Nel re-estimated the DVR model parameters by using the EBL from the estimates from C-Nel. C-Bay and IM were fitted to data comprising all lines in all environments. T-Bay re-estimated the DVR model parameters by using the EBL from the estimates by C-Bay.

Leave-one-out cross-validation

The predictive ability of the prediction methods was evaluated by using three types of leave-one-out cross-validation (LOO). The first is leave-one-environment-out cross-validation (LOEO). In this scheme, one of nine environments was removed from the data, and the DTHs of all lines in the removed environment were predicted with the prediction model derived from the DTHs of all lines of the remaining eight environments. This procedure was repeated until all environments were removed once. When T-Nel and T-Bay were tested by LOEO, the values of G, α and β estimated by C-Nel and C-Bay, respectively, were used as the response variables for regression. The fitted values, i.e. the values without residuals, were used for prediction. The second LOO is leave-one-line-out cross-validation (LOLO). One of 176 BILs was removed from the data, and the DTHs of the removed line in all environments were predicted with the prediction model derived from the retained data. This was repeated until all BILs were removed once. The third type is leave-one-'combination of an environment and a line'-out cross-validation (LOELO). In this scheme, one of nine environments and one of 176 BILs were removed from the data, and the DTH of the removed line at the removed environment was predicted with the prediction model derived from the data consisting of 175 BILs in eight environments. This was repeated until all combinations of environments and lines were removed once. LOEO, LOLO and LOELO correspond, respectively, to tests of predictive ability for tested lines in untested environments, untested lines in tested environments and untested lines in untested environments.

The phenological models C-Nel and C-Bay were evaluated using only LOEO because they were not able to predict new lines (Table 1). Conversely, GP was not evaluated using LOEO because GP could not use the environment information (Table 1). LOLO of GP was conducted at each environment separately. To assess the ability of GP for untested lines in untested environments, LOLO was conducted at the

environment most similar to the target (untested) environment. We used the root mean squared difference of DTH between environments as the metric of between-environment similarity. We report the results of this 'LOLO at the most similar environment' together with the results of LOELO for the two-step and integrated approaches, because both LOO schemes evaluated the predictive ability for untested lines in untested environments. Because we used DTH of the target environments to determine the most similar environments, the prediction ability of GP may have been biassed towards an overestimate. However, because GP showed the least accuracy, as described later in the Results section, such bias, if it existed, would not have influenced our conclusions. Another approach for constructing the GP training data set for prediction in untested environments is to average DTH across the tested (available) environments. Although we attempted this, the results were extremely poor (data not shown).

In each LOO scheme, predicted DTH values were obtained for individual lines in each environment. Predictive ability was measured as RMSE and the Pearson correlation between the predicted and observed DTH. Unlike other traits, including yield, DTH needs to be controlled at a preferable value rather than to be increased or decreased as far as possible; therefore, RMSE is more important in practice. Thus, we compared the methods by primarily using RMSE as the metric of accuracy, although we reported both metrics (RMSE and correlation). We also calculated the regression coefficients of the predicted values on observed values. Because coefficients lower than one indicate shrinkage of prediction, we reported the coefficients as the level of shrinkage.

Results

Phenotype evaluation

The distribution of DTH in each environment is shown in Fig. S1. As expected, DTH tended to be longer at higher latitudes. Correlation and root mean squared difference of DTH between environments are presented in Table S2. Scatterplots of DTH between environments are shown in Fig. S2. As described in the "Materials and methods", on the basis of the root mean squared difference, we determined the environments used for training of GP to test predictive ability of DTH for untested lines in untested environments: for prediction at Tsukuba2007, Tsukuba2009 was used as the environment for training; for Tsukuba2008E, Ishikawa; for Tsukuba2008L, Ishigaki; for Tsukuba2009, Ishikawa; for Ishikawa, Tsukuba2009; for Fukuoka, Tsukuba2008L; for Ishigaki, Ha Noi, for Thai Nguyen, Ishigaki and for Ha Noi, Ishigaki were used as the environments for training.

Full data analysis

RMSEs between the observed and estimated DTH were the smallest in C-Nel and the largest in T-Bay (Table 2). The residual variance σ_e^2 inferred by C-Bay and IM were quite small, suggesting good fitting of the models.

The Markov chains of C-Bay converged quickly (data not shown). Acceptance rates of Metropolis algorithms for the estimates of $\tilde{G}$, $\tilde{\alpha}$ and $\tilde{\beta}$ were 0.26, 0.83 and 0.67, respectively, suggesting good mixing properties. To assess the reproducibility of IM, we repeated runs five times with different initial values of G, α and β. The estimates (posterior means) of $\tilde{G}$, $\tilde{\alpha}$ and $\tilde{\beta}$ were highly consistent among the replications: mean Pearson correlation coefficients among the replications were >0.99 for all the parameters. Estimated marker effects on these parameters were also consistent among replications (mean Pearson correlation coefficients >0.98).

The marker effects inferred by T-Nel, T-Bay and IM are shown in Fig. 1. Effects estimated by these methods were largely consistent between the methods, and association signals were often observed around known heading date genes, *Hd1* (Yano et al. 1997, 2000), *Hd2* (Yano et al. 1997), *Hd3a* (Yano et al. 1997; Monna et al. 2002), *Hd5* (Yano et al. 1997) and *Hd6* (Takahashi et al. 2001) (Fig. 1).

Table 2 Root mean squared errors (RMSEs) (day) and Pearson correlation between estimated and observed values, regression coefficients of estimated values on observed values (slope), and posterior means of the residual variance of the Bayesian methods in the full data analysis

	GP[a]	Phenological (crop) model		Two-step		IM
		C-Nel[b]	C-Bay	T-Nel[c]	T-Bay[c]	
RMSE	5.6	3.6	4.1	6.4	8.7	4.0
Correlation	0.96	0.99	0.98	0.95	0.93	0.98
Slope	0.91	0.96	0.95	0.87	0.75	0.95
σ_e^2 (10^{-3})			2.39			2.82

[a] GP was fitted to data collected in each environment separately

[b] C-Nel was fitted to data collected for each line separately

[c] T-Nel and T-Bay re-estimated the DVR model parameters by fitting the extended Bayesian LASSO to the estimates by C-Nel and C-Bay, respectively, which were pooled across all the lines

Table 3 Root mean squared errors (RMSEs) (day) and Pearson correlation between predicted and observed values and regression coefficients of predicted values on observed values (slope) in the three leave-one out cross-validation schemes

	GP	Phenological (crop) model		Two-step		IM
		C-Nel	C-Bay	T-Nel	T-Bay	
LOEO						
RMSE		6.2	5.2	7.5	9.5	5.1
Correlation		0.96	0.97	0.94	0.92	0.97
Slope		0.89	0.92	0.82	0.72	0.93
LOLO						
RMSE	7.2			7.7	9.4	7.4
Correlation	0.94			0.93	0.91	0.94
Slope	0.89			0.84	0.74	0.87
LOELO						
RMSE	10.8			8.6	10.2	8.0
Correlation	0.87			0.92	0.90	0.93
Slope	0.75			0.79	0.71	0.85

RMSE, correlation and slope were calculated over all the environments

LOEO leave-one-environment-out cross-validation, *LOLO* leave-one-line-out cross-validation, *LOELO* leave-one-'combination of an environment and a line'-out cross-validation

A major difference between the methods is that a relatively large signal was detected in chromosome 12 for T-Nel and T-Bay. The associated marker was *R1684*.

Leave-one-out cross-validation

Results for LOEO are presented in Table 3. RMSE and Pearson correlation calculated for each environment are presented in Table S3. The least RMSE was observed in IM. IM also showed the regression coefficient (slope) that was closest to one, suggesting less shrunken prediction. In LOLO, GP showed the least RMSE and IM showed the second least (Tables 3 and S4). The slope of GP was also closest to one and that of IM was secondly closest. Thus, it is suggested that GP is suitable for heading date prediction of untested lines when the training data set at the target environment is available. In LOELO, IM showed the least RMSE with the least shrinkage tendency (Tables 3 and S5), suggesting the usefulness of the integrated approach in prediction of untested lines in untested environments. The RMSE of GP was greater than the RMSEs of the two-step and integrated approaches. Considering that GP showed the least RMSE in LOLO, the results suggest that genomic prediction of heading dates is sensitive to the differences in environmental conditions between the training and target environments.

In each LOO experiment, IM consistently showed higher accuracy than the two-step approach, regardless of the optimization algorithms for the DVR model, Nelder–Mead and MCMC algorithms (Table 3). This suggests that prediction accuracy can be improved by integrating genomic prediction with crop modelling in a single hierarchical model.

IM can estimate the posterior variance of predicted DTH. To investigate the empirical reliability of the posterior variance, we calculated the frequency with which the observed (true) DTH was included in the credible intervals of the predicted DTH (Table 4). In each LOO scheme, posterior variance is suggested to have been underestimated. This underestimation was most severe in LOEO. However, in LOLO and LOELO, the underestimation was relaxed and the posterior variances would have reasonable sizes.

Discussion

We developed a novel method that integrated whole-genome markers with a phenological model in predicting the heading date of Asian rice, and compared this method with conventional genomic prediction, phenological modelling and two-step approach. The proposed method IM showed the best accuracy in LOEO and LOELO and the second best in LOLO. In each LOO scheme, IM was superior to the two-step approach (T-Nel and C-Nel) pertaining to both prediction accuracy and the amount of shrinkage. An approach that integrates genomic prediction with crop models will be statistically more challenging than the two-step approach and will require the development of special algorithms and programs (e.g. the variational Bayesian inference combined with MCMC in the present study). Thus, it is important to demonstrate the superiority of the integrated approach over the two-step approach to support the development of the integrated approach. Our study has several limitations: the backcross inbred line population used here is not usually used for the development of

Table 4 Frequencies for the true DTH being included in the credible intervals estimated by the integrated model in each leave-one-out cross-validation (LOO) scheme

LOO scheme[a]	Average of posterior SD[b]	Credible interval				
		95 %	90 %	80 %	70 %	60 %
LOEO	1.6 (±0.5)	0.42	0.35	0.26	0.19	0.14
LOLO	6.0 (±2.5)	0.86	0.78	0.69	0.58	0.49
LOELO	6.1 (±2.5)	0.84	0.76	0.65	0.55	0.46

In each LOO scheme, prediction was made for each line in each environment. The frequencies that the true DTH was included in the corresponding credible intervals of predicted DTH were calculated

[a] *LOEO* leave-one-environment-out cross-validation, *LOLO* leave-one-line-out cross-validation, *LOELO* leave-one-'combination of an environment and a line'-out cross-validation

[b] Average (SD) of posterior standard deviation of predicted DTH

new cultivars; heading date in rice is considered to have a relatively simple genetic architecture (Matsubara et al. 2014) and the phenological model used in this study, the DVR model, is relatively simple, consisting of only a few functions. Nevertheless, for the first time, our results demonstrate superiority of the integrated approach in predicting plant phenotypes using real data. As the first empirical evidence, this study should encourage further development and application of the integrated approach in predicting plant phenotypes.

An issue in our integrated approach (IM) is the computational time required. When Nelder–Mead optimization is conducted for each line in tandem with the full data analysis, the computational times of IM and T-Nel are similar. However, the Nelder–Mead optimization can be conducted in parallel, which significantly reduces the time required. The computational burden of IM mainly stemmed from the MCMC simulation used to infer the posterior distribution of the DVR model parameters. This can be reduced by decreasing the number of iterations of the MCMC (600) and/or the number of entire optimization steps (100) (see Supplementary Methods), though these reductions may prevent the solution from reaching the global optimum. The number of iterations (600 and 100) was set arbitrarily in this study, but when using IM in practical applications, it would be necessary to assess the effect of the number of iterations on the efficiency of optimization.

In all the cross-validation schemes, IM outperformed T-Nel and T-Bay in terms of prediction accuracy. This superiority can be attributed to several factors: consideration of genetic dependency between lines in deriving the DVR model parameters, consideration of uncertainty of the DVR model parameter derivation in marker effect estimation and avoiding loss of information by simultaneously deriving the DVR model parameters and the marker effects on these. The relative contributions of these factors to the improvement in accuracy are unclear. Note that in C-Bay (i.e. the first step of T-Bay), the derivation of the DVR model parameters was not independent between lines because the variances of the model parameters σ_G^2, σ_α^2 and σ_β^2 were estimated from the parameter values of all the lines (see Supplementary Methods). Thus, information of the DVR model parameters was shared among lines via the variances. Nevertheless, C-Bay and T-Bay were not always superior to C-Nel and T-Nel in the cross-validation schemes tested. Thus, sharing information between lines may not largely contribute in improving prediction accuracy of IM. Further work will be needed to elucidate the reasons for the superiority of the integrated approach.

Bogard et al. (2014) predicted heading dates in wheat using a two-step approach that was based on detected QTLs. The authors reported that, when untested lines were predicted at untested environments, the RMSE was 6.3 days. Although the reported RMSE was lower than those obtained by T-Nel, T-Bay and IM in the present study, prediction accuracy will be influenced by various factors other than models and inference methods; therefore, direct comparison of accuracy between the studies would not be meaningful. These factors include: the number of lines and environments used for training (210 lines and 10 environments in Bogard et al. (2014)), the predictive ability of phenological models, genetic relationships between lines and variation in environmental conditions.

IM is a two-level hierarchical model, where a regression model is embedded in each level: the first level is the DVR model and the second level is the genomic prediction models, i.e. multi-locus regression models, which regard the DVR model parameters as the response variables. This model structure is similar to the structures used for functional QTL mapping by Malosetti et al. (2006), Sillanpaa et al. (2012) and Li et al. (2014). These studies, respectively, fitted logistic, polynomial and linear curves to temporary trends and regressed the curve parameters on multi-locus genotypes in single hierarchical models. Comparison between a single hierarchical model and a two-step approach that derives the curve parameters and the marker effects on the parameters separately was conducted by Li et al. (2014) using growth-related traits of conifers. The authors reported that both the methods detected similar sets

of QTLs when the sample size was large (≥250); whereas, when the sample size was small (<100), more association signals were detected by using the two-step approach, although it was unclear whether these associations were true. In our results, the integrated (IM) and the two-step methods (T-Nel and T-Bay) detected similar sets of association signals in the full data analysis (Fig. 1). This consistent result across the methods could be because of the relatively large sample size (176) and relatively simple genetic architecture of heading dates in the used population. The fact that both the integrated and two-step approaches captured similar association signals may suggest that the superiority of IM in prediction accuracy was because of the accuracy of marker effect size estimation rather than QTL detection power. The hierarchical model by Technow et al. (2015) also has a two-level structure; its first level is the crop growth model and the second is the genomic prediction model. However, one difference is that, although the model input traits (y_{TLN}, y_{AM}, y_{SRE}, and y_{MTU} in Technow et al. 2015) are regressed on genome-wide markers in the second level, no residual term was assumed (see Eq. 2 in Technow et al. 2015). This means that the complete variance of the model input traits was assumed to be explained by the additive effects of the markers, which may be a strong assumption in real data analysis.

In LOELO, the RMSE of GP was greater than the RMSEs of the two-step and the integrated approaches, particularly at Tsukuba2007 and Tsukuba2008E (Table S5). For prediction for these two environments, GP was trained with records at Tsukuba2009 and Ishikawa, respectively, as described in the "Results" section. However, the root mean squared differences of DTH between the target and training environments (i.e. between Tsukuba2007 and Tsukuba2009 and between Tsukuba2008E and Ishikawa), were 10.3 and 17.9, respectively, which were the largest two values among the differences between the target and training environments. Thus, the results of LOELO clearly suggest that accuracy of genomic prediction depends on the similarity between the target and training environments, which is also illustrated by morphological traits in loblolly pine (Resende et al. 2012). On the other hand, the phenological model-based methods, IM, T-Nel and T-Bay did not appear to show the dependency on the similarity. GP can predict plant performance at multiple environments by modelling the covariance structure between genotypic (polygenic) effects between environments in a mixed model framework (Burgueno et al. 2012). However, this multi-environment mixed model is unable to predict phenotypes in environments where no line has been cultivated. To overcome this deficiency, Heslot et al. (2014) proposed mixed models that incorporate stress covariates derived from crop models to model genotype–environment interaction and to predict phenotypes in untested environments. The authors performed multiple and separated steps in parameter inference, which may impair the predictive ability. Their approach is, however, an alternative to integrating genomic prediction with crop models; it will be interesting to compare the predictive ability with that of the crop model-based methods.

Author contribution statement AO conceived and developed the integrated Bayesian method, conducted all statistical analyses and drafted the manuscript. MW developed the Nelder–Mead optimization procedure, and HN collaborated on this development. TM, T. Hasegawa and HN conducted the field trials and collaborated on the preparation of the manuscript. T. Hayashi collaborated on statistical analysis. T. Hasegawa and HI were involved in the conception and design of the study, administrative support and supervision. All authors have read and approved the final manuscript.

Acknowledgments The authors thank Seishi Ninomiya and Ryo Ohsawa for the contribution to the conception and design of this study. This study was supported by JSPS KAKENHI Grant Numbers 19208003 and 25252002 and by a Grant-in-Aid for JSPS Fellows (14J10661).

Data archiving The data files, scripts and results are available at https://github.com/Onogi/HeadingDatePrediction. The data files contain heading and emergence dates and environmental information (daily mean air temperature and photoperiod).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Bishop CM (2006) Pattern recognition and machine learning. Section 10.2.1 Variational distribution. New York: Springer

Bogard M, Ravel C, Paux E, Bordes J, Balfourier F, Chapman SC, Le GJ, Allard V (2014) Predictions of heading date in bread wheat (*Triticum aestivum* L.) using QTL-based parameters of an eco-physiological model. J Exp Bot 65:5849–5865

Breunig MM, Kriegel HP, Ng RT, Sander J (2000) LOF: identifying density-based local outliers. In: Chen W, Naughton JF, Bernstein PA (eds) Proceedings of the ACM SIGMOD International Conference on Management Data, ACM, pp 93–104

Burgueno J, de los Campos G, Weigel K, Crossa J (2012) Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Sci 52:707–719

Crossa J, Deloscampos G, Perez P, Gianola D, Burgueno J, Araus JL, Makumbi D, Singh RP, Dreisigacker S, Yan J, Arief V, Banziger M, Braun HJ (2010) Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186:713–724

Ghahramani Z, Beal MJ (2001) Propagation algorithms for variational Bayesian learning. In: Leen TK, Dietterich TG, Tresp V (eds)

Advances in neural information processing systems 13, MIT press, pp 507–513
Gu J, Yin X, Zhang C, Wang H, Struik PC (2014) Linking ecophysiological modelling with quantitative genetics to support marker-assisted crop design for improved yields of rice (*Oryza sativa*) under drought stress. Ann Bot 114:499–511
Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, van Eeuwijk F, Chapman S, Podlich D (2006) Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci 11:587–593
Heslot N, Akdemir D, Sorrells ME, Jannink JL (2014) Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor Appl Genet 127:463–480
Horie T, Nakagawa H (1990) Modelling and prediction of developmental process in rice. I. Structure and method of parameter estimation of a model for simulating developmental process toward heading. Jpn J Crop Sci 59:687–695
Iwata H, Hayashi T, Terakami S, Takada N, Sawamura Y, Yamamoto T (2013) Potential assessment of genome-wide association study and genomic selection in Japanese pear *Pyrus pyrifolia*. Breed Sci 63:125–140
Li Z, Sillanpaa MJ (2012) Estimation of quantitative trait locus effects with epistasis by variational Bayes algorithms. Genetics 190:231–249
Li Z, Hallingback HR, Abrahamsson S, Fries A, Gull BA, Sillanpaa MJ, Garcia-Gil MR (2014) Functional multi-locus QTL mapping of temporal trends in Scots Pine wood traits. G3 (Bethesda) 4:2365–2379
Ly D, Hamblin M, Rabbi I, Melaku G, Bakare M, Gauch HG, Okechukwu R, Dixon AGO, Kulakow P, Jannink JL (2013) Relatedness and genotype × environment interaction affect prediction accuracies in genomic selection: a study in Cassava. Crop Sci 53:1312–1325
Ma JF, Shen R, Zhao Z, Wissuwa M, Takeuchi Y, Ebitani T, Yano M (2002) Response of rice to Al stress and identification of quantitative trait Loci for Al tolerance. Plant Cell Physiol 43:652–659
Malosetti M, Visser RG, Celis-Gamboa C, van Eeuwijk FA (2006) QTL methodology for response curves on the basis of non-linear mixed models, with an illustration to senescence in potato. Theor Appl Genet 113:288–300
Matsubara K, Hori K, Ogiso-Tanaka E, Yano M (2014) Cloning of quantitative trait genes from rice reveals conservation and divergence of photoperiod flowering pathways in Arabidopsis and rice. Front Plant Sci 5:193
Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21:1087–1092
Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829
Monna L, Lin X, Kojima S, Sasaki T, Yano M (2002) Genetic dissection of a genomic region for a quantitative trait locus, *Hd3*, into two loci, *Hd3a* and *Hd3b*, controlling heading date in rice. Theor Appl Genet 104:772–778
Mutshinda CM, Sillanpaa MJ (2010) Extended Bayesian LASSO for multiple quantitative trait loci mapping and unobserved phenotype prediction. Genetics 186:1067–1075
Nakagawa H, Yamagishi J, Miyamoto N, Motoyama M, Yano M, Nemoto K (2005) Flowering response of rice to photoperiod and temperature: a QTL analysis using a phenological model. Theor Appl Genet 110:778–786
Onogi A (2015) Documents for VIGoR. https://github.com/Onogi/VIGoR
Onogi A, Ideta O, Inoshita Y, Ebana K, Yoshioka T, Yamasaki M, Iwata H (2015) Exploring the areas of applicability of whole-genome prediction methods for Asian rice (*Oryza sativa* L.). Theor Appl Genet 128:41–53
Park T, Casella G (2008) The Bayesian lasso. Amer Stat Assoc 103:681–686
Quilot B, Kervella J, Genard M, 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
R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org/
Resende MFJ, Munoz P, Acosta JJ, Peter GF, Davis JM, Grattapaglia D, Resende MD, Kirst M (2012) Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. New Phytol 193:617–624
Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F (2003) Combining quantitative trait Loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. Plant Physiol 131:664–675
Sillanpaa MJ, Pikkuhookana P, Abrahamsson S, Knurr T, Fries A, Lerceteau E, Waldmann P, Garcia-Gil MR (2012) Simultaneous estimation of multiple quantitative trait loci and growth curve parameters through hierarchical Bayesian modeling. Heredity (Edinb) 108:134–146
Soltani A, Sinclair TR (2012) Modeling physiology of crop development, growth and yield. Chapter 6 phenology–temperature. CABI, MA, USA
Takahashi Y, Shomura A, Sasaki T, Yano M (2001) *Hd6*, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the alpha subunit of protein kinase CK2. Proc Natl Acad Sci USA 98:7922–7927
Tardieu F (2003) Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. Trends Plant Sci 8:9–14
Technow F, Messina CD, Totir LR, 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
Thornton PK, Ericksen PJ, Herrero M, Challinor AJ (2014) Climate variability and vulnerability to climate change: a review. Glob Chang Biol 20:3313–3328
Uptmoor R, Schrag T, Stützel H, Esch E (2008) Crop model based QTL analysis across environments and QTL based estimation of time to floral induction and flowering in *Brassica oleracea*. Mol Breed 21:205–216
Wei GCG, Tanner MA (1990) A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. J Amer Statist Assoc. 85:699–704
Yano M, Harushima Y, Nagamura Y, Kurata N, Minobe Y, Sakaki T (1997) Identification of quantitative trait loci controlling heading date in rice using a high-density linkage map. Theor Appl Genet 95:1025–1032
Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. Plant Cell 12:2473–2484
Yin X, Kropff MJ, Horie T, Nakagawa H, Centeno HG, Zhu D, Goudriaan J (1997) A model for photothermal responses of flowering in rice I. Model description and parameterization. Field Crop Res 51:189–200
Yin X, Chasalow SD, Dourleijn CJ, Stam P, Kropff MJ (2000) Coupling estimated effects of QTLs for physiological traits to a crop growth model: predicting yield variation among recombinant inbred lines in barley. Heredity (Edinb) 85:539–549

Yin X, Stam P, Kropff MJ, Schapendonk AH (2003) Crop modeling, QTL mapping, and their complementary role in plant breeding. Agron J 95:90–98

Yin X, Struik PC, Kropff MJ (2004) Role of crop physiology in predicting gene-to-phenotype relationships. Trends Plant Sci 9:426–432

Yin X, Struik PC, van Eeuwijk FA, Stam P, Tang J (2005) QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley. J Exp Bot 56:967–976