

1 **Utilizing random regression models for genomic prediction of a longitudinal**  
2 **trait derived from high-throughput phenotyping**

3 Malachy Campbell<sup>1,2</sup>, Harkamal Walia<sup>1</sup>, and Gota Morota<sup>2</sup>

4 <sup>1</sup>Department of Agronomy and Horticulture, University of Nebraska Lincoln, Lincoln, NE,  
5 USA 68583

6 <sup>2</sup>Department of Animal Science, University of Nebraska Lincoln, Lincoln, NE, USA 68583

7     **Corresponding author:**

8     Malachy Campbell

9     Department of Animal Science, Department of Agronomy and Horticulture

10    University of Nebraska Lincoln

11    Lincoln, Nebraska 68583

12    Email: [campbell.malachy@gmail.com](mailto:campbell.malachy@gmail.com)

## 13 Abstract

14 The accessibility of high-throughput phenotyping platforms in both the greenhouse and field,  
15 as well as the relatively low cost of unmanned aerial vehicles, have provided researchers with  
16 an effective means to characterize large populations throughout the growing season. These  
17 longitudinal phenotypes can provide important insight into plant development and responses  
18 to the environment. Despite the growing use of these new phenotyping approaches in plant  
19 breeding, the use of genomic prediction models for longitudinal phenotypes is limited in  
20 major crop species. The objective of this study is to demonstrate the utility of random  
21 regression (RR) models using Legendre polynomials for genomic prediction of shoot growth  
22 trajectories in rice (*Oryza sativa*). An estimate of shoot biomass, projected shoot area  
23 (PSA), was recored over a period of 20 days for a panel of 357 diverse rice accessions using  
24 an image-based greenhouse phenotyping platform. A RR that included a fixed second-order  
25 Legendre polynomial, a random second-order Legendre polynomial for the additive genetic  
26 effect, a first-order Legendre polynomial for the environmental effect, and heterogeneous  
27 residual variances was used to model PSA trajectories. The utility of the RR model over  
28 a single time point (TP) approach, where PSA is fit at each time point independently, is  
29 shown through four prediction scenarios. In the first scenario, the RR and TP approaches  
30 were used to predict PSA for a set of lines lacking phenotypic data. The RR approach showed  
31 a 11.6% increase in prediction accuracy over the TP approach. Much of this improvement  
32 could be attributed to the greater additive genetic variance captured by the RR approach.  
33 The remaining scenarios focused forecasting future phenotypes using a subset of early time  
34 points for known lines with phenotypic data, as well new lines lacking phenotypic data. In all  
35 cases, PSA could be predicted with high accuracy ( $r$ : 0.79 to 0.89 and 0.55 to 0.58 for known  
36 and unknown lines, respectively). This study provides the first application of RR models for  
37 genomic prediction of a longitudinal trait in rice, and demonstrates that RR models can be

38 effectively used to improve the accuracy of genomic prediction for complex traits compared  
39 to a TP approach.

40 **Keywords:** Genomic prediction, high-throughput phenotyping, phenomics, genetics

## 41 1 Introduction

42 With the advent of next-generation sequencing technologies, the biology community has  
43 experienced a rapid increase in the amount of genotypic data that is available. These de-  
44 velopments, along with the low cost of sequencing, has encouraged the adoption of genomic  
45 selection (GS) approaches in plant breeding. With these approaches, genome-wide SNP  
46 markers are used to estimate an individuals additive genetic contribution to a given trait,  
47 and genotyped individuals can be selected and advanced to further generations without  
48 phenotypic evaluation (Meuwissen et al., 2001; Jannink et al., 2010; Endelman, 2011). Al-  
49 though these approaches have increased genetic gain through the acceleration of breeding  
50 cycles, considerable resources must still be devoted to the accurate phenotypic evaluation of  
51 individuals (Furbank and Tester, 2011). This necessary step remains a major bottleneck for  
52 many breeding programs.

53 In recent years, considerable investment, in both the public and private sector, have been  
54 made to automate the phenotypic characterization of large populations. Large investments  
55 have been made to build high-throughput phenotyping facilities in both the greenhouse  
56 and field where highly controlled water, nutrient, or temperature regimes can be applied to  
57 individual plots, and plants can be routinely monitored throughout the development using  
58 imaging. Moreover, the relatively low cost of drones that can be fitted with cameras and other  
59 sensors, have provided researchers with an effective means to characterize large populations  
60 throughout the growing season (Furbank and Tester, 2011; Chapman et al., 2014; Zhang  
61 et al., 2016; Watanabe et al., 2017). These longitudinal phenotypes can provide important  
62 insight into the mechanisms that underlie physiological responses to environmental stresses  
63 and developmental processes, and can be leveraged to improve prediction accuracies for  
64 complex polygenic traits, such as yield that have been a target for most breeding programs  
65 (Fahlgren et al., 2015; Campbell et al., 2017; Sun et al., 2017). Despite the growing use of

66 these new phenotyping approaches in plant breeding, the use of models for genomic selection  
67 (GS) for longitudinal phenotypes is limited in breeding major crop species. Most conventional  
68 field studies involve one or a few evaluations throughout the growing season, thus repeated  
69 phenotypic measurements on the same plant or plot is relatively rare.

70 Several approaches have been utilized for GS using longitudinal data. A simple repeata-  
71 bility (SR) model was used by Sun et al. (2017) and Rutkoski et al. (2016) for secondary  
72 longitudinal traits. The SR model treats each time point as a repeated measure of the same  
73 trait and assumes that the variance for all records are equal and the correlation between  
74 time points is constant. However, for many traits recorded across many time points, the  
75 assumption behind SR model is not realistic. A multivariate approach can be extended  
76 to longitudinal data. However, the computational complexity of the multivariate approach  
77 increases with the number of time points, and becomes unfeasible with high frequency lon-  
78 gitudinal traits due to the large number of parameters to estimate. Often, the number of  
79 observations necessary to accurately estimate parameters exceeds the size of most studies.

80 Random regression (RR) models have proven to be an attractive alternative to the above  
81 methods, and have been utilized in livestock and tree breeding (Apiolaza et al., 2000; Bermejo  
82 et al., 2003; Nobre et al., 2003; Bohmanova et al., 2008; Costa et al., 2008; Wetten et al., 2012;  
83 Howard et al., 2015). Here, covariance functions are explicitly defined that are equivalent  
84 to the full covariance matrix of the trait across time points (Kirkpatrick et al., 1990; Meyer,  
85 1998). Covariance functions include, but are not limited to banded correlation, autoregressive  
86 models, orthogonal polynomials, or spline functions (Meyer, 1998; Apiolaza et al., 2000).  
87 Thus, these models utilize a few parameters to describe the full covariance, and are much  
88 more computationally efficient. In animal breeding, RR models have been used extensively  
89 to estimate hertiabilities and perform pedigree-based prediction of important longitudinal  
90 traits such as growth, feed intake, fat, and milk production (Bermejo et al., 2003; Nobre  
91 et al., 2003; Bohmanova et al., 2008; Costa et al., 2008; Wetten et al., 2012; Howard et al.,

92 2015).

93 The increased accessibility to high-throughput phenotyping platforms provides the plant  
94 science community with high frequency temporal measurements for complex polygenic phe-  
95 notypes. These data are very different from those typically used for genomic prediction in  
96 which phenotypes are recorded at a single time point or at harvest for large populations.  
97 However, the availability of these new data presents an opportunity to extend these ap-  
98 proaches used extensively for longitudinal traits in animal breeding to major crops. Here,  
99 we demonstrate the use of RR models to predict shoot growth trajectories in a rice diversity  
100 panel. Specifically, the aims of this study are to (1) examine the advantage of utilizing  
101 longitudinal phenotypes over single end-point measurements (cross-sectional GS), (2) de-  
102 termine whether longitudinal phenotypes collected during early time-points can be used to  
103 predict phenotypes at later time points (i.e. forecasting lines with records), and (3) predict  
104 future phenotypes for new lines using early records for existing lines.

## 105 **2 Materials and Methods**

### 106 **2.1 Plant materials and greenhouse conditions**

107 Three hundred seventy eight lines of the Rice Diversity Panel 1 were selected for this study  
108 (Zhao et al., 2011). Seed propagation is described in Campbell et al. (2015). Three uniformly  
109 germinated seedlings were selected and transplanted to pots (150mm diameter x 200 mm  
110 height) filled with approximately 2.5 kg of UC Mix (the actual weight varied from experiment  
111 to experiment by 100-200 g). Square containers were placed below each pot to allow water  
112 to collect.



## 113 2.2 Experimental Design

114 All experiments were conducted at the Plant Accelerator, Australian Plant Phenomics Fa-  
115 cility, at the University of Adelaide, SA, Australia. Each experiment consisted of 378 lines  
116 and was repeated three times from February to April 2016. Two smarthouses were used  
117 for each experiment, with 216 pots positioned across 24 lanes in each smarthouse. Each  
118 experiment consisted of a partially replicated design, with 54 randomly selected lines having  
119 two replicates in each experiment.

120 Seven days after transplant (DAT), plants were thinned to one seedling per pot. Two  
121 layers of blue mesh was placed on top of the pots to reduce soil water evaporation. The  
122 plants were loaded on the imaging system and were watered to 90% field capacity at 11  
123 DAT.

## 124 2.3 Image analysis

125 The plants were imaged daily from 13 to 33 DAT using a visible (red–green–blue camera;  
126 Basler Pilot piA2400–12 gc, Ahrensburg, Germany) from two side-view angles separated  
127 by 90° and a single top view. The three experiments produced a total of 72,537 images.  
128 "Plant pixels" were extracted from RGB images using the LemnaGrid software. Briefly,  
129 plant pixels were extracted from background objects using a color classification strategy.  
130 Two set of colors were chosen manually to represent plant and background objects. For  
131 each image, pixels were assigned as background or plant pixels using the nearest-neighbor  
132 method. For a given pixel, this method assigns the pixel to a predefined color by finding  
133 the most similar (smallest Euclidean distance) color in the set. Noise (i.e. small areas of  
134 non-plant pixels) in the image is removed using a series of erosion and dilation steps.

135 The sum of the "plant pixels" from the three RGB images were summed, and used as a  
136 measure of shoot biomass. Here this trait is referred to as projected shoot area (PSA). This

137 metric has been shown to be an accurate representation of shoot biomass (Campbell et al.,  
138 2015; Goltzarian et al., 2011; Knecht et al., 2016). Prior to downstream analyses, outlier  
139 plants at each time point were detected for each trait using the 1.5(IQR) rule. Plants that  
140 were flagged as potential outliers were plotted and inspected visually. Those that exhibited  
141 abnormal growth patterns were removed. A total of 32 plants were removed, leaving a total  
142 of 2,604 plants for downstream analyses.

## 143 2.4 Selection of random regression models

144 PSA was modeled across all twenty time points using several RR models. Following the  
145 notation of Mrode (2014), the RR models can be summarized as

$$\text{PSA}_{tjk} = \mu + \sum_{k=0}^2 \phi(t)_{jk} \beta_k + \sum_{k=0}^{nr} \phi(t)_{jk} u_{jk} + \sum_{k=0}^{nr} \phi(t)_{jk} s_{jk} + e_{tjk} \quad (1)$$

146 Here  $\beta$  is the fixed second-order Legendre polynomial to model the overall trend in the trait  
147 overtime,  $u_{jk}$  and  $s_{jk}$  are the  $k^{\text{th}}$  random regression coefficients for additive genetic effect and  
148 random experiment of line  $j$ ,  $nr$  is the order of polynomial for the random effects, and  $e_{tjk}$   
149 is the random residual. The order of  $\beta$  was selected based on visual inspection of the trends.  
150 Various polynomial functions and residual variance structures were evaluated for line and  
151 experiment, and residuals, respectively. A complete description of the models is provided  
152 in Table 1. For each trait, the models were ranked based on goodness-of-prediction using  
153 Akaike's information criterion (AIC) scores (Akaike, 1974).

## 154 2.5 Genomic selection at each time point

155 A mixed model approach was used to fit genomic best linear unbiased predictions (gBLUPs)  
156 at each time point using the following model.

$$\mathbf{y} = \mathbf{Z}\mathbf{u} + \mathbf{Q}\mathbf{s} + \mathbf{e}, \quad (2)$$

Here,  $\mathbf{y}$  is the PSA at time  $t$ ;  $\mathbf{Z}$  and  $\mathbf{Q}$  are incidence matrices corresponding to the random additive genetic effect ( $\mathbf{u}$ ), and random experimental effect ( $\mathbf{s}$ ), respectively; and  $\mathbf{e}$  is the random residual error. For the random terms we assume  $\mathbf{u} \sim N(0, \mathbf{G}\sigma_g^2)$ ,  $\mathbf{s} \sim N(0, \mathbf{I}\sigma_s^2)$ , and  $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$ . Here,  $\sigma_g^2$  is the additive genetic variance;  $\sigma_s^2$  is an environmental variance associated with experiment; and  $\sigma_e^2$  is the residual variance. A genomic relationship matrix ( $\mathbf{G}$ ) was calculated using VanRaden (2008).

$$\mathbf{G} = \frac{\mathbf{Z}_{\text{cs}}\mathbf{Z}_{\text{cs}}'}{m} \quad (3)$$

157 Here,  $\mathbf{Z}_{\text{cs}}$  is a centered and scaled  $n \times m$  matrix, where  $m$  is 33,674 SNPs and  $n$  is the 357  
158 genotyped rice lines.

## 159 2.6 Genomic selection using random regression

160 For each trait, the "best" random regression model was used to predict gBLUPs. The  
161 following mixed model was used to predict gBLUPs

$$\text{PSA}_{tjk} = \mu + \sum_{k=0}^2 \phi(t)_{jk}\beta_k + \sum_{k=0}^2 \phi(t)_{jk}u_{jk} + \sum_{k=0}^1 \phi(t)_{jk}s_{jk} + e_{tjk} \quad (4)$$

162 The variables are the same as in *Selection of random regression models*, however note that  $nr$   
163 has been replaced with 2 and 1 for the additive genetic and experiment effect, respectively.  
164 Thus the random additive genetic effects are described using a second-order Legendre poly-  
165 nomial, while a first-order Legendre polynomial is used to describe the experiment effects

166 across time points.

In matrix notation, the model is

$$\mathbf{y} = \mathbf{Z}\mathbf{u} + \mathbf{Q}\mathbf{s} + \mathbf{e}, \quad (5)$$

167 with all vectors and matrices defined as above. However here  $u$  is now a vector of random  
168 regression coefficients for the additive genetic effects. For the random terms we assume  
169  $\mathbf{u} \sim N(0, \mathbf{G} \otimes \mathbf{\Omega})$ ,  $\mathbf{s} \sim N(0, \mathbf{I} \otimes \mathbf{P})$ , and  $\mathbf{e} \sim N(0, \mathbf{I} \otimes \mathbf{D})$ . Here,  $\mathbf{\Omega}$  is a  $3 \times 3$  covariance  
170 matrix of random regression coefficients for additive genetic effects;  $\mathbf{P}$  is a  $2 \times 2$  covariance  
171 matrix of random regression coefficients for experiment effect; and  $\mathbf{D}$  is a diagonal matrix  
172 allowing for heterogeneous variances over time points.  $\mathbf{Z}$  and  $\mathbf{Q}$  are covariable matrices  
173 where the  $i$ th row contains the orthogonal polynomials for the  $i$ th day of imaging. Thus,  
174 matrix  $\mathbf{Z}$  is the covariable matrix for the additive genetic effects with a dimension of  $t \times nk$   
175 where  $nk$  is the order of Legendre polynomial for the additive genetic effect multiplied by  
176 the number of individuals with phenotypic records and  $t$  refers to the number of days of  
177 imaging. Similarly,  $\mathbf{Z}$  is a  $t \times ns$  covariable matrix for the experiment effect, where  $ns$  is  
178 the the order of the Legendre polynomial for the experiment effect (e.g. 1) time the number  
179 of experiments (e.g. 3). Variance components and gBLUPs were obtained using ASREML  
180 (Release 4.0) (Gilmour et al., 2015).

181 Using the method above, variance components were obtained for additive genetic and  
182 environmental components. For the additive genetic term, each line has three random re-  
183 gression coefficients ( $nr = 0, 1, 2$ ). gBLUPs were predicted at each time point according to  
184 Mrode (2014). For a given line,  $j$ , at time  $t$  the gBLUPs can be obtained by  $\text{gBLUP}_{jt} = \phi_t \hat{u}_j$ ;  
185 where  $\phi_t$  is the row vector of the matrix of Legendre polynomials of order 2.

## 186 2.7 Estimation of narrow-sense heritability

187 To estimate the narrow sense heritability, variance components were obtained for each ran-  
188 dom term using ASREML for the TP analyses and the RR approach. For the RR approach,  
189 additive genetic variance was obtained at each time points using methods described by Mrode  
190 (2014). Briefly, for time  $i$  the genetic variance can be obtained by  $\mathbf{t}_i\mathbf{\Omega}\mathbf{t}'_i$ , where  $\mathbf{t}_i = \phi_{ik}$ ,  
191 the  $i$ th row vector of the matrix of Legendre polynomials at different time points ( $\phi$ ) for the  
192  $i$ th day of imaging,  $\mathbf{\Omega}$  is the covariance matrix of RR coefficients for the genetic effects, and  
193  $k$  is the order of fit. The variance of the experimental effect across time points was calcu-  
194 lated using the same approach. For both the single time point analysis  $h^2$  was estimated as  
195 
$$\frac{\sigma_g^2}{\sigma_g^2 + \sigma_s^2 + \sigma_e^2}.$$

## 196 2.8 GS scenarios and cross validation

197 Four scenarios were tested using GS (Figure 1). In the first scenario (scenario A), all twenty  
198 time points were used to fit a RR model and phenotypes were predicted for a set of lines  
199 without phenotypic records. The second scenario (scenario B), the dataset was split into  
200 two datasets each consisting of ten consecutive time points. A RR model was fitted using  
201 the first ten time points and was used to predict the phenotypes for the same set of lines  
202 in the last ten time points. Scenario C, can be thought of as a combination of scenarios A  
203 and B. Here, the dataset was split into four subsets, with each quadrant consisting of 178 to  
204 179 lines and ten time points. Here, a RR model was fitted using ten early time points for  
205 half the lines with known phenotypes, and was used to predict the phenotypes in the last  
206 ten time points for the remaining 178 to 179 lines. Finally, in the last scenario (Scenario D)  
207 we sought to predict the shoot biomass at a later time points in an independent study. This  
208 can be thought of as forecasting for new lines in an independent study. A publicly available  
209 dataset was used in which 359 lines (357 lines in common between the two studies) were

210 phenotyped from 20 to 40 days after transplant, thus a 13 day overlap was available for the  
211 two datasets, and a RR model was fitted using phenotypic information from the time points  
212 in the first experiment for 179 lines, and was used to predict gBLUPs for the remaining 178  
213 lines in a second independent experiment described by Campbell et al. (2017).

214 To assess the accuracy of gBLUPs for the TP GS as well as scenarios A, C, and D, a  
215 two-fold cross validation approach was used. Briefly, the 357 lines were split into two sets,  
216 with one serving as a training set with known phenotypes and the second serving as a testing  
217 set with unknown phenotypes. Since the number of lines were not even the remaining line  
218 was assigned to the training set. The accuracy of prediction was assessed by comparing  
219 predicted gBLUPs with observed PSA at each of the three experiments using Pearson's  
220 correlation method. The lines were randomly assigned to each fold, and the process was  
221 repeated 20 times. For each fold, the average correlation over the three experiments was  
222 used, and the average over the two folds was used for each resampling run. For scenario B,  
223 half of the lines were randomly selected and the first ten time points were used to predict the  
224 phenotypes in the last ten time points for the same lines. Again, the variance in prediction  
225 accuracy was assessed by randomly sampling half the lines for analysis. Pearson's correlation  
226 was computed for the gBLUPs and PSA as described above.

## 227 **3 Results**

228 A rice diversity panel was phenotyped over a period of twenty days during the early vegetative  
229 stage using an automated high-throughput phenotyping platform. The panel consists of 357  
230 lines from 80 countries, and captures much of the genetic diversity within cultivated rice  
231 (Zhao et al., 2011).

232 The plants were imaged each day using RGB cameras from three angles (two side view  
233 angles separated by 90 degrees and one top view). The plant pixels from each image were

234 summed and used to estimate shoot biomass. Here, this metric is referred to as PSA and  
235 has been shown to be an accurate measure of shoot biomass in cereals (Berger et al., 2010;  
236 Campbell et al., 2015). This experiments captures the early vegetative stage of development,  
237 where shoot biomass increases nearly exponentially (Figure 2A, Figure S1).

### 238 **3.1 Random regression model selection**

239 RR models have been used extensively to model longitudinal phenotypes in animal breeding.  
240 These models are particularly advantageous in that differences in the shape of the curve  
241 can be accounted for, and can be solved using the conventional mixed model framework.  
242 Thus, in the scope of genetics, these models allow for inter-individual variation in the mean  
243 trend to be estimated. Here, the overall mean growth trend was modeled using a second-  
244 order Legendre polynomial. A total of eight models were evaluated to identify a model  
245 that adequately described the data and could be used for GS. Each model included a fixed  
246 second-order Legendre polynomial to describe the overall mean growth trend, while several  
247 Legendre polynomials ranging from zero to second-order Legendre polynomials were fitted for  
248 random genetic and experimental effects. The residual effects were assumed to be constant  
249 or heterogeneous across time points using an identity or diagonal matrix, respectively. The  
250 "best" model was selected based on the smallest AIC value. Table S1 provides an overall  
251 summary of the models and the corresponding AIC values. The "best" model (Model 8) was  
252 one that included a fixed second-order polynomial to model the mean trend in shoot growth,  
253 a second-order Legendre polynomial for the random additive genetic effect, a first-order  
254 Legendre polynomial for the experimental effect, and the residual variance was assumed to  
255 be heterogeneous over time points. Figure 2B shows the predicted PSA obtained with model  
256 8 for two lines with contrasting contrasting genetic values for the RR coefficients.

## 257 **3.2 Genetic correlation and narrow sense heritability of PSA**

258 To examine the relationship for PSA between time points, the phenotypic and genetic cor-  
259 relation was estimated. Estimates for the overall phenotypic correlations were high ( $r$ : 0.49  
260 - 1.0), with the highest correlation observed between adjacent time points (Figure 3A). The  
261 genetic correlation followed a similar pattern, with an overall high correlation ( $r$ : 0.84 - 1.0)  
262 observed among pairwise comparisons of all 20 time points. As above, adjacent time points  
263 exhibited the highest genetic correlation ( $r = 1$ ), while those further apart exhibited lower  
264 correlation (Figure 3B). Interestingly, a strong genetic correlation was observed between day  
265 1 and day 20 ( $r = 0.91$ ), indicating that shoot growth (e.g. PSA) may be driven by similar  
266 genetic mechanisms at the early seedling and active tillering stage in rice.

267 To evaluate the ability of the longitudinal RR approach to capture additive genetic vari-  
268 ance, the narrow sense heritability of PSA was estimated using the RR model described  
269 above and a conventional mixed model at each time point. The mixed model included ran-  
270 dom terms for the additive genetic and experimental effect. For both models, a genomic  
271 relationship matrix was generated using 33,674 markers for the 357 lines. On average, the  
272 RR approach showed a 44% increase in the heritability of PSA compared to the TP approach  
273 (Figure 4). The TP approach showed a mean  $h^2$  of 0.50 over all time points, while the RR  
274 approach showed an  $h^2$  of 0.71 on average.  $h^2$  ranged from 0.60 to 0.77 for the RR approach,  
275 while  $h^2$  ranged from 0.46 to 0.57 for the TP approach. The two approaches showed nearly  
276 identical  $h^2$  estimates on day 1, however at later time points  $h^2$  of RR was considerably  
277 higher than TP. These results suggest that the RR approach captures more additive genetic  
278 variance for PSA than the TP approach.



### 279 **3.3 Utility of longitudinal phenotypes for genomic prediction**

280 The availability of high throughput phenotyping platforms provides a means to accurately  
281 phenotype large populations for a number of traits throughout time. While phenotypes  
282 recorded at a high frequency over time will likely improve the accuracy of GS, few reports  
283 have demonstrated the advantages of longitudinal phenotypes in major crops or model plant  
284 systems. Here, the utility of longitudinal phenotypes for GS was evaluated under four hypo-  
285 theoretical scenarios (Figure 1). The first scenario can be thought of as a standard GS approach  
286 (Figure 1A). Here, all 20 time points for half of the 357 lines used to predict the phenotypes  
287 at all 20 time points for the remaining lines. The aim of scenario A is to determine whether  
288 the longitudinal RR approach provides greater prediction accuracy than a cross-sectional  
289 GS approach in which a mixed model is fit at each time point. The first training set can be  
290 thought of as existing lines with phenotypic records and the test population as a new set of  
291 lines without records. The aim of scenario B (Figure 1B), is to determine if traits at later  
292 time points can be predicted for known lines using information at early time points. Thus, it  
293 can be considered as a forecasting approach. Here, longitudinal phenotypes are available for  
294 lines during the early time points (1-10 days of imaging), and are used to predict phenotypes  
295 for the same lines at later time points. Scenario C (Figure 1C), can also be considered a  
296 forecasting approach however for new lines. Here a subset of lines with phenotypes during  
297 the first 10 time points are used to predict the phenotypes for new lines without phenotypes  
298 at the later time points. In scenario D (Figure 1D), we sought to predict the shoot biomass  
299 at a later time points in an independent study. Here, a publicly available dataset was used  
300 in which 359 lines (357 lines in common between the two studies) were phenotyped from 20  
301 to 40 days after transplant, thus a 13 day overlap was available for the two datasets. A RR  
302 model was fitted using phenotypic information from the time points in the first experiment  
303 for 179 lines, and was used to predict gBLUPs for the remaining 178 lines in the second  
304 experiment.

## 305 **Scenario A: Comparison between longitudinal RR and cross-sectional GS**

306 To evaluate the advantages of using the longitudinal phenotype for PSA for GS over a single  
307 time points, the prediction accuracy of the RR model described above was compared to a  
308 conventional cross-sectional approach in which the additive genetic effects were estimated  
309 at each time point. For both approaches, two-fold cross validation was performed in which  
310 half the lines were randomly selected as a training set, and the remaining half was used for  
311 prediction. Pearson's correlation was used to assess the accuracy between predicted gBLUPs  
312 and observed PSA in the test set for each experiment. The average correlation across all  
313 three experiments was determined for each fold. The resampling process was repeated ten  
314 times.

315 Overall, the RR model showed significantly higher prediction accuracies than the TP  
316 approach (Figure 5A). On average, the longitudinal phenotype improved prediction accuracy  
317 by 11.6% (mean across all time points) compared to the TP approach. The prediction accu-  
318 racies for the TP approach ranged from 0.40 to 0.60, while for the RR approach accuracies  
319 ranged from 0.47 to 0.58. Although the TP approach exhibited low prediction accuracies  
320 during the early time points and increasing prediction accuracies toward the end of the study,  
321 the prediction accuracy for the RR model remained relatively constant with a slight increase  
322 in  $r$  observed from day 1 to 9. The largest improvements in prediction accuracy was observed  
323 between 5 to 10 days of imaging, with the RR model showing 35% higher accuracy at day  
324 8 compared to the TP approach. Collectively, these results indicate that RR models can be  
325 used to improve the accuracy of genomic prediction for longitudinal phenotypes.

## 326 **Scenario B: Forecasting existing lines**

327 Here, the the objective is to predict future phenotypes for lines with phenotypic trajectories  
328 recorded earlier in the growing season or development. To this end, the dataset was separated  
329 into two, with the first ten time points serving as a training set to predict the phenotypes

330 for the last ten days. This approach is described in Figure 1B. The RR model described  
331 above was fit to the data. To assess the accuracy of prediction, two-fold cross validation was  
332 performed in which 50% of the lines were randomly selected for training and prediction, and  
333 the resampling process was repeated ten times. The accuracy of prediction was very high,  
334 ranging from 0.79 to 0.82 for the last ten time points without phenotypic records (Figure  
335 5B). A slight decline in prediction accuracy was observed after day 10, with day 11 exhibiting  
336 the highest accuracy ( $r = 0.82$ ) and the lowest accuracy on day 20 ( $r = 0.79$ ). This trend in  
337 prediction accuracy is expected, given that the phenotypic records at day 11 should be very  
338 highly correlated with those at day 10, with the correlation declining as time progresses.  
339 The high predictive ability observed indicates that the first ten time points is sufficient to  
340 accurately predict future phenotypes for known lines.

### 341 **Scenario C: Forecasting new lines**

342 As shown above, future phenotypes can be accurately predicted from longitudinal traits at  
343 early time points for existing lines. While the knowledge of performance of known lines at fu-  
344 ture time points may be beneficial in some applications, GS is most often used to select lines  
345 without prior knowledge of the phenotype. Previously in scenario A, we showed that pheno-  
346 types could be predicted accurately for new lines using the complete longitudinal phenotype.  
347 Here, the aim is to predict future phenotypes for new lines with no phenotypic records using  
348 early phenotypic records for existing lines. To this end, the dataset was partitioned into  
349 two temporal datasets, with the first ten time points serving as a training set to predict the  
350 phenotypes for the last ten days (Figure 1C). As above, a two-fold cross validation approach  
351 was used to assess prediction accuracy. Half the lines were randomly assigned to each fold,  
352 and the first ten time points from the first fold were used to predict the phenotypes at the  
353 last ten time points in the second fold. The prediction accuracies for scenario C were very  
354 similar to those observed for scenario A. Accuracies ranged from 0.48 to 0.57, with the pre-

355 diction accuracy ranging from 0.55 to 0.57 in the last ten days (Figure 5C). The prediction  
356 accuracies showed a slight increase from day 1 to day 9. The highest prediction accuracy  
357 was observed at day 15, while the lowest accuracy was observed at day 1. These results  
358 suggest that future phenotypes can be forecast for new lines with reasonable accuracy using  
359 phenotypic records from earlier time points for a set of known lines.

#### 360 **Scenario D: Forecasting new lines at later time points in an independent study**

361 In scenario C, we have shown that gBLUPs for new lines can be accurately predicted using  
362 phenotypes for a set of known lines at a subset of early time points. Here, the objective  
363 was to expand this approach and evaluate the utility of the RR model to predict gBLUPs  
364 for new lines at future time points in an independent study. Here, we utilized an existing  
365 dataset where 359 lines from the Rice Diversity Panel 1 were phenotyped from 20 to 40 days  
366 after transplant (Figure 1D.). Although there is overlap between developmental stages of  
367 this dataset and the dataset used for scenarios A-C, this experiment was conducted at a  
368 different time of year and therefore the photoperiod and light intensity should be different  
369 between the two.

370 A RR model was fitted that was identical to that used for scenarios A-C, in that it  
371 included a fixed second-order polynomial to model the mean trend in PSA, a second-order  
372 Legendre polynomial for the random additive genetic effect, a first-order Legendre polynomial  
373 for the random experimental effect, and a heterogeneous residual variance over time points.  
374 The RR model was fitted using phenotypes for 179 lines from the early vegetative stage  
375 experiment (i.e. 13 to 32 DAT), and the genetic values for the RR coefficients were used to  
376 predict the phenotypes for the remaining 178 lines in the second experiment (i.e. 20 to 40  
377 DAT). A two-fold cross validation approach was used in which phenotypes across all twenty  
378 days were selected for 179 lines in the first experiment and were used to predict gBLUPs for  
379 the remaining 178 lines in the second experiment.

380 The prediction accuracy was high with  $r$  values ranging from 0.51 to 0.59 (Figure 5D).  
381 The prediction accuracy was relatively constant, but showed a slight increase in accuracy  
382 from 22 to 29 days after transplant. An increase in the prediction accuracy was observed  
383 from 13 to 31 DAT, after which the prediction accuracy declined slightly. The second time  
384 point (22 DAT) exhibited the lowest prediction accuracy ( $r = 0.51$ ). The highest prediction  
385 accuracy was observed on day 34 after transplant ( $r = 0.59$ ). Collectively, these results  
386 suggest that longitudinal phenotypes can be accurately predicted in an independent study  
387 using the RR approach.

### 388 **3.4 Discussion**

389 High-throughput phenotyping platforms provide an accessible means to record traits non-  
390 destructively for large populations throughout development. Such longitudinal data provide  
391 an opportunity to understand the genetics of the development of a phenotype, and identify  
392 individuals that exhibit desirable trait trajectories. However, such data provides new chal-  
393 lenges to adapt approaches utilized for single time point phenotypes in plant genomics and  
394 breeding to accommodate longitudinal data. This study provides the first application of RR  
395 models for genomic prediction of a longitudinal trait in rice.

#### 396 **Advantages of RR over univariate genomic prediction**

397 The predictive ability in GS is dependent on the heritability of the trait, the number of  
398 markers, population size, linkage disequilibrium (LD), and the number of QTL influencing  
399 the trait (Daetwyler et al., 2008, 2010). Here, the RR model using longitudinal phenotypes  
400 provided greater prediction accuracy compared to the TP gBLUP. The predictive ability  
401 of the RR approach improved prediction accuracies by 11.6% on average compared to TP  
402 analysis. The number of markers, population size, LD, and the number of QTL influencing  
403 PSA are held constant between the two models. Thus, the difference in prediction accuracy

404 hold be largely attributed to the differences in heritability between the RR approach and  
405 TP analysis. As shown in Figure 4, the RR approach accounted for more additive genetic  
406 variance than the TP analysis. Similar gains in heritability for height in Swedish Scots pine  
407 has been reported by Wang et al. (2009) with RR models that utilize B-splines or Legendre  
408 polynomials over TP analyses. Moreover, when the prediction accuracy is expressed as  
409 the ratio of the correlation of gBLUPs and observed PSA to the square root of  $h^2$ , both  
410 approaches were nearly equivalent (Figure S2). Thus, the higher prediction accuracy is due  
411 to the higher  $h^2$  of the RR approach relative to the TP approach.

412 With both methods (RR and TP), we observed high prediction accuracies ranging from  
413 0.4 - 0.6 (Arruda et al., 2015; Duhnen et al., 2017; Kristensen et al., 2018; Leplat et al.,  
414 2016). While similar accuracies have been reported by other studies for complex traits,  
415 it is important to note that the current study utilized a diversity panel with considerable  
416 population stratification and the prediction models did not account for population structure.  
417 Accounting for population structure is important in genome wide association studies to  
418 reduce spurious associations (Yu et al., 2006). However, these corrections can often hinder  
419 the ability to detect true QTL that are correlated with population structure (Zhao et al.,  
420 2011). With GS, the aim is to achieve high prediction accuracies across subpopulations rather  
421 than to detect QTL associated with the trait (Hayes et al., 2009; Lorenz et al., 2011). Thus,  
422 the high prediction accuracies observed for the models used in this study may be due, in part,  
423 to population structure, however the random sampling of individuals across subpopulations  
424 during CV should reduce the possibility of having a training set that is strongly imbalanced  
425 by a given subpopulation.

## 426 **Utilizing RR prediction for forecasting phenotypes**

427 The utilization of genomic information to predict future outcomes is not new. Considerable  
428 effort in the field of personalized medicine has been devoted to predict disease risks for in-

429 individuals based on genomic information. Here, disease-associated loci are used to predict  
430 potential future outcomes for individuals (Moser et al., 2015). The ability to predict future  
431 phenotypes using phenotypic information collected early in the life cycle may be advanta-  
432 geous in plants, particularly perennial species with long life cycles. Selection during the early  
433 developmental stage can shorten evaluation times.

434 Here, we evaluated the ability of RR models to predict future phenotypes using pheno-  
435 typic records collected during the early time points. This was performed for known lines  
436 (e.g. those with early records; Scenario B), as well as new lines (Scenario C and D). We  
437 observed high prediction accuracies for each forecasting scenario. As expected the highest  
438 accuracy was observed for Scenario B, in which early phenotypic records are used to predict  
439 future phenotypes for the same set of lines. Surprisingly, high prediction accuracies were  
440 also observed when early records for known lines were used to predict future phenotypes  
441 for unknown lines (Scenarios C and D). In both cases, the accuracies were not significantly  
442 different from those achieved when using phenotypic information for all time points. These  
443 results collectively indicate that the future phenotypes can be accurately predicted using a  
444 subset of the temporal phenotypes. While these results are encouraging, these forecasting  
445 approaches will be highly dependent on the temporal genetic architecture of the trait. The  
446 lack of decline by utilizing only a subset of time points is likely due to the high genetic cor-  
447 relation observed between time points. The similar genetic architecture between the early  
448 and late time points that is evidenced by the strong positive genetic correlation (Figure 3B)  
449 estimated between early (1-10 days) and late (11-20 days) time points. Thus, we suggest  
450 to first evaluate the genetic correlation between time points for the trait of interest before  
451 utilizing such forecasting approaches.

## 452 **3.5 Conclusion**

453 High throughput phenomics platforms have provided the plant science community with a  
454 means to generate high resolution temporal phenotypes for large populations at a relatively  
455 low cost. RR models that utilize Legendre polynomials provide a flexible for genomic pre-  
456 diction of longitudinal traits. These approaches provide several advantages over single time  
457 point analyses: (1) these models account for more additive genetic variance compared to  
458 the TP analysis, which translates to higher predictive accuracies; (2) future phenotypes can  
459 be accurately predicted using phenotypic information for earlier time points for known and  
460 unknown lines. TP



## 461 References

462 Akaike, H., 1974: A new look at the statistical model identification. *IEEE Transactions on*  
463 *Automatic Control*, **19 (6)**, 716–723.

464 Apiolaza, L. A., A. R. Gilmour, and D. J. Garrick, 2000: Variance modelling of longitudinal  
465 height data from a *Pinus radiata* progeny test. *Canadian Journal of Forest Research*,  
466 **30 (4)**, 645–654.

467 Arruda, M. P., P. J. Brown, A. E. Lipka, A. M. Krill, C. Thurber, and F. L. Kolb, 2015:  
468 Genomic selection for predicting fusarium head blight resistance in a wheat breeding pro-  
469 gram. *The Plant Genome*, **8 (3)**.

470 Berger, B., B. Parent, and M. Tester, 2010: High-throughput shoot imaging to study drought  
471 responses. *Journal of Experimental Botany*, **61 (13)**, 3519–3528.

472 Bermejo, J. L., R. Roehe, V. Schulze, G. Rave, H. Looft, and E. Kalm, 2003: Random  
473 regression to model genetically the longitudinal data of daily feed intake in growing pigs.  
474 *Livestock Production Science*, **82 (2)**, 189–199.

475 Bohmanova, J., F. Miglior, J. Jamrozik, I. Misztal, and P. Sullivan, 2008: Comparison of  
476 random regression models with legendre polynomials and linear splines for production  
477 traits and somatic cell score of Canadian Holstein cows. *Journal of Dairy Science*, **91 (9)**,  
478 3627–3638.

479 Campbell, M. T., Q. Du, K. Liu, C. J. Brien, B. Berger, C. Zhang, and H. Walia, 2017: A  
480 comprehensive image-based phenomic analysis reveals the complex genetic architecture of  
481 shoot growth dynamics in rice. *The Plant Genome*, **10 (2)**.

482 Campbell, M. T., A. C. Knecht, B. Berger, C. J. Brien, D. Wang, and H. Walia, 2015: Inte-

- 483 grating image-based phenomics and association analysis to dissect the genetic architecture  
484 of temporal salinity responses in rice. *Plant Physiology*, **168** (4), 1476–1489.
- 485 Chapman, S. C., and Coauthors, 2014: Pheno-copter: a low-altitude, autonomous remote-  
486 sensing robotic helicopter for high-throughput field-based phenotyping. *Agronomy*, **4** (2),  
487 279–301.
- 488 Costa, C. N., C. M. R. d. Melo, I. U. Packer, A. F. d. Freitas, N. M. Teixeira, and J. A. Cobuci,  
489 2008: Genetic parameters for test day milk yield of first lactation Holstein cows estimated  
490 by random regression using Legendre polynomials. *Revista Brasileira de Zootecnia*, **37** (4),  
491 602–608.
- 492 Daetwyler, H. D., R. Pong-Wong, B. Villanueva, and J. A. Woolliams, 2010: The impact of  
493 genetic architecture on genome-wide evaluation methods. *Genetics*, **185** (3), 1021–1031.
- 494 Daetwyler, H. D., B. Villanueva, and J. A. Woolliams, 2008: Accuracy of predicting the  
495 genetic risk of disease using a genome-wide approach. *PloS One*, **3** (10), e3395.
- 496 Duhnen, A., A. Gras, S. Teyssède, M. Romestant, B. Claustres, J. Daydé, and B. Mangin,  
497 2017: Genomic selection for yield and seed protein content in soybean: A study of breeding  
498 program data and assessment of prediction accuracy. *Crop Science*, **57** (3), 1325–1337.
- 499 Endelman, J. B., 2011: Ridge regression and other kernels for genomic selection with R  
500 package rrBLUP. *The Plant Genome*, **4** (3), 250–255.
- 501 Fahlgren, N., M. A. Gehan, and I. Baxter, 2015: Lights, camera, action: high-throughput  
502 plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology*, **24**, 93–99.
- 503 Furbank, R. T., and M. Tester, 2011: Phenomics–technologies to relieve the phenotyping  
504 bottleneck. *Trends in Plant Science*, **16** (12), 635–644.

- 505 Gilmour, A., B. Gogel, B. Cullis, S. Welham, and R. Thompson, 2015: Asreml user guide  
506 release 4.1 structural specification. *Hemel Hempstead: VSN International Ltd.*
- 507 Golzarian, M. R., R. A. Frick, K. Rajendran, B. Berger, S. Roy, M. Tester, and D. S. Lun,  
508 2011: Accurate inference of shoot biomass from high-throughput images of cereal plants.  
509 *Plant Methods*, **7** (1), 2.
- 510 Hayes, B. J., P. J. Bowman, A. C. Chamberlain, K. Verbyla, and M. E. Goddard, 2009:  
511 Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics*  
512 *Selection Evolution*, **41** (1), 51.
- 513 Howard, J. T., S. Jiao, F. Tiezzi, Y. Huang, K. A. Gray, and C. Maltecca, 2015: Genome-  
514 wide association study on Legendre random regression coefficients for the growth and feed  
515 intake trajectory on Duroc Boars. *BMC Genetics*, **16** (1), 59.
- 516 Jannink, J.-L., A. J. Lorenz, and H. Iwata, 2010: Genomic selection in plant breeding: from  
517 theory to practice. *Briefings in Functional Genomics*, **9** (2), 166–177.
- 518 Kirkpatrick, M., D. Lofsvold, and M. Bulmer, 1990: Analysis of the inheritance, selection  
519 and evolution of growth trajectories. *Genetics*, **124** (4), 979–993.
- 520 Knecht, A. C., M. T. Campbell, A. Caprez, D. R. Swanson, and H. Walia, 2016: Image  
521 Harvest: an open-source platform for high-throughput plant image processing and analysis.  
522 *Journal of Experimental Botany*, **67** (11), 3587–3599.
- 523 Kristensen, P. S., A. Jahoor, J. R. Andersen, F. Cericola, J. Orabi, L. Janss, and J. Jensen,  
524 2018: Genome-wide association studies and comparison of models and cross-validation  
525 strategies for genomic prediction of quality traits in advanced winter wheat breeding lines.  
526 *Frontiers in Plant Science*, **9**, 69.

- 527 Leplat, F., J. Jensen, and P. Madsen, 2016: Genomic prediction of manganese efficiency in  
528 winter barley. *The Plant Genome*, **9** (2).
- 529 Lorenz, A. J., and Coauthors, 2011: Genomic selection in plant breeding: knowledge and  
530 prospects. *Advances in Agronomy*, Vol. 110, Elsevier, 77–123.
- 531 Meuwissen, M., B. Hayes, and Goddard, 2001: Prediction of total genetic value using  
532 genome-wide dense marker maps. *Genetics*, **157** (4), 1819–1829.
- 533 Meyer, K., 1998: Estimating covariance functions for longitudinal data using a random  
534 regression model. *Genetics Selection Evolution*, **30** (3), 221.
- 535 Moser, G., S. H. Lee, B. J. Hayes, M. E. Goddard, N. R. Wray, and P. M. Visscher,  
536 2015: Simultaneous discovery, estimation and prediction analysis of complex traits us-  
537 ing a Bayesian mixture model. *PLoS Genetics*, **11** (4), e1004969.
- 538 Mrode, R. A., 2014: *Linear models for the prediction of animal breeding values*. CABI.
- 539 Nobre, P., I. Misztal, S. Tsuruta, J. Bertrand, L. Silva, and P. Lopes, 2003: Analyses of  
540 growth curves of Nelore cattle by multiple-trait and random regression models. *Journal*  
541 *of Animal Science*, **81** (4), 918–926.
- 542 Rutkoski, J., J. Poland, S. Mondal, E. Autrique, L. G. Pérez, J. Crossa, M. Reynolds,  
543 and R. Singh, 2016: Canopy temperature and vegetation indices from high-throughput  
544 phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat.  
545 *G3: Genes, Genomes, Genetics*, **6** (9), 2799–2808.
- 546 Sun, J., J. E. Rutkoski, J. A. Poland, J. Crossa, J.-L. Jannink, and M. E. Sorrells, 2017: Mul-  
547 titrait, random regression, or simple repeatability model in high-throughput phenotyping  
548 data improve genomic prediction for wheat grain yield. *The Plant Genome*, **10** (2).

- 549 VanRaden, P. M., 2008: Efficient methods to compute genomic predictions. *Journal of Dairy*  
550 *Science*, **91** (11), 4414–4423.
- 551 Wang, C., B. Andersson, and P. Waldmann, 2009: Genetic analysis of longitudinal height  
552 data using random regression. *Canadian Journal of Forest Research*, **39** (10), 1939–1948.
- 553 Watanabe, K., and Coauthors, 2017: High-throughput phenotyping of sorghum plant height  
554 using an unmanned aerial vehicle and its application to genomic prediction modeling.  
555 *Frontiers in Plant Science*, **8**.
- 556 Wetten, M., J. Ødegård, O. Vangen, and T. H. Meuwissen, 2012: Simultaneous estimation of  
557 daily weight and feed intake curves for growing pigs by random regression. *Animal*, **6** (3),  
558 433–439.
- 559 Yu, J., and Coauthors, 2006: A unified mixed-model method for association mapping that  
560 accounts for multiple levels of relatedness. *Nature Genetics*, **38** (2), 203.
- 561 Zhang, J., C. Yang, H. Song, W. C. Hoffmann, D. Zhang, and G. Zhang, 2016: Evaluation  
562 of an airborne remote sensing platform consisting of two consumer-grade cameras for crop  
563 identification. *Remote Sensing*, **8** (3), 257.
- 564 Zhao, K., and Coauthors, 2011: Genome-wide association mapping reveals a rich genetic  
565 architecture of complex traits in *Oryza sativa*. *Nature Communications*, **2**, 467.

## 566 Figure Legends

567 Figure 1: Graphic representation of cross validation schemes for predicting longitudinal phe-  
568 notypes using random regression. In (A), (C), and (D) two-fold cross validation was used,  
569 where phenotypes for 179 lines were used as a training set to predict phenotypes for the  
570 remaining 178 lines. In (A), all twenty time points for the training set were used to predict  
571 the phenotypes at each of the twenty time points for an new set of lines. The second scenario  
572 (B) can be thought of as a forecasting approach where the dataset was split into two longi-  
573 tudinal datasets each consisting of ten time points. The first ten time points for 179 lines  
574 and were used to predict the phenotypes at the last ten time points for the same 179 lines.  
575 In (C), a forecasting approach was again used, however the lines were randomly split in two,  
576 and the first ten time points were used to predict phenotypes in the last ten time points for  
577 a group of new lines. In (D) the first 20 time points was used to predict gBLUPs at a later  
578 time points in an independent study. Here, a publicly available dataset was used as a testing  
579 set in which 357 lines were phenotyped from 20 to 40 days after transplant, thus a 13 day  
580 overlap was available for the two datasets. Here, the independent dataset is indicated with  
581  $PSA_{LaterVeg.}$ . Excluded indicates that these data points were not included for analyses.

582

583 Figure 2: Projected shoot area (PSA) across twenty days of imaging. (A) Population mean  
584 for PSA across the twenty days of imaging. Here, the shaded region represents the standard  
585 deviation of PSA at each time point. (B) Predicted PSA for two contrasting lines using a  
586 random regression (RR) model. The RR model included a fixed second-order polynomial to  
587 model the mean trend in shoot growth, a second-order Legendre polynomial for the random  
588 additive genetic effect, a first-order Legendre polynomial for the experimental effect, and  
589 the residual variance was assumed to be heterogeneous over time points. The predicted RR  
590 coefficients for each line are provided in the figure legend. The shaded regions represent the

591 standard error of predicted PSA at each time point. Here, PSA is defined as the sum of  
592 plant pixel from three images (two side-view images and one top-view). The shaded region  
593 represents the standard deviation of PSA at each time point.

594

595 Figure 3: Phenotypic and genetic correlations between each time point. (A) Phenotypic  
596 correlations were estimated between time points using Pearson's method. (B) The inferred  
597 genetic correlation matrix of random regression terms for the additive genetic effects were  
598 used to estimate the genetic correlations between time points. The scale on the left of each  
599 panel indicates the strength of the correlations ( $r$ ).

600

601 Figure 4: Narrow sense heritability and variance components estimated using the single time  
602 point (TP) and random regression (RR) approaches. The narrow sense heritability ( $h^2$ ) is  
603 presented in panel A. Variance components for the TP and RR approaches are pictured  
604 in panels B and C, respectively. For the single time point analysis, a conventional mixed  
605 model was used to estimate the narrow sense heritability of PSA at each time point. The  
606 TP model included a random additive genetic effect and experimental effect. The RR model  
607 included a fixed second-order Legendre polynomial, the random additive genetic effect were  
608 modeled using a second-order Legendre polynomial, a first-order random effect was used for  
609 experiment, and the residual variance was assumed to be heterogeneous over time points.  
610 For both models, the experimental term was considered as an environmental effect.

611

612 Figure 5: Prediction accuracies of scenarios A to D. For the random regression (RR) ap-  
613 proach, a RR model was fit using phenotypic records for 178-179 lines over 20 days. A  
614 univariate single time point (TP) run using phenotypic records for 178-179 lines at each day.  
615 In both cases, genetic effects from each model were used to predict gBLUPs for the remain-  
616 ing 178-179 lines. Prediction accuracy was assessed using Pearson's correlation between the

617 predicted gBLUPs and observed PSA for the test set. Resampling was done twenty times.  
618 The error bars represent the standard deviation where  $n = 20$ . A comparison of prediction  
619 accuracies for TP and RR approaches is presented in (A). Panels B and C present the pre-  
620 diction accuracies for forecasting future phenotypes using phenotypic information at early  
621 time points for known lines (B) and new lines (C). Panel D provides prediction accuracies  
622 for forecasting future phenotypes in an independent study using phenotypes from an earlier  
623 developmental period.

624

625 Figure S1: Projected shoot area for a subset of 12 lines. The line identifier (NSFTV\_),  
626 experiment (Exp), and replicate (Rep) are provided in the plot titles.

627 Figure S2: Predictive ability of the random regression (RR) and single time point (TP)  
628 approaches expressed as a function of heritability: The analysis followed the same approach  
629 as that for scenario A, however for each fold the correlation between gBLUP and observed  
630 PSA was divided by the square root of heritability. The error bars represent the standard  
631 deviation where  $n = 20$ .

632



## 633 Acknowledgements

634 Funding for this research was provided by the National Science Foundation (United States)  
635 through Award No. 1238125 to Harkamal Walia, and Award No. 1736192 to Harkamal Walia  
636 and Gota Morota.

## 637 Data Accessibility

638 The full datasets and all code used in this study is available via [GitHub](https://github.com/malachycampbell/random-regression-models-for-genomic-prediction-of-a-longitudinal-trait-derived-from-HTP) ([https://github.com/malachycampbell](https://github.com/malachycampbell/random-regression-models-for-genomic-prediction-of-a-longitudinal-trait-derived-from-HTP)  
639 [random-regression-models-for-genomic-prediction-of-a-longitudinal-trait-derived-from-HTP](https://github.com/malachycampbell/random-regression-models-for-genomic-prediction-of-a-longitudinal-trait-derived-from-HTP))  
640 and the [WRCHR](http://WRCHR.org) website ([WRCHR.org](http://WRCHR.org)).

641 **Figures**

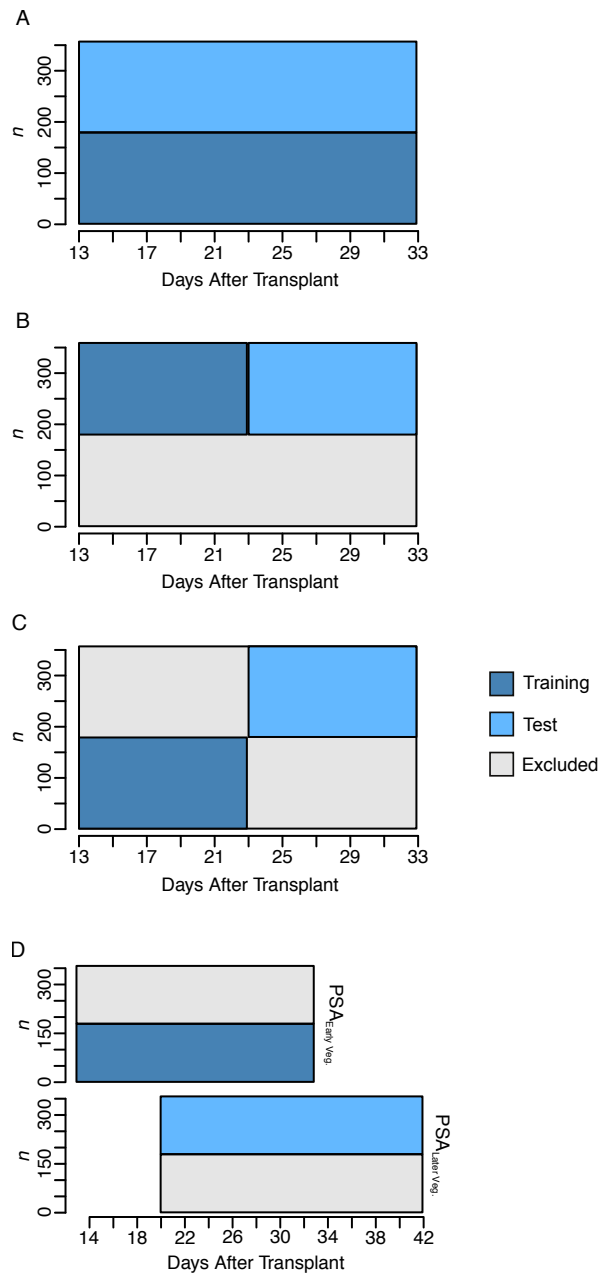


Figure 1

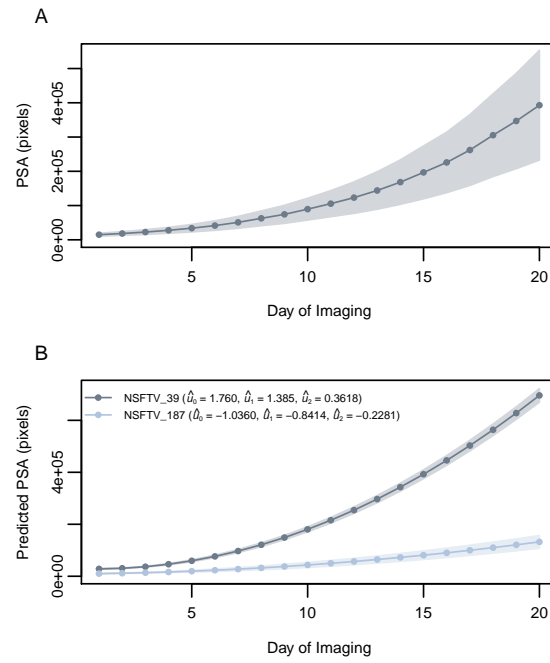


Figure 2

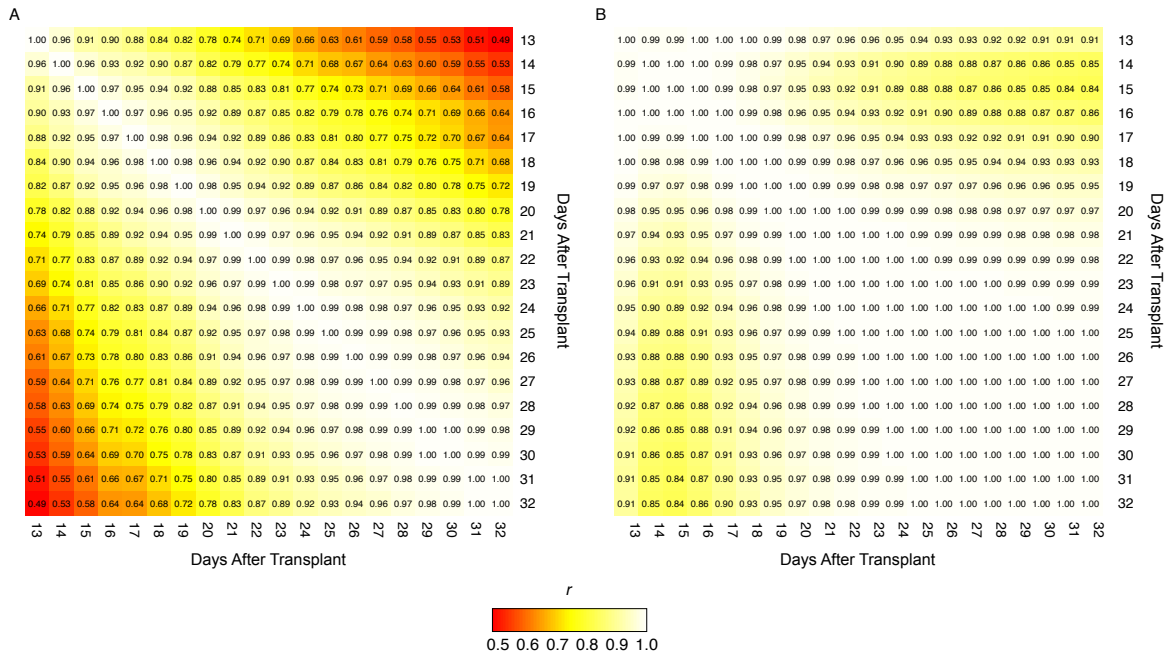


Figure 3

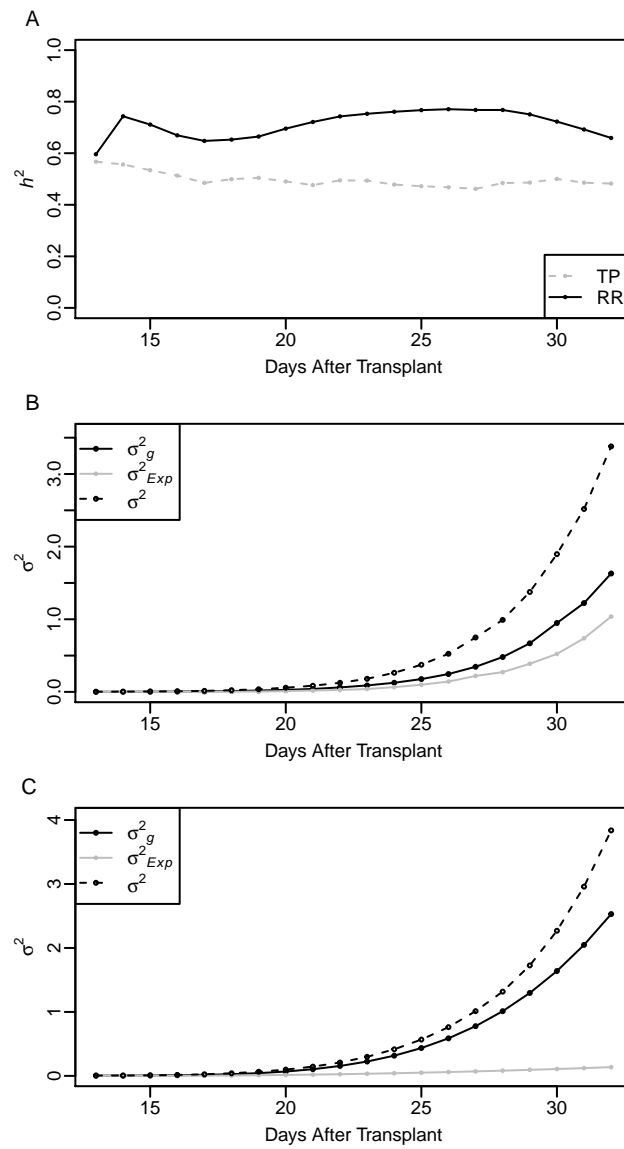


Figure 4

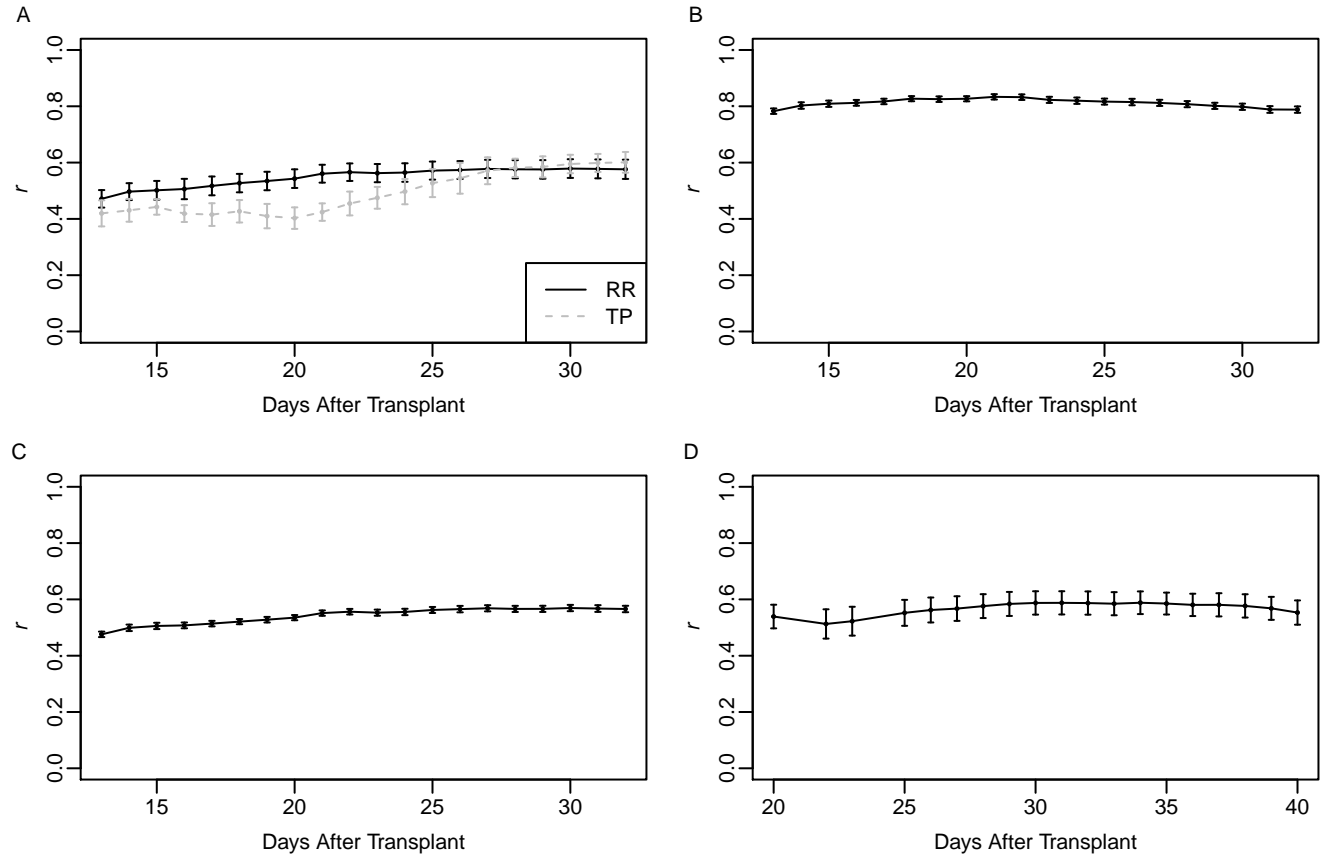


Figure 5

## 642 Supplemental Data

Table S1: Random regression model selection. Each of the four random regression models included a fixed second-order polynomial to model the mean trend in PSA over the twenty time points, indicated by the column  $f$ .  $G$  refers to the random additive genetic effect,  $Exp$  the random experimental effect, and  $e$  error term. Models with  $Diag$  assumed heterogeneous residual variance over time points, while those with  $I$  assumed the residual variance was constant.  $pol^n$  refers to a Legendre polynomial of order  $n$ .

<i>Model</i>	<i>f</i>	<i>G</i>	<i>Exp</i>	<i>e</i>	<i>LogREML</i>	<i>AIC</i>	<i>BIC</i>
Model 1	$pol^2$	$pol^0$	$pol^0$	$I$	2026.97	-4047.93	-4023.65
Model 2	$pol^2$	$pol^0$	$pol^0$	$Diag$	19358.83	-38673.65	-38495.60
Model 3	$pol^2$	$pol^1$	$pol^0$	$I$	7345.85	-14681.69	-14641.23
Model 4	$pol^2$	$pol^1$	$pol^0$	$Diag$	23273.62	-46499.24	-46305.01
Model 5	$pol^2$	$pol^2$	$pol^0$	$I$	8204.64	-16393.28	-16328.54
Model 6	$pol^2$	$pol^2$	$pol^0$	$Diag$	24718.93	-49383.86	-49165.35
Model 7	$pol^2$	$pol^2$	$pol^0$	$I$	12700.64	-25381.28	-25300.35
Model 8	$pol^2$	$pol^2$	$pol^1$	$Diag$	27537.59	-55017.19	-54782.49

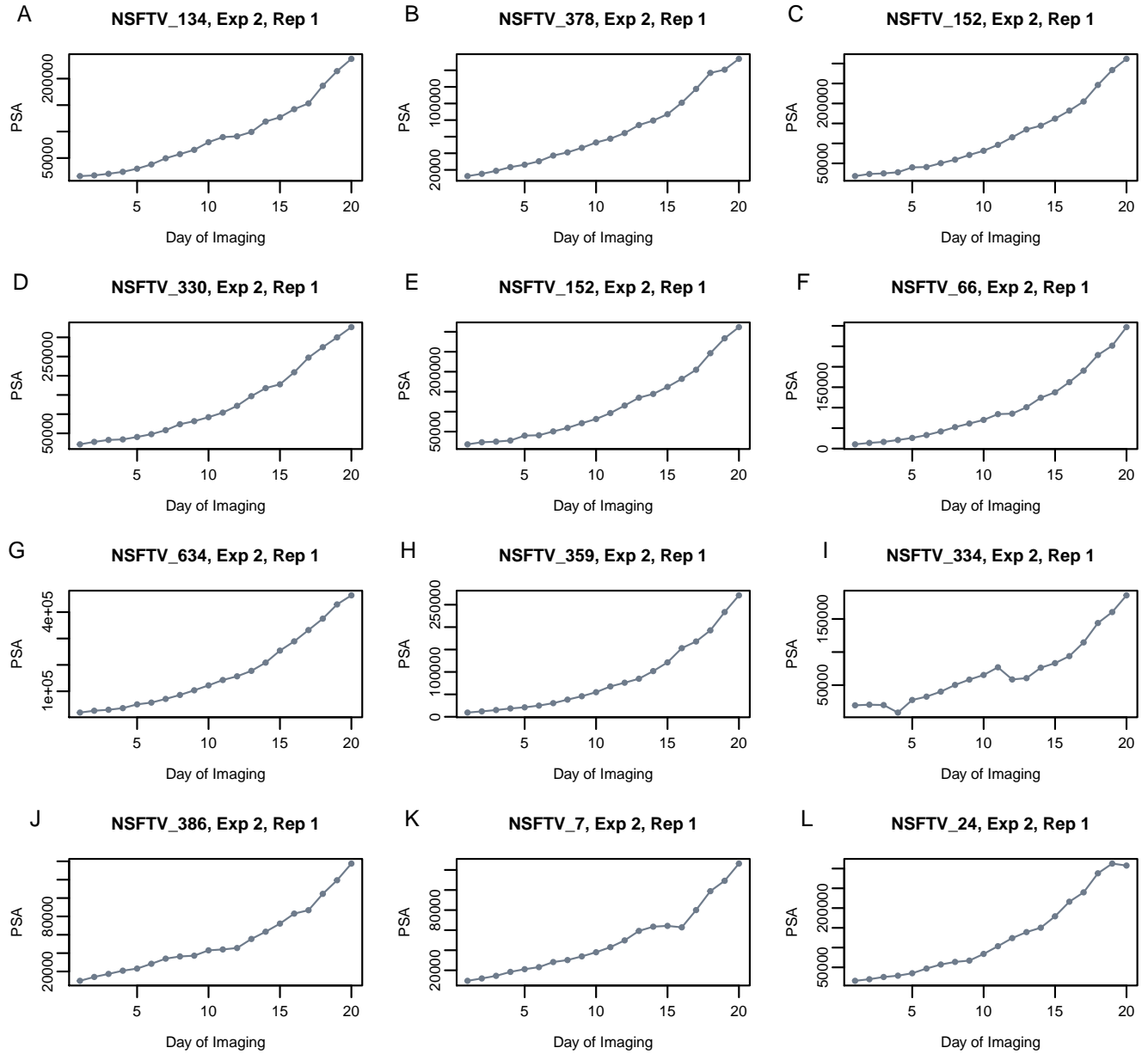


Figure S1: Projected shoot area for a subset of 12 lines. The line identifier (NSFTV\_), experiment (Exp), and replicate (Rep) are provided in the plot titles.



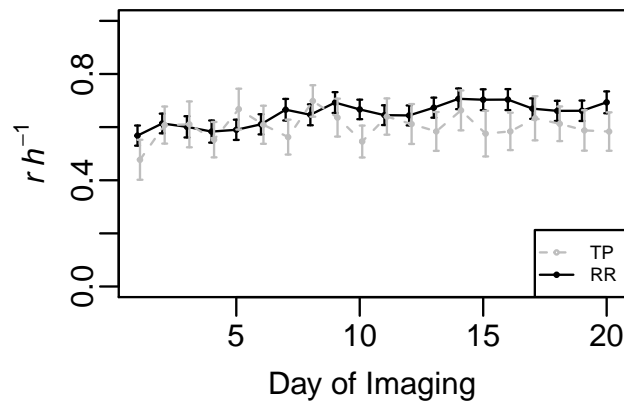


Figure S2: Predictive ability of the random regression (RR) and single time point (TP) approaches expressed as a function of heritability: The analysis followed the same approach as that for scenario A, however for each fold the correlation between gBLUP and observed PSA was divided by the square root of heritability. The error bars represent the standard deviation where  $n = 20$ .