Improving selection efficiency of crop breeding with a genomic prediction aided partial phenotyping strategy

Sang He¹,²,* Yong Jiang³, Rebecca Thistlethwaite⁴, Matthew J. Hayden¹,⁵, Richard Trethewan⁴,⁶, Hans D. Daetwyler¹,⁵,*

¹Agriculture Victoria, Centre for AgriBioscience, AgriBio, Bundoora, VIC 3083, Australia
²CAAS-IRRI Joint Laboratory for Genomics-Assisted Germplasm Enhancement, Agricultural Genomics Institute in Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China
³Department of Breeding Research, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany
⁴School of Life and Environmental Sciences, Plant Breeding Institute, Sydney Institute of Agriculture, The University of Sydney, Narrabri, NSW, Australia
⁵School of Applied Systems Biology, La Trobe University, Bundoora, VIC 3083, Australia
⁶School of Life and Environmental Sciences, Plant Breeding Institute, Sydney Institute of Agriculture, The University of Sydney, Cobbitty, NSW, Australia

*Correspondence (Tel +61 390327037, Email hans.daetwyler@agriculture.vic.gov.au; Tel +86 0755-28394429, Email hesang@caas.cn)

Abstract

Increasing the number of environments for phenotyping of crop lines in earlier stages of breeding programs can improve selection accuracy. However, this is often not feasible due to cost. In our study, we investigated a partial phenotyping strategy that does not test all entries in all environments, but instead capitalizes on genomic prediction to predict missing phenotypes in additional environments without extra phenotyping expenditure. The breeders’ main interest – response to selection – was directly simulated to evaluate the effectiveness of the partial genomic phenotyping strategy in a wheat dataset. Whether the partial phenotyping strategy
resulted in more selection response depended on the correlations of phenotypes between environments. The partial phenotyping strategy consistently showed statistically significant higher simulated responses to selection, compared to complete phenotyping, when the majority of completely phenotyped environments were negatively correlated and any extension environment was highly positively correlated with any of the completely phenotyped environments. Our results indicate that genomics-based partial phenotyping can improve selection response at middle stages of crop breeding programs.

**Key words:** partial phenotyping strategy; genomic prediction; response to selection; correlations between environments

**Declarations**

**Funding and acknowledgment**

This study is funded by the Grain Research Development Corporation (GRDC, US00081), the University of Sydney and Agriculture Victoria. We thank staff and students at the Plant Breeding Institute at Narrabri and AgriBio for phenotyping and genotyping, respectively.

**Conflict of interest**

All authors declare there is no conflict of interest

**Availability of data and material**

Data is available upon request for non-commercial purposes

**Code availability**

Available upon request

**Ethics approval**

Not applicable

**Consent to participate**
Author contribution statement

SH, HDD and YJ designed the study. SH conducted genomic prediction analyses and responses simulations. RTr and RTh developed the plant populations and collected the phenotypes. MJH oversaw genotyping. SH and HDD wrote the manuscript. All authors have read and approved the final manuscript.

Key message

A partial phenotyping strategy based on genomic prediction in multi-environment trials deployed in middle stages of crop breeding programs improves selection response.

Introduction

Genomic selection is a promising tool to assist plant breeding by accelerating selection gain per unit time (Crossa et al. 2017; Endelman et al. 2014; Slater et al. 2016; Voss-Fels et al. 2019). In wheat breeding programs, there is a consensus that genomic selection should be applied in the early stages as phenotyping intensity during this period is low, especially for grain yield and hard to measure traits (Endelman et al. 2014; He et al. 2016). However, this genomic selection strategy depends on an independent and robust reference population, normally consisting of historical data collected across several years (Dawson et al. 2013; Jarquin et al. 2016; Rutkoski et al. 2015).

Another way to deploy genomic selection in breeding is through phenotype imputation (Hori et al. 2016), which does not require an independent reference population. In the middle stages
of breeding programs (e.g. sometimes referred to as stages 1 or 2), wheat lines are regularly phenotyped in only a few environments. Increasing the number of testing environments during these stages with genomic selection could markedly boost selection accuracy, compared to the advanced stages where most selection candidates are intensively tested in many environments (He et al. 2016). However, budget and seed availability constraints makes complete phenotyping of all selection candidates in many environments impractical earlier in the breeding program. Nevertheless, the phenotype imputation scheme proposed by (Hori et al. 2016) suggests that lines do not need to be tested in each environment, i.e. a partial phenotyping strategy. Instead, the phenotype of untested lines in environments are reliably predicted using methods such as multi-environment genomic prediction approaches based on the remaining observations in tested environments. Consequently, a multi-environment trial (MET) with more testing environments could improve overall selection accuracy.

Traditionally, the correlation between the best linear unbiased estimation (BLUE) of genetic value and the genomic estimated genetic value (GEGV) is used to evaluate genomic prediction accuracy (He et al. 2016; Heslot et al. 2012; Jarquin et al. 2016; Rutkoski et al. 2015). BLUEs are assumed to be the best benchmark of GEGV because they are derived directly from per se performance, which is trusted by plant breeders. However, the true genetic value is unknown and whether BLUE or GEGV is closer to the true genetic value is difficult to establish. Thus, rather than prediction accuracy, the focus could be on the actual breeders’ interest, e.g. the response to selection, which can be inferred from a simulation-based approach (Piepho and Möhring 2007) to directly evaluate the effectiveness of genomic selection. To our knowledge, no study has applied this approach to assess the effectiveness of genomic selection.

Our study utilised an Australian pre-breeding wheat population with complete and orthogonal phenotypic records of grain yield across three years and two sowing times to investigate the potential of the genomics-assisted partial phenotyping strategy to improve selection response in the context of multi-environment trials. We also investigate the relationship among environments and how this affects the effectiveness of the proposed genomics-assisted partial phenotyping strategy.
Materials and methods

Plant materials

The wheat grain yield dataset used in this study was a subset of the data set used in He et al. (2019), which also described details such as location, the number of lines in each data set, phenotype test year, dates of sowing per year, and experimental design. The experiments in this study were based on 189 lines consistently tested from year 2015 to 2017 at two times of sowing (TOS) per year. These lines composed an orthogonal data set with a dimension of 189 lines and six environments.

Phenotypic analysis was implemented for each dataset to derive the repeatability estimate per environment (year-TOS combination) and best linear unbiased estimates (BLUEs) per line in each environment, as described in He et al. (2019).

Genotypic data and correlations between environments

The genotypic data of the 189 lines used in this study was drawn from the genotypic data of 2,412 wheat lines fingerprinted with 41,666 90K single nucleotide polymorphisms (SNP) in He et al. (2019). As the number of genotypes was reduced, SNPs were refiltered by removing those with a minor allele frequency of less than 0.05, which left 32,800 SNP for subsequent analyses. The genetic diversity of the 189 genotypes was inspected based on a cluster analysis using Rogers’ distance (Roger 1972) estimated by the 32,800 SNP. The correlation between environments was estimated by Pearson correlation coefficient between the BLUEs of the 189 genotypes in different environments.

Multi-environment genomic prediction model

A multi-environment genomic prediction model explicitly describing genotype-by-environment interactions was used:

\[ y = 1_{mn} \mu + Z_v v + Z_g g + g v + e \]
where \( m \) is the number of environments, \( n \) is the number of genotypes, \( \mathbf{y} \) is a \( m \times n \) vector of BLUEs of genotypes in each environment, \( \mu \) is the common intercept, \( \mathbf{v} \) is the \( m \)-dimensional vector of environment main effect, \( \mathbf{g} \) is the \( n \)-dimensional vector of additive genetic main effect of genotypes, \( \mathbf{gv} \) is the \( m \times n \) vector of genotype-by-environment interaction effects, \( \mathbf{e} \) is the random residual, \( \mathbf{Z}_v \) is the incidence matrices for \( \mathbf{v} \), \( \mathbf{Z}_g \) is the incidence matrices for \( \mathbf{g} \). We assumed \( \mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}\sigma_v^2) \), \( \mathbf{g} \sim \mathcal{N}(\mathbf{0}, \mathbf{G}\sigma_g^2) \), \( \mathbf{gv} \sim \mathcal{N}(\mathbf{0}, \mathbf{Z}_g\mathbf{G}\mathbf{Z}_g^\prime \odot \mathbf{Z}_v\mathbf{Z}_v^\prime \sigma_{gv}^2) \), and \( \mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}\sigma_e^2) \), where \( \odot \) is the Hadamard product of matrices, \( \sigma_g^2 \), \( \sigma_{gv}^2 \) and \( \sigma_e^2 \) are their variance components, respectively, for genotype, genotype-by-environment interaction effects and random residual. \( \mathbf{G} \) is the genomic relationship matrix proposed by VanRaden (2008) constructed based on 32,800 SNP genotypic profiles. The genomic prediction model was run in R (R Core Team 2016) using the BGLR package (de los Campos and Pérez-Rodríguez 2016). Iteration times were fixed to 30,000 and the first 5,000 times were set as burn-in.

**Partial phenotyping strategy**

We compared the selection response of the complete phenotyping trial in fewer environments with a partial genomic phenotyping strategy in additional environments. In this sense, all possible combinations of three environments out of the total six environments were used as the complete phenotyping trials, which retained total phenotypic values (BLUEs per environment). Phenotypic values in combinations of four, five and six environments (there is just one combination using all six environments) were proportionally masked to create the partial phenotyping trials. The percentage of phenotypic values retained in the 4-, 5- and 6-environment combinations was 75%, 60% and 50% respectively, which made the phenotyping intensity in all 3-, 4-, 5- and 6-environment combinations equivalent. Thus, the number of BLUEs and the amount of phenotype data collected was the same in all scenarios. There were twenty different combinations of three environments out of the total six environments. Each 3-environment combination was extended to three 4- or 5-environment combinations by including one or two environments from the remaining three environments. According to the phenotyping proportions (75%, 60% and 50%) of 4-, 5- and 6-environment combinations, phenotypic values in each 4- and 5-environment combination were randomly masked one
hundred times, and in the 6-environment combination were stochastically masked three hundred times. This resulted in the same replication level (300) for each 3-environment combination and its three extended 4- and 5-environment combinations, as well as the single 6-environment combination. The random masking strategy of phenotypic values was based on cross validation strategy two (CV2) in He et al. (2019). Specifically, in this study, each genotype has six environment-specific BLUEs. We first attempted to randomly mask one BLUE of genotypes in the 4-, 5- and 6-environment combinations to make the phenotyping proportions the same as the 3-environment complete phenotyping trial. If masking one BLUE was insufficient to meet the required phenotyping proportion, another BLUE of genotypes was masked until the required phenotyping proportion was reached.

**Response to selection**

The genomic prediction model, also known as a mixed linear model, can be used to directly estimate the response to selection through a simulation-based approach following Piepho and Möhring (2007). Briefly, the multi-environment genomic prediction model was fitted using phenotypic records of complete phenotyping trial (3-environment combination) and phenotypic records of partial phenotyping trials (4-, 5- and 6-environment combinations). We were mainly interested in the relationship between the true genetic main effect $g$ and its best linear unbiased prediction (BLUP) $\hat{g}$, because the selection was based on the BLUP, while the response of selection was determined by the true values. In fact, the joint distribution of $g$ and $\hat{g}$ is multivariate normal and the corresponding variance-covariance matrix $Ω = \text{var}(\begin{pmatrix} g \\ \hat{g} \end{pmatrix})$ can be derived from the mixed model equations. Then, $Ω$ was eigendecomposed as $Ω = DΛD' = ΓΓ'$, where $D$ is the matrix of eigenvectors and $Λ$ is the diagonal matrix of eigenvalues, $Γ = D\sqrt{Λ}$. The vector combining the true and estimated genetic main effects $w = \begin{pmatrix} g \\ \hat{g} \end{pmatrix}$ could be simulated by $w = Γz$, where $z$ is a 2n-dimensional vector of independent standard normal deviates because $\text{var}(w) = \text{var}(Γz) = Γ\text{var}(z)Γ' = ΓΓ' = Ω$ as desired.

For each 3-environment complete phenotyping trial, the responses to selection under selection ratios ranging from 10% to 90% with a gap of 10% were simulated 10,000 times. In each
simulation run, a subset of $S_q$ genotypes with top $p\%$ ($p$=10-90) of $\mathbf{g}$ was selected. The response to selection of the simulation run ($q^{th}$) was calculated as $R_q = \frac{\sum_{i \in S_q} g_i}{\#(S_q)}$, where $\#(S_q)$ is the size of $S_q$. The average value of the 10,000 runs were finally used as the achieved responses to selection of the complete phenotyping trial, i.e. $R = \frac{\sum_{q=1}^{10000} R_q}{10000}$. The responses to selection of each extended 4-, 5- and 6-environment partial genomic phenotyping trial scenario were simulated in the same manner based on only unmasked phenotypic values. The effectiveness of genomic selection was determined by comparing the achieved selection response between each complete phenotyping trial and its extended different partial phenotyping trials. The difference between the achieved responses of complete and partial phenotyping trials was statistically tested with Student’s t-tests.

Result

Phenotypic data and population structure

The repeatability of each environment was above 0.4, indicating that the phenotypic data was of high quality (Fig. 1a). The distribution of BLUEs in different environments was asymptotically normal (Fig. 1b). Several large families were identified by clustering analysis and linkages existed across families (Supplementary Fig. S1). The Rogers’ distance values between any pair of genotypes ranged from 0.01 to 0.53.

Correlations between environments

Pairwise correlations ranged from -0.35 to 0.84 among the six environments (Fig. 2). Among the 3-environment combinations, five combinations showed all positive pairwise correlations. Each 3-environment combination displayed at least one positive pairwise correlation (Supplementary Table S1).

Inspecting the pairwise correlations within the twenty 3-environment combinations, four groupings became clear: 1) one pair of environments had high positive correlation 0.84, i.e. combinations 1-4; 2) environments where all pairwise correlations were positive, i.e.
combinations 5, 11 and 19; 3) one pair of environments had negative correlations, i.e.
combinations 6-7, 12-13 and 17-18; and 4) two pairs of environments had negative correlations,
i.e. combinations 8-10, 14-16 and 20 (Supplementary Table S1).

Simulated response to selection

Twenty one 4-environment combinations with partial phenotyping applied had statistically
significant higher responses to selection, compared to their equivalent 3-environment
combination with complete phenotyping under each selection ratio, i.e. 10%-90% (Table 1).
For the 5- and 6-environment combinations, this number was twenty three and seven,
respectively (Table 2; Table 3). Comparison of the responses of all 3-environment
combinations and their extended 4-, 5- and 6-environment combinations identified five 3-
environment combinations where the partial phenotyping combinations did not result in a
significantly higher response than the corresponding full 3-environment scenarios
(combinations 1-4, 19) (Supplementary Table S2).

Discussion

Our study investigated the potential of a genomics-assisted partial phenotyping strategy via
simulated selection responses. Partial phenotyping can lead to a similar or greater response and
provides information on genotype performance in more environments, compared to fully
replicated trials. As the level of phenotyping (i.e. the number of observations) was the same in
all scenarios, the advantage of partial phenotyping was achieved with a similar budget. While
families existed in our population, our partial phenotyping strategy tested each genotype in at
least one environment. Consequently, as all genotypes were included in the reference set, the
families did not introduce bias due to relatedness discrepancy to genomic prediction in the
different phenotype masking scenarios.

Simulated response to selection can be used to compare the effectiveness of phenotypic
and genomic selection
Conventional plant breeders make selections based on *per se* plant performance, often in the form of BLUEs and are therefore less familiar with selection on GEGVs. However, whether BLUEs or GEGVs are closer to true genetic value is unknown. In this sense, genomic prediction accuracy, normally denoted as the Pearson correlation between BLUE and GEGV, is affected by the accuracy of both BLUEs and GEGVs. Response to selection, which is the plant breeders’ main interest, can be directly derived through simulation. It can be used to evaluate and compare the effectiveness of phenotypic and genomic selection, which circumvents the dilemma of BLUE versus GEGV accuracy.

To further improve simulations of response to selection, other more complicated genomic prediction approaches such as those that accommodate environmental covariance and heterogeneous residual variance (Burgueño et al. 2012; Cuevas et al. 2017; Cuevas et al. 2018) are worthwhile investigating. For simplicity, the model used in this our study did not contain covariances between environments as the simulation of response to selection would too complex and difficult to interpret. Apart from prediction approaches, further improvement of response to selection could be achieved by optimizing the partial phenotyping design (Jarquin et al. 2020). Our study ensured each line was tested in at least one environment. However, complete phenotyping of a small proportion of lines across all environments could improve the estimation of environmental correlations and subsequently enhance genomic predictability (Jarquin et al. 2020). Therefore, further study is required to understand the impact of different partial phenotyping designs on the estimation of response to selection.

**The benefit of partial phenotyping can be anticipated from correlations between environments**

The correlations between environments in our study included high (e.g. 0.84), moderate (e.g. 0.32, 0.38), low (e.g. 0.04, 0.06) and negative (e.g. -0.28, -0.35), which is representative of the types of environments encountered in plant breeding. These four groupings of 3-environment combinations are illustrated in Table 4 and can be used to understand when partial phenotyping can be beneficial.
Group 1 had a highly positive correlation (0.84) between environments and the partial phenotyping strategy did not result in additional selection response, regardless of the number of expansion environments added (Tables 1-4).

In group 2, all pairwise correlations were positive and when the extended environment was highly positively correlated (0.84) with any of the complete phenotyping environments, the partial phenotyping strategy was always superior (Table 1; Supplementary Table S1). However, this superiority was not maintained when additional environment(s) were included that were only poorly correlated with the complete phenotyping environments (Table 2; Table 3; Supplementary Table S1). As there was no expansion environment with a high positive correlation (0.84) with the complete phenotyping environments in combinations 1-4, it was not possible to determine if adding such a highly positively correlated expansion environment would be beneficial or not. It is therefore possible the efficacy of partial phenotyping is actually very similar in groups 1 and 2.

Group 3 had two pairs of environments with a positive correlation and one pair with a negative correlation. Here, the partial phenotyping strategy consistently resulted in an additional selection response when the expansion environment was highly positively correlated (0.84) or even when several expansion environments were moderately positively correlated with the complete phenotyping environments (Tables 1; Table 2; Table 4). This suggests that the robustness of group 3 is less than groups 1 and 2, and the superiority of including two expansion environments in group 3 depends on the relationship between the two expansion environments.

In combination 17-18, no expansion environment was highly positively correlated with any of the complete phenotyping environments. However, two expansion environments were highly correlated (0.84), i.e. Year2015_TOS1 and Year2015_TOS3, and each was moderately positively correlated with one of the complete phenotyping environments, which made the partial phenotyping strategy superior (Table 2). In contrast, their per se 4-environment partial phenotyping scenario did not show superiority (Table 1).

For group 4, where one pair of environments had a positive correlation and two pairs a negative correlation, i.e. combinations 8-10, 14-16 and 20, partial phenotyping resulted in a greater
response when one expansion environment was highly correlated (0.84) or all expansion environments had moderate positive correlations with the complete phenotyping environments (Table 1). In some cases, such as combination 16 and 20, even one extended environment with a moderate positive correlation with the complete phenotyping environments was superior (Table 1). This suggest that when environments are dissimilar the partial phenotyping strategy is particularly useful; a finding corroborated by the largest number of superior 5- and 6-environment combinations in group 4 (Table 2; Table 3).

Breeders are advised to consider the expected phenotypic correlation between environments when deciding whether genomics-assisted partial phenotyping is of value. As shown in Table 4, when the environments projected for complete phenotyping contain a highly positive correlation, the partial phenotyping strategy does not increase selection response. For any other combination of complete phenotyping environments, adding one expansion environment that is positively highly correlated with any of the complete phenotyping environments will always be beneficial. When most complete phenotyping environments are negatively correlated, including more (≤3) expansion environments also consistently improved the response as long as one positive highly correlated expansion environment was added. It is worth noting that while adding one highly positively correlated expansion environment was of benefit, breeders could choose this environment for complete phenotyping if some prior knowledge was available, which would revert the combination to group 1. Nevertheless, adding positive correlation partial phenotyping scenarios was generally of benefit (group 4, Table 1). However, in practice, breeders tend to choose environments that are distinct to select germplasm that are widely adapted.

Finally, although the budgets of a partial phenotyping strategy with different number of expansion environments are theoretically identical, the actual cost would rise if the number of environments was increased, regardless of size. Hence, breeders should assess the practicality of the genomics-assisted partial phenotyping strategy based on both the relationship between testing environments and complexity of breeding program deployment.

Conclusion
Our study demonstrated a genomics-assisted partial phenotyping strategy can improve selection effectiveness for crop breeding, especially at the middle stages of a breeding program when multi-environment trials are not feasible due to cost. The partial phenotyping strategy was optimal when most of the complete phenotyping environments were negatively correlated and at least one of the extension environments was positively highly correlated with any of the complete phenotyping environment.
Tables

Table 1 3-environment combinations with complete phenotypic values showing statistically significant (P<0.05) lower response to selection than their extended 4-environment combinations using genomics-assisted partial phenotyping strategy under each selection ratio (10%-90%)

| Combination   | Environments in the 3-environment combination                                      | Expansion Environment |
|---------------|-----------------------------------------------------------------------------------|-----------------------|
| Combination 5 | Year2015_TOS1, Year2016_TOS1, Year2016_TOS3                                      | Year2015_TOS3         |
| Combination 6 | Year2015_TOS1, Year2016_TOS1, Year2017_TOS1                                      | Year2015_TOS3         |
| Combination 7 | Year2015_TOS1, Year2016_TOS1, Year2017_TOS3                                      | Year2015_TOS3         |
| Combination 8 | Year2015_TOS1, Year2016_TOS3, Year2017_TOS1                                      | Year2015_TOS3, Year2016_TOS1 |
| Combination 9 | Year2015_TOS1, Year2016_TOS3, Year2017_TOS3                                      | Year2015_TOS3, Year2016_TOS1 |
| Combination 10| Year2015_TOS1, Year2017_TOS1, Year2017_TOS3                                      | Year2015_TOS3, Year2016_TOS1 |
| Combination 11| Year2015_TOS3, Year2016_TOS1, Year2016_TOS3                                      | Year2015_TOS1         |
| Combination 12| Year2015_TOS3, Year2016_TOS1, Year2017_TOS1                                      | Year2015_TOS1         |
| Combination 13| Year2015_TOS3, Year2016_TOS1, Year2017_TOS3                                      | Year2015_TOS1         |
| Combination 14| Year2015_TOS3, Year2016_TOS3, Year2017_TOS1                                      | Year2015_TOS1         |
| Combination 15| Year2015_TOS3, Year2016_TOS3, Year2017_TOS3                                      | Year2015_TOS1         |
| Combination 16| Year2015_TOS3, Year2017_TOS1, Year2017_TOS3                                      | Year2015_TOS1         |
| Combination 20| Year2016_TOS3, Year2017_TOS1, Year2017_TOS3                                      | Year2015_TOS1         |

333
Table 2 3-environment combinations with complete phenotypic values showing statistically significant (P<0.05) lower response to selection than their extended 5-environment combinations using genomics-assisted partial phenotyping strategy under each selection ratio (10%-90%)

| Combination 6 | Environments in the 3-environment combination | Expansion Environments |
|---------------|-----------------------------------------------|------------------------|
|               | Year2015_TOS1, Year2016_TOS1, Year2017_TOS1   | Year2015_TOS3, Year2016_TOS3 |
| Combination 7 | Year2015_TOS1, Year2016_TOS1, Year2017_TOS3  | Year2015_TOS3, Year2016_TOS3 |
| Combination 8 | Year2015_TOS1, Year2016_TOS3, Year2017_TOS1  | Year2015_TOS3, Year2017_TOS3 |
| Combination 9 | Year2015_TOS1, Year2016_TOS3, Year2017_TOS3  | Year2015_TOS3, Year2017_TOS3 |
| Combination 10| Year2015_TOS1, Year2017_TOS1, Year2017_TOS3 | Year2015_TOS3, Year2017_TOS3 |
| Combination 12| Year2015_TOS3, Year2016_TOS1, Year2017_TOS1 | Year2015_TOS1, Year2016_TOS3 |
| Combination 13| Year2015_TOS3, Year2016_TOS1, Year2017_TOS3 | Year2015_TOS1, Year2016_TOS3 |
| Combination 14| Year2015_TOS3, Year2016_TOS3, Year2017_TOS1 | Year2015_TOS1, Year2016_TOS3 |
| Combination 15| Year2015_TOS3, Year2016_TOS3, Year2017_TOS3 | Year2015_TOS1, Year2017_TOS3 |
| Combination 16| Year2015_TOS3, Year2017_TOS1, Year2017_TOS3 | Year2015_TOS1, Year2017_TOS3 |
| Combination 17| Year2016_TOS1, Year2016_TOS3, Year2017_TOS1 | Year2015_TOS1, Year2015_TOS3 |
| Combination 18| Year2016_TOS1, Year2016_TOS3, Year2017_TOS3 | Year2015_TOS1, Year2015_TOS3 |
| Combination 20| Year2016_TOS1, Year2017_TOS1, Year2017_TOS3 | Year2015_TOS1, Year2015_TOS3 |
Table 3 3-environment combinations with complete phenotypic values showing statistically significant (P<0.05) lower response to selection than using total six environments with genomics-assisted partial phenotyping strategy under each selection ratio (10%-90%)

| Combination | Environments in the 3-environment combination | Expansion Environments |
|-------------|-----------------------------------------------|------------------------|
| 8           | Year2015_TOS1, Year2016_TOS3, Year2017_TOS1   | Year2015_TOS3, Year2016_TOS1, Year2017_TOS3 |
| 9           | Year2015_TOS1, Year2016_TOS3, Year2017_TOS3   | Year2015_TOS3, Year2016_TOS1, Year2017_TOS1 |
| 10          | Year2015_TOS1, Year2017_TOS1, Year2017_TOS3   | Year2015_TOS3, Year2016_TOS1, Year2016_TOS3 |
| 14          | Year2015_TOS3, Year2016_TOS3, Year2017_TOS1   | Year2015_TOS1, Year2016_TOS1, Year2017_TOS3 |
| 15          | Year2015_TOS3, Year2016_TOS3, Year2017_TOS3   | Year2015_TOS1, Year2016_TOS1, Year2017_TOS3 |
| 16          | Year2015_TOS3, Year2017_TOS1, Year2017_TOS3   | Year2015_TOS1, Year2016_TOS1, Year2016_TOS3 |
| 20          | Year2016_TOS3, Year2017_TOS1, Year2017_TOS3   | Year2015_TOS1, Year2015_TOS3, Year2016_TOS1 |
Table 4 Grouping of 3-environment combinations according to their utility of genomics-assisted partial phenotyping strategy over complete phenotyping

| Group | Complete Phenotyping Three Environment Combinations | Genomic Partial Phenotyping Better? |
|-------|-----------------------------------------------------|-----------------------------------|
|       | Plus 1 partial environment                           | Plus 2 partial environments       | Plus 3 partial environments |
| 1     | One highly positive correlation                      | No                                | No                          |
|       | 1, 2, 3, 4                                          |                                    |                             |
| 2     | All correlations positive                            | Yes, when additional environment was positively highly correlated with the complete phenotyping environment | No                          | No                          |
|       | 5, 11, 19                                           |                                    |                             |
| 3     | One negative correlation                             | Yes, when additional environment was positively highly correlated with the complete phenotyping environment | Yes, when additional environments were positively highly or moderately correlated with the complete phenotyping environment, where the two moderately correlated environments need to be highly correlated | No                          |
|       | 6, 7, 12, 13, 17, 18                                |                                    |                             |
| 4     | Two negative correlations                            | Yes, when additional environment was positively highly correlated with any or positively correlated with all complete phenotyping environments | Yes, when one additional environment was positively highly correlated with the complete phenotyping environment | Yes, when one additional environment was positively highly correlated with the complete phenotyping environment |
Figure captions

**Fig. 1** (a) Repeatability of grain yield in each environment. The highest and lowest repeatability of specific environments evaluated in different datasets are shown in two grayscales; (b) Distribution of best linear unbiased estimate (BLUE) of genotypes in different environments

**Fig. 2** Pairwise correlation between environments

Reference

Burgueño 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 Science 52:707-719

Crossa J, Perezrodriguez P, Cuevas J, Montesinoslopez OA, Jarquin D, Campos GDL, Burgueno J, Gonzalecamucho JM, Perezelizalde S, Beyene Y (2017) Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. Trends in Plant Science 22:961-975

Cuevas J, Crossa J, Montesinos-López OA, Burgueño J, Pérez-Rodríguez P, de los Campos G (2017) Bayesian genomic prediction with genotype× environment interaction kernel models. G3: Genes, Genomes, Genetics 7:41-53

Cuevas J, Granato I, Fritsche-Neto R, Montesinos-Lopez OA, Burgueño J, e Sousa MB, Crossa J (2018) Genomic-enabled prediction Kernel models with random intercepts for multi-environment trials. G3: Genes, Genomes, Genetics 8:1347-1365

Dawson JC, Endelman JB, Heslot N, Crossa J, Poland J, Dreisigacker S, Manès Y, Sorrells ME, Jannink J-L (2013) The use of unbalanced historical data for genomic selection in an international wheat breeding program. Field Crops Research 154:12-22

He S, Schultess AW, Mirdita V, Zhao Y, Korzun V, Bothe R, Ebmeyer E, Reif JC, Jiang Y (2016) Genomic selection in a commercial winter wheat population. Theoretical and applied genetics 129:641-651
He S, Thistlethwaite R, Forrest K, Shi F, Hayden MJ, Trethowan R, Daetwyler HD (2019) Extension of a haplotype-based genomic prediction model to manage multi-environment wheat data using environmental covariates. Theoretical and Applied Genetics 132:3143-3154

Heslot N, Yang H-P, Sorrells ME, Jannink J-L (2012) Genomic Selection in Plant Breeding: A Comparison of Models. Crop Science 52:146-160

Hori T, Montcho D, Agbangla C, Ebana K, Futakuchi K, Iwata H (2016) Multi-task Gaussian process for imputing missing data in multi-trait and multi-environment trials. Theoretical and Applied Genetics 129:2101-2115

Jarquin D, Howard R, Crossa J, Beyene Y, Gowda M, Martini JWR, Covarrubias Pazaran G, Burgueño J, Pacheco A, Grondona M, Wimmer V, Prasanna BM (2020) Genomic Prediction Enhanced Sparse Testing for Multi-environment Trials. G3: Genes|Genomes|Genetics 10:2725

Jarquin D, Specht J, Lorenz A (2016) Prospects of Genomic Prediction in the USDA Soybean Germplasm Collection: Historical Data Creates Robust Models for Enhancing Selection of Accessions. G3: Genes|Genomes|Genetics 6:2329-2341

Piepho HP, Möhring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881-1888

R Core Team (2016) R: A language and environment for statistical computing, Vienna, Austria. https://www.R-project.org/

Roger J (1972) Measure of genetic similarity and genetic distance. Studies in genetics VII. University of Texas publication 7213:145-153

Rutkoski J, Singh RP, Huerta-Espino J, Bhavani S, Poland J, Jannink JL, Sorrells ME (2015) Efficient Use of Historical Data for Genomic Selection: A Case Study of Stem Rust Resistance in Wheat. The Plant Genome 8:plantgenome2014.2009.0046

Slater AT, Cogan NO, Forster JW, Hayes BJ, Daetwyler HD (2016) Improving Genetic Gain with Genomic Selection in Autotetraploid Potato. Plant Genome 9

VanRaden PM (2008) Efficient methods to compute genomic predictions. Journal of dairy science 91:4414-4423

Voss-Fels KP, Cooper M, Hayes BJ (2019) Accelerating crop genetic gains with genomic selection. Theoretical and Applied Genetics 132:669-686