

# **Accounting for Epistasis in Genomic Phenotype Prediction**

Dissertation

to attain the degree Doctor of Philosophy (Ph. D.)

of the Faculty of Agricultural Sciences

Georg-August-Universität Göttingen

Submitted by

**Elaheh Vojgani**

born on the May 10<sup>th</sup>, 1990 in Tehran, Iran

Göttingen, November 2020

1<sup>st</sup> Referee: Prof. Dr. Henner Simianer, Georg-August University of Göttingen

2<sup>nd</sup> Referee: Prof. Dr. Timothy Beissinger, Georg-August University of Göttingen

3<sup>rd</sup> Referee: Prof. Dr. Thomas Kneib, Georg-August University of Göttingen

Date of oral examination: 22. January 2021

# Declaration

1. I hereby declare that this work has not been submitted to other examination authorities in the same or a similar form.  
I further declare that I have not applied for a doctorate at any other university.
2. I hereby declare under oath that this dissertation was prepared independently and without unauthorized assistance.

Göttingen, 20.11.2020

Elaheh Vojgani





# Acknowledgments

“Acknowledgment is the only way to keep the love alive”

Barry Long

First of all, I would like to appreciate my supervisor **Prof. Henner Simianer** for his precious supervision, guidance and encouragement during my PhD and providing me many opportunities for participation in various courses and conferences.

I would like to thank **Prof. Timothy Beissinger** for being my second supervisor and providing me valuable support and supervision specially during the last year of my PhD. I would like to specially thank **Dr. Torsten Pook** for being the member of my thesis committee and providing outstanding support and guidance during different phases of my PhD. I also would like to thank **Prof. Thomas Kneib** for becoming my referee.

Moreover, I would like to thank all my colleagues at the Animal Breeding and Genetics Group of Göttingen for the interaction and the pleasant working atmosphere specially my former colleague **Dr. Johannes W.R. Martini** who had a great impact on my research specially in the first year of my PhD. I further would like to specifically thank **Mrs. Ute Döring** who took care of all the administrative work.

I would like to thank all the MAZE partners especially KWS SAAT SE, Misión Biológica de Galicia, Spanish National Research Council (CSIC) and Technical University of Munich for organizing annual meetings and providing the data of this study.

I would like to appreciate my father **Mohammad Vojgani** and my mother **Ashraf Asheghabadi** as the first teachers of my life who have supported me and guided me during whole my life with overflowing love.

Finally, I would like to appreciate **Alireza Eshtehardiha** for all his precious supports.

## Funding

I gratefully acknowledge the financial support I received during my PhD. With this, I am thankful to the German Federal Ministry of Education and Research (BMBF) for funding my project within the scope of the funding initiative “Plant Breeding Research for the Bioeconomy” (MAZE – “Accessing the genomic and functional diversity of maize to improve quantitative traits”; Funding ID: 031B0195).



# Contents

Declaration.....	i
Acknowledgments .....	iii
Summary.....	9
Zusammenfassung.....	11
<b>1 Introduction.....</b>	<b>15</b>
1.1 Availability of single nucleotide polymorphism (SNP) data.....	15
1.2 Genomic breeding value prediction.....	16
1.3 Genomic prediction of phenotypes .....	17
1.4 Mixed models.....	19
1.5 Additive genomic prediction models .....	20
1.6 Epistasis genomic prediction models.....	23
1.7 Univariate and multivariate genomic prediction models .....	26
<b>2 Phenotype prediction under epistasis.....</b>	<b>29</b>
2.1 Abstract.....	30
2.2 Introduction.....	30
2.3 Methods.....	32
2.3.1 Genomic Best Linear Unbiased Prediction (GBLUP) .....	33
2.3.2 Epistatic Random Regression BLUP (ERRBLUP) .....	33
2.3.3 selective Epistatic Random Regression BLUP (sERRBLUP) .....	35
2.4 Notes .....	41
<b>3 Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments .....</b>	<b>45</b>
3.1 Key Message .....	46
3.2 Abstract.....	46
3.3 Introduction.....	46
3.4 Materials and Methods.....	49
3.4.1 Data used for analysis.....	49
3.4.2 Quality control, coding and imputing .....	50
3.4.3 Linkage disequilibrium pruning.....	50
3.4.4 Univariate statistical models for phenotype prediction .....	51

3.4.5	Bivariate statistical models for phenotype prediction .....	51
3.4.6	Model 1: Genomic Best Linear Unbiased Prediction (GBLUP) .....	53
3.4.7	Model 2: Epistatic Random Regression BLUP (ERRBLUP).....	53
3.4.8	Model 3: selective Epistatic Random Regression BLUP (sERRBLUP) .....	54
3.4.9	Assessment of predictive ability via 5-fold random cross validation with 5 replicates .....	55
3.4.10	Assessment of GBLUP, ERRBLUP and sERRBLUP predictive abilities.....	55
3.4.11	Use of multiple environments jointly.....	56
3.4.12	Estimation of variance and covariance components.....	57
3.5	Results .....	57
3.6	Discussion.....	67
3.7	Supplementary material .....	71
3.7.1	Supplementary Tables .....	71
3.7.2	Supplementary Figures.....	81
<b>4</b>	<b>Bivariate genomic prediction of phenotypes by selecting epistatic interactions across years based on haplotype blocks and pruned sets of SNPs .....</b>	<b>107</b>
4.1	Key Message .....	108
4.2	Abstract.....	108
4.3	Introduction.....	108
4.4	Materials and Methods.....	110
4.4.1	Data used for analysis .....	110
4.4.2	Statistical models for phenotype prediction .....	113
4.4.3	Assessment of genomic prediction models .....	113
4.4.4	Variance component estimation.....	114
4.4.5	Genomic correlation estimation.....	114
4.5	Results .....	114
4.6	Discussion.....	122
4.7	Supplementary material .....	125
4.7.1	Supplementary Tables .....	125
4.7.2	Supplementary Figures.....	134
<b>5</b>	<b>Discussion .....</b>	<b>157</b>
5.1	Different methods for pairwise SNP interaction selection .....	157
5.2	Linkage disequilibrium based SNP pruning and its influence on genomic prediction model's performance .....	160
5.3	Comparison of cross validation scenarios in bivariate model framework.....	164

5.3.1	Evaluating cross validation scenario in bivariate models.....	167
5.4	The required genotype overlap in bivariate model framework.....	167
5.4.1	Genomic and phenotypic correlation impacts on bivariate model’s predictive ability under the high level of genotype overlap .....	170
5.5	Potential future research topics .....	174
5.6	Outlooks and Conclusion.....	174
5.6.1	Outlook for epistasis models .....	174
5.6.2	Outlook for EpiGP R-package.....	175
5.6.3	Outlook for influential factors on the model’s predictive ability.....	175
5.6.4	Concluding remarks .....	176
6	References.....	177
7	List of Figures.....	191
8	List of Tables .....	195
9	List of Supplementary Figures .....	197
10	List of Supplementary Table.....	203
	Curriculum Vitae .....	207



# Summary

Wide availability of genomic data has had a considerable impact on plant and animal breeding programs which enables the study of genotypes and their relationships with phenotypes. Improving genomic prediction accuracy is of great interest in plant and animal breeding for selection purposes. In quantitative genetics, the standard models account for additive genetic effects while epistasis effects have been widely ignored due to their computational load. In this thesis, the significance of incorporating epistasis interactions in the genomic prediction of phenotypes are investigated.

Chapter 1 presents a general introduction to the significant effects of genomic data specifically in animal and plant studies in both breeding value prediction and genomic prediction of phenotypes. Then different additive and epistasis models are reviewed and the challenges they encounter when considering epistasis are detailed. Finally, the univariate and multivariate statistical settings for genomic prediction of phenotypes are compared in their predictive abilities. The main chapters of this thesis are the three corresponding articles presented in Chapters 2, 3, and 4.

In Chapter 2, “Phenotype Prediction under Epistasis” is discussed through developed epistatic models defined as Epistatic Random Regression BLUP (ERRBLUP) and selective Epistatic Random Regression BLUP (sERRBLUP) implemented in the developed R-package named “EpiGP”, which is able to process large scale genomic data in a computationally efficient manner. ERRBLUP is considered as a full epistatic model which incorporates all pairwise SNP interactions, while sERRBLUP is a selective epistatic model which incorporates a subset of pairwise SNP interactions selected according to their absolute effect sizes or the effect variances. These models are compared to GBLUP as an additive model in univariate statistical framework with the genotypes from the publicly available wheat dataset and respective simulated phenotypes. The results indicate that sERRBLUP leads to a considerable increase in predictive ability compared to ERRBLUP and GBLUP when the optimum proportion of SNP interactions is maintained in the model.

GBLUP, ERRBLUP and sERRBLUP are developed in bivariate statistical setting in Chapter 3 in the article “Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments” where two environments are modeled as two separate traits in multi-trait model. In Chapter 3, GBLUP, ERRBLUP and sERRBLUP are compared in both univariate and bivariate statistical frameworks in maize dataset derived from 910 doubled haploid lines of two European landraces Kemater Landmais Gelb and Petkuser Ferdinand Rot grown in six locations in Germany and Spain in the year 2017 for eight phenotypic traits. In the maize dataset, pairwise SNP interaction selection based on effect variances is considered as the selection criteria due to its robustness compared to selection based on effects sizes in sERRBLUP model. Our results indicate the superiority of the sERRBLUP over GBLUP and ERRBLUP in both univariate and bivariate statistical settings when selecting the subset of interactions with the

highest effect variances. The comparison between univariate and bivariate models also reveals the superior predictive abilities of bivariate models over univariate models.

In chapter 4, we analyze the utility of haplotype blocks in contrast to LD-pruning in the article "Bivariate genomic prediction of phenotypes by selecting epistatic interactions across years based on haplotype blocks and pruned sets of SNPs". For this, we consider a model in which observations of the same trait in different years (2017 & 2018) are considered as two separate traits in a multivariate model. This is done in the 873 doubled haploid lines in the respective maize dataset in four locations in Germany and Spain in both years 2017 and 2018. The results are in line with our finding from the bivariate model when considering two environments as the two separate traits indicating the superiority of bivariate sERRBLUP over GBLUP in most cases. Overall, the prediction accuracies obtained by LD-pruning and haplotype blocks are similar. However, the use of haplotype blocks can significantly reduce the computation time. Moreover, we explore genomic correlation, phenotypic correlation and trait's heritability as three influential factors on bivariate model's prediction accuracy. The results illustrate the significance of genomic correlation between growing seasons in the bivariate model's prediction accuracy. Phenotypic correlation and heritability of the traits also affect this increase in predictive ability to some extent.

In this thesis, the main studied trait in the maize dataset is plant height at V4 growth stage (PH\_V4) and the results for series of other phenotypic traits are presented in supplementary material in Chapter 3 and Chapter 4.

Finally, the general discussion is presented in Chapter 5 in which our proposed selection method in sERRBLUP model is compared with other methods of variable selection indicating the superiority of our proposed selection method in sERRBLUP. Furthermore, the influential factors on the predictive ability of the genomic prediction models are investigated. In this regard, linkage disequilibrium based SNP pruning as a potential approach to reduce the number of SNPs in order to make the application of epistasis models feasible is shown to result in predictive abilities as good as or better than those obtained from utilizing full panel of SNPs. Moreover, the cross validation scenario in bivariate statistical settings is shown to be an important factor affecting the bivariate models' predictive abilities. In addition, the level of genotype overlap is found to be significantly correlated with the increase in the bivariate model's predictive ability under the cross validation scenario which leads to higher predictive ability. Under the assumption of high level of genotype overlap, the genomic correlation is significantly correlated to the bivariate models' predictive abilities for highly heritable traits. Phenotypic correlation is also shown to be an influential factor in this context. Finally, incorporating transcriptomic data into epistasis genomic prediction models, incorporating weather data into epistasis multi-trait genomic prediction models and exploring single-trait and multi-trait epistasis GWAS are proposed as the potential field of research and further investigations for future studies in the context of epistasis models.



# Zusammenfassung

## **Titel der Arbeit: Berücksichtigung von Epistasie in der genomischen Phänotypvorhersage.**

Die breite Verfügbarkeit genomischer Daten hat einen erheblichen Einfluss auf Pflanzen- und Tierzuchtprogramme, da hierdurch Untersuchung von Genotypen und deren Beziehungen zu Phänotypen ermöglicht wurden. Die Verbesserung der Genauigkeit genomischer Zuchtwerte ist in der Pflanzen- und Tierzucht zu Selektionszwecken von großem Interesse. In der quantitativen Genetik berücksichtigen die Standardmodelle additive genetische Effekte, während epistatische Effekte aufgrund des damit verbundenen rechentechnischen Aufwands meist ignoriert werden. In dieser Arbeit wird die Bedeutung der Einbeziehung von Epistasie-Interaktionen in die genomische Vorhersage von Phänotypen untersucht.

Kapitel 1 enthält eine allgemeine Einführung in die Nutzung genomischer Daten speziell in Tier- und Pflanzenstudien sowohl für die Zuchtwertschätzung als auch für die genomische Vorhersage von Phänotypen. Anschließend werden verschiedene rein additive und epistatische Modelle zur Zuchtwertschätzung dargestellt und die Herausforderungen bei der Berücksichtigung von epistatischen Effekten werden detailliert beschrieben. Schließlich werden univariate und multivariate Modelle für die genomische Vorhersage von Phänotypen aufgrund ihrer jeweiligen Genauigkeiten miteinander verglichen. Die Hauptkapitel dieser Arbeit sind die drei separaten wissenschaftlichen Artikel, die in den Kapiteln 2, 3 und 4 vorgestellt werden.

In Kapitel 2 wird die Phänotypvorhersage unter Epistasie anhand neu entwickelter epistatischer Modelle diskutiert, die als "Epistatic Random Regression BLUP" (ERRBLUP) und "selective Epistatic Random Regression BLUP" (sERRBLUP) bezeichnet werden. Alle Methoden wurden im assoziierten R-Paket "EpiGP" implementiert, das in der Lage ist, große Mengen genomischer Daten auf rechnerisch effiziente Weise zu verarbeiten. ERRBLUP ist ein vollständig epistatisches Modell, das alle paarweisen SNP-Interaktionen enthält, während sERRBLUP ein selektives epistatisches Modell ist, das eine Untermenge von paarweisen SNP-Interaktionen enthält, die nach ihren absoluten Effektgrößen oder den Effektvarianzen ausgewählt werden. Diese Modelle werden mit dem additiven GBLUP-Modell in einem univariaten statistischen Rahmen miteinander verglichen. Hierfür wurde der öffentlich verfügbare Weizendatensatz aus dem R-Paket BGLR mit simulierten Phänotypen genutzt. Die Ergebnisse deuten darauf hin, dass sERRBLUP im Vergleich zu ERRBLUP und GBLUP zu einer erheblichen Steigerung der Vorhersagefähigkeit führt, wenn der optimale Anteil an SNP-Interaktionen im Modell berücksichtigt wird.

Ähnlich wie GBLUP können auch ERRBLUP und sERRBLUP in einem multivariaten Setting genutzt werden. Hierzu werden die entsprechenden Modelle in einem bivariaten Setting in Kapitel 3 in dem Artikel "Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments" entwickelt, in dem zwei verschiedene Umwelten als zwei getrennte Merkmale im multivariaten Ansatz modelliert werden. In Kapitel 3 werden GBLUP, ERRBLUP und sERRBLUP sowohl im univariaten als auch im bivariaten

statistischen Rahmen in Mais-Datensätzen verglichen, die von 910 doppelhaploiden Linien der beiden europäischen Landrassen Kemater Landmais Gelb und Petkuser Ferdinand Rot erzeugt wurden, welche im Jahr 2017 an sechs Standorten in Deutschland und Spanien angebaut wurden und an denen acht phänotypische Merkmale erfasst wurden. Bei der Anwendung des sERRBLUP-Modells auf den Maisdatensatz erwies sich die Auswahl von SNP-Interaktionen auf Grundlage von Effektvarianzen aufgrund ihrer Robustheit gegenüber der Auswahl auf der Grundlage von Effektgrößen als überlegen. Unsere Ergebnisse zeigen die Überlegenheit von sERRBLUP gegenüber GBLUP und ERRBLUP sowohl in univariaten als auch in bivariaten statistischen Modellen. Der Vergleich zwischen univariaten und bivariaten Modellen zeigt auch die überlegenen prädiktiven Fähigkeiten bivariater Modelle gegenüber univariaten Modellen.

In Kapitel 4 analysieren wir den Nutzen von Haplotypenblöcken anstelle von LD-pruning im Artikel “Bivariate genomic prediction of phenotypes by selecting epistatic interactions across years based on haplotype blocks and pruned sets of SNPs”. Hierzu betrachten wir ein Modell in dem Beobachtungen des gleichen Merkmals in unterschiedlichen Jahren (2017 & 2018) als zwei separate Merkmale in einem multivariaten Modell betrachtet werden. Dies geschieht auf Grundlage von 873 doppelhaploiden Linien des jeweiligen Maisdatensatzes an vier Standorten in Deutschland und Spanien in den Jahren 2017 und 2018. Die Ergebnisse stimmen mit unseren Erkenntnissen aus dem bivariaten Modell in der Anwendung auf zwei Umwelten im gleichen Jahr überein, die in den meisten Fällen eine Überlegenheit des bivariaten sERRBLUP gegenüber GBLUP ergaben. Insgesamt sind die Vorhersagegenauigkeiten, die durch LD-pruning und Haplotypenblöcke erzielt werden, ähnlich. Allerdings kann durch die Nutzung von Haplotypenblöcken die Rechenzeit deutlich stärker reduziert werden. Darüber hinaus untersuchen wir die genomische Korrelation, die phänotypische Korrelation und die Heritabilität des Merkmals als drei Einflussfaktoren auf die Genauigkeit der Vorhersage im bivariaten Modell. Die Ergebnisse betonen die Bedeutung der genomischen Korrelation zwischen den Jahren für die Vorhersagegenauigkeit des bivariaten Modells. Daneben beeinflussen auch die phänotypische Korrelation und die Heritabilität der Merkmale die Zunahme der Vorhersagegenauigkeit bis zu einem gewissen Grad.

In dieser Arbeit ist das wichtigste untersuchte Merkmal im Maisdatensatz die Pflanzenhöhe im Wachstumsstadium V4 (PH\_V4), die Ergebnisse für anderer phänotypischer Merkmale werden im Anhang von Kapitel 3 und Kapitel 4 vorgestellt.

Abschließend wird in der allgemeine Diskussion (Kapitel 5) unsere vorgeschlagene Auswahlmethode im sERRBLUP-Modell mit anderen Methoden der Variablenauswahl verglichen, was die Überlegenheit unserer vorgeschlagenen Auswahlmethode in sERRBLUP nochmals verdeutlicht. Darüber hinaus werden die Einflussfaktoren auf die Vorhersagegenauigkeit der genomischen Vorhersagemodelle untersucht. In dieser Hinsicht hat sich gezeigt, dass eine auf Kopplungsungleichgewicht basierende SNP Auswahl (LD-pruning) zur Verringerung der Anzahl der SNPs eingesetzt werden kann, um die Anwendung von Epistasiemodellen rechentechnisch zu ermöglichen. Die damit erhaltenen Vorhersagegenauigkeiten sind vergleichbar oder sogar besser als die, die durch die Verwendung eines vollständigen Panels von SNPs erreicht werden. Darüber

hinaus wird gezeigt, dass das Kreuzvalidierungsszenario in bivariaten statistischen Ansätzen ein wichtiger Faktor für die Abschätzung der Vorhersagefähigkeiten in bivariaten Modellen ist. Weiterhin wird gezeigt, dass der Grad der Überlappungen der in den beiden Umwelten beobachteten Linien signifikant mit der Zunahme der Vorhersagefähigkeit des bivariaten Modells unter dem Kreuzvalidierungsszenario korreliert ist. Unter der Annahme eines hohen Grades an Überlappung ist die genomische Korrelation signifikant mit der Güte der Vorhersage des bivariaten Modells für Merkmale mit hoher Heritabilität korreliert. Auch die phänotypische Korrelation erweist sich in diesem Zusammenhang als relevanter Faktor. Schließlich hinaus werden die Einbeziehung von Transkriptomdaten und Wetterdaten in die Vorhersagemodelle diskutiert und die Erforschung von epistatischen Modellen für GWAS Analysen als potentiell Forschungsbereich vorgeschlagen.



# 1 Introduction

“All models are wrong, but some are useful”

George Edward Pelham Box

Modern technologies pave the way for cheaper and easier methods of data collection. Now it is data scientists' turn to extract the information out of the valuable data resources. In this regard, genomic data is drawing the attention of human biologists, animal and plant specialists and data scientists. Therefore, efficient statistical models are required to investigate and compile the information hidden in of these huge amounts of data. Fortunately, developments in computing powers and the continuing development of different statistical software packages make the application of statistical models feasible for such huge datasets.

In the context of plant and animal breeding, generation of dense molecular markers has replaced the traditional methods based on the pedigree information. Several statistical models have been developed over the last decades based on the genomic data in order to increase the statistical model's accuracy and reliability.

In this chapter, the significance of wide availability of genomic data in human, animal and plant studies is reviewed with a focus on how utilization of these genomic data in plant breeding can potentially lead to an increase in the accuracy of breeding value prediction and phenotype prediction. In this context, additive and epistasis models are reviewed and univariate and multivariate statistical settings are compared in their prediction accuracy.

## 1.1 Availability of single nucleotide polymorphism (SNP) data

Genotyping technology took a huge step in its development at the beginning of this century which resulted in the wide availability of marker data. New technologies made it possible to obtain thousands of SNP as the most common type of genetic variation representing a difference in a single nucleotide in an individual's genome at a reasonable cost by using SNP array with high throughput genotyping platform. SNPs are considered to play a major role in the induction of phenotypic variations in human, animals and plants (Huq *et al.*, 2016). In the context of plant breeding, the discovery and application of SNPs helps to increase the knowledge about genetic diversity and gain a better understanding about crops developments (Morgil, 2020). In this regard, Illumina (<https://www.illumina.com>) and Affymetrix (<http://www.affymetrix.com>) are two world-wide companies which provide SNP arrays. To illustrate this, Barley 9K illumine array (Comadran *et al.*, 2012), Maize 50K illumine array (Ganal *et al.*, 2011), 600K Affymetrix® Axiom® Maize

Genotyping Array (Unterseer *et al.*, 2014), 600K Affymetrix® Axiom® SNP genotyping array for chicken (Kranis *et al.*, 2013) are some of low and high density genotyping arrays.

## 1.2 Genomic breeding value prediction

Genomic breeding value prediction is an important tool in improving the livestock species genetics based on the relationship between the individuals. This enables estimating the proportion of phenotypic variance which is heritable (Calus, 2010). In fact, the animals' and plants' important economic quantitative traits have been selected traditionally from their phenotypic records and their relatives' phenotypic records in order to be the parents of the next generation. These phenotypic data were used to estimate breeding values through best linear unbiased prediction (BLUP, Henderson, 1984) which was initially introduced by Henderson in 1975 (Henderson, 1975) resulting in a milestone in breeding models' development. In fact, the application of BLUP in the context of linear mixed model originated from Henderson's work in 1950s (Henderson *et al.*, 1959). In this framework, linear mixed model random terms are assumed to be multivariate normally distributed and the covariance structures are predetermined by measuring the relationship between individuals (Henderson *et al.*, 1959). Therefore, the maximum achievable reliability of breeding value estimation through BLUP was obtained by using pedigree information across many generations and phenotypic data from individuals or their relatives (Henderson, 1975).

Breeding value estimation is based on a covariance matrix describing the (additive) relationship between the individuals in the population (Henderson, 1984). This additive relationship matrix can be constructed from pedigree information which has been collected over multiple generations (Hayes and Goddard, 2008). However, the pedigree information might not be always available or complete in plants and livestock population. Therefore, a reliable alternative approach is required. In 2001, the dense marker set was proposed to predict breeding values which revolutionized animal breeding by replacing the pedigree based prediction of breeding values (Meuwissen *et al.*, 2001). Marker information is used as an alternative to infer relationships instead of pedigree to construct the additive relationship matrix for breeding value prediction (Hayes and Goddard, 2008). A single step approach also has been proposed to construct the relationship matrix based on the combination of both available pedigree and genotypic information (Legarra *et al.*, 2014).

In fact, genomic selection which plays an important role in plant and animal breeding (Burgueño *et al.*, 2012) requires breeding value estimation of selection candidates, and its potential benefits rely on the accuracy of breeding value estimation (Wolc *et al.*, 2011) and reducing the generation interval (Schaeffer, 2006). In this regard, the accuracy of estimated breeding values when using pedigree and high density SNP genotypes was compared. The use of dense marker sets potentially results in more accurate breeding value estimation, since they can capture past relationships which are not in the pedigree (Hayes and Goddard, 2008). This fact led to a rapid development of genomic selection tools during the first decade of 21<sup>st</sup> century (Wolc *et al.*, 2011). In this context, several BLUP models have been proposed utilizing inclusion of marker information (Meuwissen *et al.*,

2001). Genomic best linear unbiased prediction (GBLUP) has been proposed in this regard by constructing an additive genomic relationship matrix (VanRaden, 2008).

Overall, several studies indicate that the reliability of breeding value estimation is enhanced by the usage of marker based relationship matrices in breeding value prediction (Meuwissen *et al.*, 2001; VanRaden, 2007; Hayes and Goddard, 2008). This is due to the fact that quantitative traits as the most prevalent form of traits in plant and animal breeding are usually affected by a high number of quantitative trait loci (QTL) (Meuwissen *et al.*, 2001). Hundreds of QTLs as the gene loci which contribute to the specific trait's variation are assumed to influence most of the quantitative traits in which their location and their contribution to the genetic variation are unknown. However, mapping the QTLs can be done by linkage to the polymorphic marker loci such as molecular polymorphisms (Mackay *et al.*, 2009). Therefore, utilizing marker information from high density SNP genotyping was shown to increase the prediction accuracy of breeding values, which helps selecting the young animals with higher breeding values as the parents of the next generation in addition to shortening the generation intervals as an application of genomic selection (Meuwissen *et al.*, 2001; Daetwyler *et al.*, 2007; Wolc *et al.*, 2011).

### **1.3 Genomic prediction of phenotypes**

The World Summit on Food Security declared that by 2050 an increase of 70 to 100 percent in the level of food production is predicted to be essential for food security. However, increasing world food production is challenging especially due to the increasing population and climate change (Tester and Langridge, 2010; Hickey *et al.*, 2017). The 20<sup>th</sup> century faced hunger as a poverty problem rather than absolute food scarcity (Koning *et al.*, 2008). Therefore, using new technologies to increase food production to meet the food demand for the increasing population is of special importance (A. Montesinos-López *et al.*, 2018). In this context, livestock and crop breeding is one of the efficient ways to increase food production. Plant breeding needs to utilize efficient selection strategies to increase yields in different environments (Tester and Langridge, 2010).

Enhancing genetic gain – defined as the amount of increase in performance which is achieved annually through artificial selection – as one of the important concepts in conventional quantitative genetics and breeding is essential to fill the gap between the growing population food demand and food production (Xu *et al.*, 2017). In this regard, phenotype prediction is the core of genetic improvement in crop breeding. Utilizing genomic information in order to predict new potential genotypes which are superior in yield performance is of special interest for plant breeders. In fact, the phenotypic and genomic information can be used for predicting own phenotypic performance as well as breeding value prediction (Calus, 2010) which is significantly important in breeding programs. Plant breeders' interest in genomic prediction of phenotypes is focused on predicting phenotypic trait performance for different genotypes for selection purposes rather than accurately predicting phenotypic trait for individual genotypes within environments. Therefore, genomic

prediction of phenotypes is required in breeding programs for genomic improvement of complex traits such as yield (Hammer *et al.*, 2019).

In fact, phenotype prediction plays an important role in wide variety of fields of life science from artificial selection programs (Meuwissen *et al.*, 2001) to risk prediction in human medicine (Wray *et al.*, 2007). In plant and animal breeding, phenotype prediction of lines and individuals has been widely developed (Crossa *et al.*, 2010; Daetwyler *et al.*, 2013; de los Campos, Hickey, *et al.*, 2013) as well as human health related traits (Wray *et al.*, 2007; de los Campos *et al.*, 2010). In plant breeding, genotyping costs have decreased in the last 20 years, whereas phenotyping is still one of the bottlenecks in breeding programs which requires optimization or minimization of costs (Akdemir and Isidro-Sánchez, 2019). Therefore, maximization of genomic prediction accuracy not only will help genomic improvement in plant breeding but will also lead to reduction in phenotyping costs which gives the genomic prediction of phenotypes considerable importance in breeding programs.

Understanding how genetic variation causes phenotypic variations of quantitative traits is a major challenge of contemporary biology. Therefore, the development of a genomic prediction model which can capture the genetic variation of phenotypes results in more accurate phenotype prediction. In this regard, the variation of observed phenotypes has shown to be caused by many loci and most single SNPs only have a small effect on phenotype. Consequently, phenotypic variation should be captured by combining the effects of multiple SNPs; and the proper models which help to find the set of SNPs best explaining and predicting the phenotypic variation have to be selected (Lee *et al.*, 2008). In this context, wide availability of dense markers brings new opportunities and challenges on how to include this information in the statistical phenotype prediction models (de los Campos *et al.*, 2009) to model the relationship between genome wide marker data and phenotypes (Lee *et al.*, 2008).

Several studies have been conducted based on whole genome prediction methods in which their performance has been positively evaluated (Wray *et al.*, 2007; de los Campos *et al.*, 2010; Ober *et al.*, 2012; de los Campos, Hickey, *et al.*, 2013; Liu *et al.*, 2019). One of the first applications of genomic prediction of phenotypes by whole genome sequence data was for starvation stress resistance and startle response in *Drosophila melanogaster* using approximately 2.5 million SNPs (Ober *et al.*, 2012). This was done by GBLUP using a genomic relationship matrix constructed from SNP data indicating a potential benefit of sequence based phenotype prediction (Ober *et al.*, 2012). Using all markers simultaneously rather than considering only significant marker-trait associations for genomic prediction of phenotypes (Windhausen *et al.*, 2012) was successfully implemented in Holstein and Jersey dairy cattle (Goddard and Hayes, 2009; Hayes, Bowman, *et al.*, 2009; Habier *et al.*, 2010), plant breeding (Lorenzana and Bernardo, 2009; Windhausen *et al.*, 2012) and human diseases (Daetwyler *et al.*, 2008) which indicates an improvement in prediction accuracy. Two examples of this approach are: Accurate prediction of maize grain yield based on whole genome regression models (Millet *et al.*, 2019) and prediction of unobserved phenotypes based on all genomic information (SNP) across the whole genome simultaneously reported to be significantly better than prediction based on the close relatives' phenotypes (Lee *et al.*, 2008).



In plant and animal breeding, genomic selection utilizes phenotype prediction of quantitative traits based on all markers to enhance phenotypic traits in breeding populations which increase breeding efficiency (Liu *et al.*, 2019). Accurately predicting complex phenotypic traits based on genotype data will revolutionize plant and animal breeding and will also result in personalized medicine (Ober *et al.*, 2012). Therefore, the importance of genomic prediction of phenotypes as a tool in genomic selection in animal and plant breeding (Meuwissen *et al.*, 2001), disease risk prediction (Vazquez *et al.*, 2012; de los Campos, Hickey, *et al.*, 2013; Wray *et al.*, 2013) and personalized medicine (Burke and Psaty, 2007; Bielinski *et al.*, 2014) is undeniable.

The accuracy of genomic prediction of phenotypes can be affected by many factors. Heritability of the desired trait is one of the important factors which influences the accuracy of phenotype prediction. Traits with low heritabilities will never have accurate prediction of unobserved phenotypes from genetic data even if genetic effect is predicated with 100 percent accuracy (Lee *et al.*, 2008). The genetic architecture of the complex trait defined as a genotype-phenotype relationships that includes the loci contributing to phenotypic variation is another factor which could be influential in the accuracy of genomic prediction of phenotypes (Lee *et al.*, 2008; Momen *et al.*, 2018). However, the traits which are relevant for breeding programs have different genetic architectures, most of which remain unknown. Therefore, it is important to search for a prediction model which is robust and stable (Momen *et al.*, 2018). Studies indicate that under additive gene action, parametric prediction models such as GBLUP (VanRaden, 2008; Habier *et al.*, 2013), ridge regression BLUP (rrBLUP) (Meuwissen *et al.*, 2001; Endelman, 2011), Bayesian ridge regression (BRR) (Gianola *et al.*, 2003; Zou and Hastie, 2005; de los Campos *et al.*, 2009), BayesA (Meuwissen *et al.*, 2001; Habier *et al.*, 2011), BayesB (Meuwissen *et al.*, 2001; Habier *et al.*, 2011), and BayesC (Meuwissen *et al.*, 2001; Habier *et al.*, 2011) outperformed non parametric ones and under epistasis gene action non parametric models such as reproducing kernel Hilbert space regression (RKHS) (Gianola *et al.*, 2006; Gianola and van Kaam, 2008; de Los Campos *et al.*, 2010) outperformed parametric models (Momen *et al.*, 2018).

## 1.4 Mixed models

Mixed models are the models which consider both fixed and random effects. Fixed effects are unknown constants and random effects are the values which are drawn from an underlying distribution. This distribution is often considered as a normal distribution with mean zero and an unknown variance. From the statistical point of view, fixed effects will be estimated, while random effects will be predicted. Mixed models make it possible to predict random effects by using the genomic covariance between the observations (Walsh and Lynch, 2018). The standard mixed model is given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \boldsymbol{\epsilon} \quad ,$$

where  $\mathbf{y}$  is a vector of observation of size  $n$ ,  $\boldsymbol{\beta}$  is an unknown fixed effects vector of size  $q$  for  $q$  fixed effects,  $\mathbf{a}$  is an unknown random effects vector of size  $p$  for  $p$  random effects and  $\boldsymbol{\epsilon}$  is a

residual vector of size  $n$ .  $\mathbf{X}$  is a known design matrix of dimension  $n \times q$  which is associated with fixed effects and  $\mathbf{Z}$  is an incidence matrix of dimension  $n \times p$  which is associated with random effects. The two random vectors  $\mathbf{a}$  and  $\boldsymbol{\epsilon}$  are independent of each other with the mean equal to zero and the unknown variance of  $\mathbf{A}\sigma_a^2$  and  $\mathbf{I}\sigma_e^2$ , respectively.  $\mathbf{A}$  is a known matrix of dimension  $n \times n$  giving the covariance between the individuals and  $\mathbf{I}$  is a diagonal matrix of dimension  $n \times n$  with ones for the diagonal and zeros for off diagonal indicating that residuals are uncorrelated.

Based on the mixed model, any number of fixed effects such as environmental factors can be accounted for, which leads to more accurate prediction of an individual's breeding value and estimation of a population's genetic response (Walsh and Lynch, 2018).

Henderson (1975) suggests BLUP in which residuals are assumed to be identically and independently distributed, while the random effects vectors have a more complicated structure in which their covariance structure need to be specified. In animal and plant breeding, the random effects vector represents the breeding values and the covariance matrix  $\mathbf{A}$  is the relationship matrix whose elements are given by pedigree structure. Therefore, Henderson's mixed model equation (Henderson, 1975) is given by

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \lambda\mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

where  $\lambda = \sigma_e^2/\sigma_a^2$  and  $\mathbf{A}$  is the pedigree based relationship matrix of dimension  $n \times n$ . Meuwissen, *et al.* (2001) proposed using a genome wide dense marker for prediction of genetic value, so that the relationship matrix could be estimated from dense marker information instead of pedigree information.

The BLUP selection which is used in the mixed model aims to find the individuals who possess the highest estimated breeding values in order to use them as the parents of the next generation. This is the main way of selection used by animal breeders and also plant breeders who are working with outcrossing species (Walsh and Lynch, 2018).

## 1.5 Additive genomic prediction models

Breeding values are additive by definition (Falconer and Mackay, 1996), so that the early developments of prediction methods, which took place in dairy cattle breeding to select sires with high breeding values, exclusively accounted for the additive effects underlying relevant quantitative traits (Schaeffer, 2006; VanRaden, 2007; Filho *et al.*, 2016). Additionally, phenotypic variation has been shown to be caused mostly by additive genetic variation (Hill *et al.*, 2008; Mäki-Tanila and Hill, 2014). In this context, GBLUP and Bayesian regressions have been the most commonly used models among additive genomic prediction models (de los Campos, Hickey, *et al.*, 2013).

GBLUP (Meuwissen *et al.*, 2001; VanRaden, 2007) is the widely used prediction method considered as a linear mixed model which assumes additive marker effects (Da *et al.*, 2014; Rönnegård and Shen, 2016; Covarrubias-Pazaran *et al.*, 2018). The GBLUP model is given by

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{g} + \boldsymbol{\epsilon}$$

where  $\mathbf{y}$  is a vector of phenotypes of size  $n$ ,  $\mathbf{1}$  is a vector of ones of the same size as  $\mathbf{y}$ ,  $\mu$  is an unknown population mean considered as a fixed effect,  $\mathbf{Z}$  is a diagonal matrix of dimension  $n \times n$ ,  $\mathbf{g}$  is an unknown vector of breeding values of size  $n$  considered as the random effect which is normally distributed with mean zero and variance  $\mathbf{G}\sigma_g^2$ , and  $\boldsymbol{\epsilon}$  is a vector of errors of size  $n$  which is derived from identically and independently normal distribution of mean zero and variance  $\mathbf{I}\sigma_e^2$ .  $\mathbf{G}$  is a genomic relationship matrix of dimension  $n \times n$  and  $\mathbf{I}$  is an identity matrix of the same dimension as  $\mathbf{G}$ . It is also assumed that  $\mathbf{g}$  and  $\boldsymbol{\epsilon}$  are independent.

Therefore, GBLUP uses relationships between individuals in a genomic relationship matrix  $\mathbf{G}$  which is calculated from SNPs to estimate breeding values. The VanRaden (2008) relationship matrix which is utilized in GBLUP model is given by

$$\mathbf{G} = \frac{(\mathbf{M} - \mathbf{P})(\mathbf{M} - \mathbf{P})'}{2 \cdot \sum_{i=1}^m (p_i(1 - p_i))'}$$

where  $\mathbf{M}$  is a marker matrix of dimension  $n \times p$  which gives  $p$  marker values for  $n$  lines,  $\mathbf{P}$  is the matrix of the same dimension as  $\mathbf{M}$  with  $2 \cdot p_i$  in the  $i^{th}$  column, and  $p_i$  represents allele frequency of minor allele of SNP  $i$ .

Therefore, the population mean and breeding values are estimated by

$$\begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{1} & \mathbf{Z}'\mathbf{Z} + \lambda_{GBLUP}\mathbf{G}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

where  $\lambda_{GBLUP} = \sigma_e^2/\sigma_g^2$ . Based on this equation, the breeding values are estimated for all phenotyped lines and predicted for unphenotyped lines based on their genotypes used to construct the  $\mathbf{G}$  matrix.

rrBLUP (Meuwissen *et al.*, 2001; Endelman, 2011) is another additive genomic prediction model given by

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{W}\mathbf{s} + \boldsymbol{\epsilon}$$

where  $\mathbf{y}$ ,  $\mathbf{1}$ ,  $\mu$  and  $\boldsymbol{\epsilon}$  are as defined in GBLUP,  $\mathbf{W} = \mathbf{M} - \mathbf{P}$  and  $\mathbf{s}$  is an unknown vector of marker effects of size  $n$  considered as a random effect which is identically and independently normally distributed with mean zero and variance  $\mathbf{I}\sigma_s^2$ . With this, the equation is solved by

$$\begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{s}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{W} \\ \mathbf{W}'\mathbf{1} & \mathbf{W}'\mathbf{W} + \lambda_{RRBLUP}\mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}'\mathbf{y} \\ \mathbf{W}'\mathbf{y} \end{bmatrix}$$

Where  $\lambda_{RRBLUP} = \sigma_e^2/\sigma_s^2$ . By solving this equation  $\hat{\mathbf{s}}$  is given by

$$\hat{\mathbf{s}} = \frac{\hat{\sigma}_g^2}{2 \cdot \sum_{i=1}^{m^*} (p_i^* (1 - p_i^*))} \mathbf{W}' (\hat{\sigma}_g^2 \mathbf{G} + \hat{\sigma}_e^2 \mathbf{I})^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})$$

And the variance of  $\hat{\mathbf{s}}$  is given by

$$\hat{\sigma}_s^2 = (\hat{\mathbf{s}} \circ \hat{\mathbf{s}}) 2\mathbf{P}(1 - \mathbf{P})$$

with  $\circ$  denoting the Hadamard product. Therefore, SNP effects can be estimated based on rrBLUP model.

GBLUP which computes additive genetic merit based on the genomic relationship matrix was shown to be equivalent to rrBLUP (Habier *et al.*, 2007; Goddard *et al.*, 2009; Tan *et al.*, 2017). Therefore, the breeding values can also be estimated based on SNP effects using  $\hat{\mathbf{g}} = \mathbf{W}\hat{\mathbf{s}}$  and the variance component  $\sigma_g^2$  can also be estimated by

$$\hat{\sigma}_g^2 = 2 \times \sum_{i=1}^p p_i (1 - p_i) \cdot \hat{\sigma}_s^2$$

Moreover, Bayesian methods (Meuwissen *et al.*, 2001; Habier *et al.*, 2011; Wang *et al.*, 2018) such as ‘‘Bayesian Alphabet’’ (Gianola *et al.*, 2009; Gianola, 2013) and Bayesian Lasso (Park and Casella, 2008) are also utilized as additive genomic prediction models which consider markers as random effects and offer the flexibility of using different priors. Tsai *et al.* (2020) showed that Bayesian Lasso provides higher accuracy than rrBLUP for powdery mildew and yield traits in spring barley, while they had similar prediction accuracies for yield traits in winter wheat. Daetwyler *et al.* (2010) have compared GBLUP and Bayes B in three different effective population sizes including a wide range of numbers of additive quantitative trait loci (QTLs). Their study indicated that GBLUP has a stable accuracy regardless of the number of QTLs, while Bayes B outperformed GBLUP only when the number of QTLs was small. Karaman *et al.* (2016) also compared GBLUP with Bayes B and Bayes C in human height prediction indicating that in the small reference population size (< 6,000 individuals) Bayes B and Bayes C show no advantage over GBLUP, while including more individuals in the reference population results in the superiority of Bayes B and Bayes C over GBLUP. Overall, GBLUP is superior in computing speed to Bayesian method and hard to beat in prediction accuracy for complex traits (Wang *et al.*, 2018) and it is found perform well for the traits with mostly additive genetic background (Ober *et al.*, 2012; Momen *et al.*, 2018).

In this thesis, GBLUP as an additive genomic prediction model with its predictive ability is considered as a reference compared to which the proposed models try to increase the predictive ability (chapter 2, chapter 3 and chapter 4).

## 1.6 Epistasis genomic prediction models

Based on the theory and data, additive genetic variance has shown to be the main source of genetic variation (Hill *et al.*, 2008; Bloom *et al.*, 2013, 2015). However, most of the biological traits have been shown to be influenced by a complex interplay between multiple genes (Forsberg *et al.*, 2017). This contradiction between the biological complexity of the quantitative traits and observation of additive genetic variance capturing most of genetic variation led to a lot of debates in genetics. In fact, many traits of different species show heritable variation in which most of them have complex inheritance patterns with multiple underlying genetic factors (Mackay *et al.*, 2009; Hill, 2010). Therefore, discovering these factors in humans, in model organisms and in agriculturally important species is receiving special attention in genetic research (Buckler *et al.*, 2009; Atwell *et al.*, 2010; Aylor *et al.*, 2011; Mackay *et al.*, 2012). Many loci which have been discovered for a wide range of traits typically explain a minority of each trait's heritability and indicate the existence of other undiscovered genetic factors considered to be the sources of missing heritability (Manolio *et al.*, 2009).

Epistasis as a non-additive interaction between loci can be a potential reason for the gap between the complex trait's heritability and the identified genetic loci's variation (Zuk *et al.*, 2012; Hemani *et al.*, 2013; Brown *et al.*, 2014). "Epistasis" was first coined by Bateson in the studies of multilocus genotype-phenotype maps by which he figured out that the effects of one locus' alleles could be masked by the other loci's alleles (Bateson, 1909). Epistasis is defined as the gene-gene interactions by which the additive effect of a single locus is changed by the additive effect of the interacting locus such that the estimated effects of a single interacting locus will be different between populations with different allele frequencies (Mackay, 2014). Several studies were conducted to shed light on the role of epistasis in genetic variation of complex traits (Cheverud and Routman, 1995; Carlborg and Haley, 2004; Nelson *et al.*, 2013; Mackay, 2014; Forsberg *et al.*, 2017).

The identifiable fractions of human gene expression were demonstrated to be explained by epistasis, while this has proven to be more difficult in human complex traits compared to model organisms. This can result from larger number of possible interactions which have to be tested in the human genome and also from the different genetic architecture in a homogeneous outbred population compared to crossbreeding between inbred lines (Brown *et al.*, 2014). In model organisms epistasis has been detected (Mackay, 2014) and experiments have reported a large number of genetic interactions underlying important biological traits (Forsberg *et al.*, 2017). Many studies have highlighted the importance of epistasis in the genetic architecture of traits in model organisms (Fisher, 1930; Wright, 1931; Carlborg and Haley, 2004; Hill *et al.*, 2008; Huang *et al.*, 2012; Mackay, 2014). Rice yields have also seen to be affected by significant interactions between variants (Huang *et al.*, 2014) as well as metabolic traits in Yeast (Wentzell *et al.*, 2007). It has been found that accounting for epistasis improves phenotype prediction of quantitative traits in Yeast (Forsberg *et al.*, 2017). In fact, epistasis seems to be the most prevalent form of genetic architecture of quantitative traits (Flint and Mackay, 2009; Huang *et al.*, 2012).

The significance of epistasis is undeniable but its use faces some challenges. Mapping epistasis is experimentally, statistically and computationally challenging due to high number of interactions which have to be evaluated. Experimental challenges are caused by requiring large sample sizes for detecting significant interactions and sampling the landscape of possible genetic interactions. Statistical challenges are due to testing multiple hypothesis and computational challenges are caused by a high number of tests having to be evaluated (Mackay, 2014). In fact, epistasis can be represented as any statistical interaction between genotypes at two or more loci (Cheverud and Routman, 1995; Falconer and Mackay, 1996; Lynch and Walsh, 1998) influencing the additive and/or dominance effects of the interacting loci. In quantitative traits, epistasis interaction can influence the effects' magnitude by which a single locus phenotype can be enhanced or suppressed by the other locus genotype, or it can influence the effects' direction. Without considering epistasis, the additive and dominance effects' estimates at each locus remain the same regardless of the genotype of the other locus, whereas the effects of one locus rely on the genotype of its interacting locus when considering epistasis (Mackay, 2014).

Among the genomic prediction models for modeling epistasis, RKHS as non-parametric models (Gianola *et al.*, 2006; Gianola and van Kaam, 2008; de Los Campos *et al.*, 2010) and extended genomic best linear unbiased prediction (EG-BLUP) reduces the required computational load (Jiang and Reif, 2015; Martini *et al.*, 2016). RKHS model based on Gaussian kernel were shown capable of capturing the epistasis effects (de Los Campos *et al.*, 2010). Some studies which were conducted in wheat and maize datasets indicated that RKHS and EG-BLUP have similar predictive abilities (Crossa *et al.*, 2010; Jiang and Reif, 2015; Martini *et al.*, 2016).

EG-BLUP is based on the epistasis relationship matrix given by

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{g}_1 + \mathbf{g}_2 + \boldsymbol{\epsilon}$$

where  $\mathbf{y}$  is a vector of phenotypes of size  $n$ ,  $\mathbf{1}$  is a vector of ones of the same size as  $\mathbf{y}$ ,  $\mu$  is the unknown population mean considered as the fixed effect,  $\mathbf{g}_1$  is the random effect vector of additive genotypic values of size  $n$  which has a normal distribution with mean zero and variance  $\mathbf{G}\sigma_1^2$ ,  $\mathbf{g}_2$  is the random effect vector of additive-by-additive epistasis genotypic values of size  $n$  which has a normal distribution with mean zero and variance  $\mathbf{H}\sigma_2^2$ , and  $\boldsymbol{\epsilon}$  is an errors vector of size  $n$  which driven from identically and independently normal distribution of mean zero and variance  $\mathbf{I}\sigma_e^2$ . It is assumed that  $\mathbf{g}_1$ ,  $\mathbf{g}_2$  and  $\boldsymbol{\epsilon}$  are independent.  $\mathbf{G}$  is an additive relationship matrix is calculated by VanRaden (2008) which was fully described in GBLUP model (section 1.5). Henderson, (1985) proposed the Hadamard product of the additive relationship matrix by itself as the epistasis relationship matrix  $\mathbf{H}$ . Based on the additive genomic relationship matrix  $\mathbf{G}$  this is given by  $\mathbf{H} = \mathbf{G} \circ \mathbf{G}$ .

EG-BLUP can also be given by

$$y_i = \mu + \sum_{j=1}^p M_{i,j}\beta_j + \sum_{k=1}^p \sum_{j=k}^p M_{i,j} M_{i,k} h_{j,k} + \epsilon_i$$

where incorporates the product of marker values with itself as a predictor variable in which  $M_{i,j}$  is the  $j^{th}$  marker value of the  $i^{th}$  line,  $\beta_j$  is an unknown vector of additive marker effects of size  $n$  which are identically and independently normally distributed with mean zero and variance  $\sigma_\beta^2$  and  $h_{j,k}$  is an unknown vector of epistasis marker effects of size  $n$  which are also identically and independently normally distributed with mean zero and variance  $\sigma_h^2$ .

With this, EG-BLUP was shown to exhibit coding-dependent performance indicating that different marker coding lead to different prediction accuracies (He *et al.*, 2015; He and Parida, 2016), since the choice of marker coding specifies how the effects of specific allele combinations at different locus are to be captured for phenotype prediction. Martini *et al.* (2017) discussed this undesirable feature of EG-BLUP and compared different marker coding prediction accuracies indicating that the symmetric coding  $\{-1, 1\}$  or  $\{-1, 0, 1\}$  lead to the highest accuracy in EG-BLUP and standardization by allele frequencies should be avoided (Martini *et al.*, 2017).

The Categorical Epistasis model (CE) was then proposed by Martini *et al.* (2017) which eliminates the undesirable features of EG-BLUP. The CE model treats marker data as a categorical variable, not as a numerical value with modeling allele combination effects as independently normally distributed variable with mean zero. The CE model introduces a dummy variable  $\{0, 1\}$  by which the presence of each combination of alleles of two loci is indicated. This results in an increase in the number of variables in the model. Martini *et al.* (2017) showed that CE can be a valuable alternative for EG-BLUP, since it does not possess the undesired properties of EG-BLUP and also its predictive ability is comparable to EG-BLUP with symmetric coding indicating very high closeness.

Overall, a full epistasis model with all pairwise SNP interactions faces high computational load due to high number of interactions (Rönnegård and Shen, 2016) in addition to the difficult inference of biological interpretations due to over-parametrization. This can introduce a large number of unimportant variables into the model (Martini *et al.*, 2016) which might prevent a considerable gain in prediction accuracy. However, Martini *et al.* (2016) showed that a subset of epistasis interactions with the largest absolute interaction effects sizes has the potential to increase the prediction accuracy.

In this thesis two epistasis models were developed: Epistatic Random Regression BLUP (ERRBLUP) as a full epistasis model which incorporates all pairwise SNP interactions (chapter 2, chapter 3 and chapter 4); and selective Epistatic Random Regression BLUP (sERRBLUP) as a selective epistasis model which incorporates a subset of pairwise SNP interactions selected based on their absolute effects sizes (chapter 2) or effect variances (chapter 3 and 4). ERRBLUP and sERRBLUP are considered as CE models and their predictive abilities are compared to GBLUP.

## 1.7 Univariate and multivariate genomic prediction models

In animal and plant breeding programs, multi-trait and multi-environment data are quite common. Therefore, powerful statistical models are required to use these data and exploit the correlation between the traits to improve the prediction accuracy for genomic selection purposes (Montesinos-López *et al.*, 2018).

Genomic prediction models can be classified into two main categories of univariate models and multivariate models based on the number of desired traits to be analyzed. The univariate models are designed to predict a single phenotypic trait, while the multivariate models as multi-trait and multi-environment models are designed to predict multiple phenotypic traits simultaneously. Utilizing multi-trait models helps to capture the complex relationships between the traits more efficiently than univariate models and mostly results in more accurate prediction. Multi-trait models have been recently more popular in genomic selection due to their capacity of predicting multiple traits concurrently and the ability to increase the prediction accuracy compared to univariate models when the genetic correlation between the traits is high (Jia and Jannink, 2012; Jiang *et al.*, 2015; Montesinos-López *et al.*, 2018). Multi-trait models mostly provide higher prediction accuracy for correlated traits compared to univariate models (He *et al.*, 2016; Schulthess *et al.*, 2018), although some studies reported just a modest increase in their prediction accuracy (Calus and Veerkamp, 2011; Montesinos-López *et al.*, 2016).

Henderson and Quaas (1976) proposed the first application of mixed models for multi-trait evaluation. Multi-trait models were initially proposed in animal breeding to model genetic correlation among traits and to model genotype by environment interactions across multiple years or environments (Mrode, 2014; Lee and van der Werf, 2016). The initial multivariate models which were applied to plant and animal species were based on available pedigree information to infer relationships among individuals and traits in mixed model framework (Mrode, 2014). However, the wide availability of dense molecular markers led to a replacement of the limited pedigree information to construct genomic relationship matrices resulting in new options for analyzing crops with restricted pedigree information (Endelman and Jannink, 2012). Velazco *et al.* (2019) reported an improvement in predictive ability of multi-trait GBLUP compared to single-trait GBLUP in sorghum. Their study illustrated that multi-trait GBLUP increases the predictive ability of grain yield up to 16 percent by including plant height information into a multi-trait GBLUP model. This might be due to the strong genetic correlation between grain yield and plant height in sorghum hybrids (Velazco *et al.*, 2019). Covarrubias-Pazarán *et al.* (2018) also showed that under medium or high genetic correlation, multivariate GBLUP exhibited higher accuracy than univariate GBLUP.

In plant breeding, one of the breeders' major challenges is the difference in genotype performance from one environment to the other environments which is known as  $\mathbf{G} \times \mathbf{E}$  interaction (Kang and Gorman, 1989). Multi-environment models are usually employed to assess  $\mathbf{G} \times \mathbf{E}$  interaction for a single trait when the information on multiple genotype is recorded in multiple environments (Montesinos-López *et al.*, 2016; Hassen *et al.*, 2018). Inclusion of  $\mathbf{G} \times \mathbf{E}$  interaction in genomic



prediction models help selection of lines with optimal overall performance across target environments in genomic selection context (Roorkiwal *et al.*, 2018).

Several statistical models have been used to estimate  $\mathbf{G} \times \mathbf{E}$  interaction in plant breeding such as linear regression, Analysis of Variance (ANOVA) models and linear mixed models (Elias *et al.*, 2016). Incorporating genotype  $\times$  environment ( $\mathbf{G} \times \mathbf{E}$ ) interaction into additive genomic prediction models in multi environment analysis has been reported to be potentially successful in increasing predictive ability (Hallauer *et al.*, 2010). Burgueño *et al.* (2012) proposed the first statistical framework to model  $\mathbf{G} \times \mathbf{E}$  using a linear mixed model for genomic prediction so that the single-trait, single-environment GBLUP model was extended to the multi-environment context. This approach was based on borrowing information across environments which resulted in higher prediction accuracy (Burgueño *et al.*, 2012). Days to heading, and days to maturity in Iranian and Mexican wheat landraces in drought and heat environments has been evaluated by Crossa *et al.* (2016) which indicated that inclusion of  $\mathbf{G} \times \mathbf{E}$  interaction in genomic prediction model lead to substantial and consistent increase in prediction accuracy compared to models without the  $\mathbf{G} \times \mathbf{E}$  term. Inclusion of  $\mathbf{G} \times \mathbf{E}$  interaction in a whole regression approach also leading to accurate prediction of maize yield (Millet *et al.*, 2019), and the highly significant effect of  $\mathbf{G} \times \mathbf{E}$  interaction on grain yield for single cross maize hybrids across environments with low and optimum availability of nitrogen in the soil (Mafouasson *et al.*, 2018) are some examples of multi environment models. Moreover, multi-environment analysis can also be utilized for multi-year analysis in the scenario of changing environmental conditions (Elias *et al.*, 2016). In fact, gathering phenotypic data over the years to predict the lines in the upcoming years is a potential approach to increase the prediction accuracy, such as including historical phenotypic data in genomic prediction of hybrids in grain maize which has shown to increase its prediction accuracy (Schrag *et al.*, 2019a).

Additionally, Martini *et al.* (2016) showed the feasibility of borrowing information across environments in EG-BLUP without incorporating additional terms such as  $\mathbf{G} \times \mathbf{E}$  interaction into the epistasis genomic prediction model. This method resulted in an increase in predictive ability in one environment by variable selection in the other environment under the assumption of positive correlation of phenotypes in different environments which was demonstrated with the publicly available wheat data set (Pérez and de los Campos, 2014).

Overall, in the context of crop and livestock breeding, developing efficient selection strategies and powerful statistical models with higher prediction accuracy for which the costly and time consuming phenotyping of numerous selection candidates in multiple environment could be mitigated deserves special attention.

In this thesis, GBLUP, ERRBLUP and sERRBLUP models have been compared in their predictive abilities in the univariate statistical framework for the simulated phenotypes from genotypes of the publicly available wheat dataset (Pérez and de los Campos, 2014) (chapter 2). We further compared GBLUP, ERRBLUP and sERRBLUP models in both univariate and bivariate statistical frameworks for prediction across environments in 910 doubled haploid lines from European maize

landraces Kemater Landmais Gelb and Petkuser Ferdinand Rot in six locations in Germany and Spain for series of eight phenotypic traits gathered in the year 2017 (Chapter 3). Bivariate GBLUP, ERRBLUP and sERRBLUP models have been finally compared for prediction across years in the maize dataset in four locations in Germany and Spain by modeling the years 2017 and 2018 as two separate traits in in multi-trait model (chapter 4).

## 2 Phenotype prediction under epistasis

This chapter contains the manuscript “Phenotype prediction under epistasis” which has been published as a chapter of the book “Epistasis: Methods and Protocols” (Vojgani *et al.*, 2021). In order to have uniform style in the thesis, the journal style is not used in this chapter.

This manuscript is the joint work of Elaheh Vojgani<sup>1\*</sup>, Torsten Pook<sup>2</sup> and Henner Simianer<sup>1</sup> which is focused on developing epistasis models in univariate statistical framework and implementing in the associated R-Package “EpiGP” (Vojgani *et al.*, 2019) which is available at: <https://github.com/evoigani/EpiGP>.

1,2: University of Goettingen, Center for Integrated Breeding Research, Animal Breeding and Genetics Group, Goettingen, Germany

### **Author contributions by EV**

EV developed and wrote the associated R-package “EpiGP”, analyzed the data, derived the results, wrote the initial manuscript and led the revision of the manuscript.

## 2.1 Abstract

Reliable methods of phenotype prediction from genomic data play an increasingly important role in many areas of plant and animal breeding. Thus, developing methods which enhance prediction accuracy is of major interest. Here, we provide three methods for this purpose: (i) Genomic Best Linear Unbiased Prediction (GBLUP) as a model just accounting for additive SNP effects; (ii) Epistatic Random Regression BLUP (ERRBLUP) as a full epistatic model which incorporates all pairwise SNP interactions, and (iii) selective Epistatic Random Regression BLUP (sERRBLUP) as an epistatic model which incorporates a subset of pairwise SNP interactions selected based on their absolute effect sizes or the effect variances which is computed based on solutions from the ERRBLUP model. We compared the predictive ability obtained from GBLUP, ERRBLUP and sERRBLUP with genotypes from a publicly available wheat dataset and respective simulated phenotypes. Results showed, that sERRBLUP provides a substantial increase in prediction accuracy compared to the other methods when the optimum proportion of SNP interactions are kept in the model, especially when an optimal proportion of SNP interactions is selected based on the SNP interaction effect sizes. All methods described here are implemented in the R-package EpiGP, which is able to process large scale genomic data in a computationally efficient way.

**Keywords:** GBLUP, Epistasis model, Genomic prediction, Phenotype prediction, R-package, EpiGP

## 2.2 Introduction

Genomic prediction of phenotypes is of interest in all species and especially in breeding programs. In practice, there is an option to reduce breeding cycle time by genomic prediction since the individuals can be selected and crossed without being phenotyped (Edwards *et al.*, 2019). Moreover, it can enhance selection accuracy since genomic data provide more powerful statistical models in comparison to experimental designs which use more observation than what can be phenotyped for a single trait. This is resulting from reducing the cost of evaluating individuals by reducing the number of phenotyped individuals and/or their replicates leading to an increase in selection intensity by genomic selection application (Edwards *et al.*, 2019). Additionally, genomic prediction models could be updated by utilizing the trait's data from previous years and become more accurate (Edwards *et al.*, 2019). Accordingly, one of the major challenges of contemporary biology is understanding the effect of genetic variation on phenotypic variation in quantitative traits (Mackay, 2014). Therefore, genomic prediction of complex phenotypes has been developed by using all markers simultaneously instead of only considering significant marker-trait association (Windhausen *et al.*, 2012) in plants (Crossa *et al.*, 2010), animals (Daetwyler *et al.*, 2013) and humans (de los Campos, Vazquez, *et al.*, 2013). In this context, several statistical models have been developed for genomic prediction of phenotypes which were compared in terms of predictive ability, defined as the correlation between predicted and observed phenotypes in a test set not used for the training of the model. Initially prediction models exclusively accounted for

additive effects on phenotypic traits, while non-additive effect were typically ignored (Filho *et al.*, 2016). In this regard, GBLUP (VanRaden, 2008) as a linear mixed model which assumes additive marker effects is one of the widely used genomic prediction models (Da *et al.*, 2014; Rönnegård and Shen, 2016; Covarrubias-Pazarán *et al.*, 2018). GBLUP has been superior in computing speed and prediction accuracy of complex traits compared to Bayesian methods which consider markers as random effects and offer the flexibility of using different priors (Wang *et al.*, 2018). It has been shown that BayesB (Meuwissen *et al.*, 2001) has higher accuracy than GBLUP for traits controlled by smaller numbers of quantitative trait nucleotides, otherwise GBLUP has the higher accuracy (Daetwyler *et al.*, 2010). In this regard, BayesB and BayesC methods in human height prediction show no advantage over GBLUP in small reference population (<6,000 individuals), while these methods outperformed GBLUP when more individuals were included in the reference population (Karaman *et al.*, 2016). Bayesian Monte Carlo approaches also have been considered for prediction of complex traits in humans, but they only produced a modest increase in predictive ability at the cost of large computation times (Lello *et al.*, 2019).

Moreover, it has been shown that developing models which incorporate epistatic interactions in addition to additive marker effects can enhance the prediction accuracy, since epistasis contributes substantially to the genetic variation of quantitative traits (Mackay, 2014). Many studies have highlighted the importance of epistasis in the genetic architecture of quantitative traits in model organisms (Fisher, 1930; Wright, 1931; Carlborg and Haley, 2004; Hill *et al.*, 2008; Huang *et al.*, 2012; Mackay, 2014). Therefore, prediction accuracy of the genomic prediction models incorporating interactions compared to models with only main effects have been widely discussed over the last years, and evidence was provided that including epistasis for prediction of complex trait phenotypes has the potential to increase the predictive ability (Carlborg and Haley, 2004; Hu *et al.*, 2011; Wang *et al.*, 2012; Mackay, 2014; Jiang and Reif, 2015; Rönnegård and Shen, 2016). However, epistasis has often been ignored due to the high computational load, especially in the case of large sets of marker data. Additionally, since breeding values are additive by design, higher accuracy in phenotype prediction does not automatically lead to higher genomic gain in the breeding program. The extended GBLUP (EGBLUP) model reduces the computational load by using an equivalent marker-based epistatic relationship matrix (Jiang and Reif, 2015; Martini *et al.*, 2016) rather than all pairwise marker interaction effects. But still, full epistatic models have been largely ignored in practical applications due to the computational load through the high number of interactions to be included (Rönnegård and Shen, 2016). Subsequently it was found with a small example data set that reducing full epistatic models by incorporating certain subnetworks of pairwise SNP interactions with highest pairwise interaction effects increases the predictive ability (Martini *et al.*, 2016).

Here, we provide full epistatic models by utilizing marker based epistatic relationship matrices which are applicable to large dataset and large scale genotype data. We further developed a reduced epistatic model which is based on the selection of different subset of the pairwise SNP interactions based on their estimated effect sizes or estimated effect variances, to assess the optimum proportion of SNP interactions to be kept in the model resulting in maximum predictive ability.

Additionally, we explain step by step our R-package EpiGP for epistatic genomic prediction of phenotypes based on our developed methods.

## 2.3 Methods

Genetic analyses are generally based on mixed models which consider both fixed and random effects (Henderson, 1975). Fixed effects are unknown constants, while random effects are random variables which are sampled from a certain distribution. In fact, the variance components in the mixed model indicate how much of the observed variance can be attributed to the random effects and how much remains in the residual variance, where It is often assumed that the residual variance is independent of genetic variance. The standard assumption in the covariance matrix of residuals is that all the residuals are uncorrelated with common variance. If existing random effects are not specified in the model, the corresponding covariance structure not covered by the model is inadvertently captured by the residuals parts. Therefore, an inadequate specification of random effects in the model may lead to a violation of the assumption that residuals are independently distributed. The covariance structure of random effects in genomic models is given by genetic variance multiplied by the genomic relationship matrix constructed from SNP information, which is reflecting the relatedness between the individuals (Walsh and Lynch, 2018).

Here, a linear mixed model with the assumption of  $n$  genotyped lines in which a subset of  $n_1$  lines are phenotyped can be represented as follows. These  $n_1$  lines are used to train the model which then is used to predict phenotypes for the remaining  $n_2 = n - n_1$  lines. The basic model is

$$\mathbf{y} = \mathbf{1}\mu + (\mathbf{I} \quad \mathbf{0})\mathbf{g} + \boldsymbol{\epsilon}, \quad (\text{eq. 1})$$

where  $\mathbf{y}$  is an  $n_1 \times 1$  vector of phenotypes (any outcome variable),  $\mathbf{1}$  is an  $n_1 \times 1$  vector with all entries equal to 1,  $\mu$  is a scalar population mean which is modelled as a fixed effect,  $\mathbf{I}$  is an identity matrix of dimension  $n_1 \times n_1$  and  $\mathbf{0}$  is a matrix of dimension  $n_1 \times n_2$  of zeros.  $\mathbf{g} \sim N(\mathbf{0}, \boldsymbol{\Gamma}\sigma_g^2)$  is an  $n \times 1$  vector of genetic effects which are modelled as random effect, and  $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma_\epsilon^2)$  is a random error vector, where  $\boldsymbol{\Gamma}$  and  $\mathbf{I}$  are the respective dispersion matrices and  $\sigma_g^2$  and  $\sigma_\epsilon^2$  are the corresponding variance components.

The population mean and the genetic effects for all phenotyped and non phenotyped lines are estimated using

$$\begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{g}}_1 \\ \hat{\mathbf{g}}_2 \end{bmatrix} = \begin{bmatrix} n_1 & \mathbf{1}' & \mathbf{0} \\ \mathbf{1} & \mathbf{I} + \lambda\boldsymbol{\Gamma}^{11} & \lambda\boldsymbol{\Gamma}^{12} \\ \mathbf{0} & \lambda\boldsymbol{\Gamma}^{21} & \lambda\boldsymbol{\Gamma}^{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}'\mathbf{y} \\ \mathbf{y} \\ \mathbf{0} \end{bmatrix}, \quad (\text{eq. 2})$$

where  $\lambda = \sigma_\epsilon^2/\sigma_g^2$ ,  $\boldsymbol{\Gamma}^{-1} = \begin{bmatrix} \boldsymbol{\Gamma}^{11} & \boldsymbol{\Gamma}^{12} \\ \boldsymbol{\Gamma}^{21} & \boldsymbol{\Gamma}^{22} \end{bmatrix}$ ,  $\mathbf{g} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix}$  and the indices pertain to the subset of individuals with (1) or without (2) phenotypes, respectively.

Then, the phenotype prediction for the set of unphenotyped individuals can be achieved by

$$\hat{\mathbf{y}}_2 = \mathbf{1}_2 \hat{\mu} + \hat{\mathbf{g}}_2, \quad (\text{eq. 3})$$

where  $\hat{\mathbf{y}}_2$  is the  $n_2 \times 1$  vector of predicted phenotypes and  $\mathbf{1}_2$  is an  $n_2 \times 1$  vector of ones.

If all  $n$  lines are phenotyped and genotyped which means  $n = n_1$  and  $n_2 = 0$ , eq. 1 provides estimates of genetic effects.

GBLUP, ERRBLUP and sERRBLUP as three different phenotype prediction models can be represented by eq. 1 by differing only in the specification of the dispersion matrix  $\mathbf{\Gamma}$  reflecting the covariance structure of the genetic effects.

### 2.3.1 Genomic Best Linear Unbiased Prediction (GBLUP)

GBLUP as linear mixed model which assumes additive marker effects uses the genomic relationship matrix  $\mathbf{G}$  suggested by VanRaden (VanRaden, 2008) as  $\mathbf{\Gamma}$ :

$$\mathbf{\Gamma}_{VR} = \frac{(\mathbf{M} - \mathbf{P})(\mathbf{M} - \mathbf{P})'}{2 \cdot \sum_{i=1}^m (p_i(1 - p_i))},$$

where  $\mathbf{M}$  is an  $n \times m$  marker matrix which gives  $m$  marker values for  $n$  lines.  $\mathbf{P}$  is a matrix of equal size as  $\mathbf{M}$  with  $2 \cdot p_i$  in the  $i^{th}$  column, and  $p_i$  is allele frequency of the minor allele of SNP  $i$ .

### 2.3.2 Epistatic Random Regression BLUP (ERRBLUP)

ERRBLUP model is a full epistatic model which incorporates all pairwise SNP interactions in the prediction model. First, a marker combination matrix  $\mathbf{M}^*$  of dimension  $n \times m^*$  is generated whose element  $i, j$  is 1 if genotype combination  $j$  is present in individual  $i$  and is 0 otherwise. If the  $m$  markers have 2 possible genotypes at a single locus, i.e. 0 or 2 when coded as the counts of the minor allele, then for each pair of loci, there will be 4 different genotype combinations: {00, 02, 20, 22}. The total number of pairs of loci is  $\frac{m \times (m+1)}{2}$  if interaction of a locus with itself is also considered. Since for each of these pairs there are four possible genotype combinations, the total number of combinations to be considered as dummy variables is

$$m^* = 4 \times \frac{m \times (m+1)}{2} = 2m \times (m + 1).$$

If the heterozygous are also possible at a single locus, then there will be 9 different genotype combinations instead of 4: {00, 01, 02, 10, 20, 11, 12, 21, 22}. Therefore, the total number of combinations is

$$m^* = 9 \times \frac{m \times (m+1)}{2}.$$

Then, VanRaden (VanRaden, 2008) relationship matrix based on the marker combination matrix  $\mathbf{M}^*$  which includes all pairwise SNP interactions is calculated as

$$\mathbf{\Gamma}_{ERR} = \frac{(\mathbf{M}^* - \mathbf{P}^*)(\mathbf{M}^* - \mathbf{P}^*)'}{\sum_{i=1}^{m^*} (p_i^*(1 - p_i^*))},$$

where  $p_i^*$  is the average value giving the frequency of respective genotype combination in the population for column  $i$  of matrix  $\mathbf{M}^*$ .  $\mathbf{P}^*$  is a matrix of equal size as  $\mathbf{M}^*$  with  $p_i^*$  in the  $i^{th}$  column. As  $m^*$  can be extremely large, we highly recommend the use of the R-package “miraculix” (Schlather, 2020) as it is about 15 times as fast as the regular matrix multiplications on genotype data in R.

This matrix is used in the ERRBLUP model as the dispersion matrix for the genetic effects, which now are based on epistatic interaction effects. It should be noted that including the interaction of each locus with itself replaces the additive effect, so that it is not necessary to use a model that separately accounts for both additive and epistatic effects.

ERRBLUP genomic prediction of phenotypes can be done by the package EpiGP (Vojgani *et al.*, 2019) in R. The EpiGP R-package depends on the R-packages MASS, EMMREML and stats. First of all, the package should be installed in R

```
1. devtools::install_github("evojgani/EpiGP", subdir="pkg")
```

The package expects SNP markers to be coded {0, 1, 2} or {0, 2} in fully inbred lines. If coded differently, the following function of EpiGP provides the desired marker matrix from {-1, 0, 1} or {0, 1} or character coded marker matrix

```
2. Recodemarkers (M)
```

```
# M is the original marker matrix which could be {-1, 0, 1} or {0, 1} or character coded marker matrix.
```

```
# The output of the “Recodemarkers” function is a {0, 1, 2} or {0, 2} coded marker matrix.
```

```
The function also report how the original markers have been recoded.
```

Afterwards,  $\mathbf{\Gamma}_{ERR}$  can be calculated using

```
3. Gall(m, cores=1)
```

```
# m is a {0,1,2} or {0,2} coded marker matrix with individuals in rows and the markers in columns.
```

```
# cores is the number of cores and the default value is 1.
```

```
# The outputs of the “Gall” function are two components, G as the ERRBLUP relationship matrix, and P as a vector of all genotype combinations frequencies in the population.
```

Finally, ERRBLUP genomic prediction of phenotypes can be done by

```
4. ERRBLUP_Stepwise(Pheno, G_ERRBLUP)
```

```
# Pheno is a numeric vector of phenotypes.
```



# **G\_ERRBLUP** is the ERRBLUP relationship matrix which is one of the outputs of “Gall” function.

# The output of the “ERRBLUP” function is a vector of both phenotype estimations for training set and phenotype predictions for test sets based on ERRBLUP method.

### 2.3.3 selective Epistatic Random Regression BLUP (sERRBLUP)

The sERRBLUP model is based on the same approach as ERRBLUP, but here the  $\mathbf{\Gamma}$ -matrix is constructed from a selected subset of genotype interactions. Selection of genotype interactions can be either based on the effects estimates or their variances estimates. In this regard, interaction effects  $\hat{\mathbf{t}}$  and their variances  $\hat{\sigma}_i^2$  from the ERRBLUP model (Fragomeni *et al.*, 2014; Mrode, 2014) are calculated as

$$\hat{\mathbf{t}} = \frac{\widehat{\sigma}_g^2}{\sum_{i=1}^{m^*} (p_i^* (1 - p_i^*))} (\mathbf{M}^* - \mathbf{P}^*)' \left( \widehat{\sigma}_g^2 \mathbf{\Gamma}_{ERR} + \widehat{\sigma}_\epsilon^2 \mathbf{I} \right)^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}),$$

$$\hat{\sigma}_i^2 = \hat{\mathbf{t}}_i^2 2p_i^* (1 - p_i^*),$$

Then, those interactions whose absolute effect estimates or their respective variance estimates are in the top desired proportion  $\pi$  of all interaction effects are selected. We suggest to use a range of  $\pi = \{0.1, 0.05, 0.01 \text{ or } 0.001\}$  since it was observed in preliminary analyses that they cover the most relevant range, but any value  $0 < \pi < 1$  is possible. For a desired subset, specified by the chosen criterion and  $\pi$ , a reduced matrices  $\mathbf{M}_\pi^*$  of dimension  $n \times \pi m^*$  is generated, containing only those columns of  $\mathbf{M}^*$  and pertaining to the selected subset of genotype interactions, and then a VanRaden (VanRaden, 2008) type relationship matrix based on a reduced set of epistatic interactions

$$\mathbf{\Gamma}_{sERR} = \frac{(\mathbf{M}_\pi^* - \mathbf{P}_\pi^*)(\mathbf{M}_\pi^* - \mathbf{P}_\pi^*)'}{\sum_{i=1}^{\pi m^*} (p_{\pi i}^* (1 - p_{\pi i}^*))},$$

is calculated, where  $p_{\pi i}^*$  is the average frequency of the selected genotype combination in the population for column  $i$  of matrix  $\mathbf{M}_\pi^*$ .  $\mathbf{P}_\pi^*$  is a matrix of equal size as  $\mathbf{M}_\pi^*$  with  $p_{\pi i}^*$  in the  $i^{th}$  column.

sERRBLUP genomic prediction of phenotypes can also be done by R-package EpiGP (Vojgani *et al.*, 2019). First of all, SNP interaction effects and their variances can be estimated by the following functions

#### 5. SNP\_Effect\_Var (**m**, **Pheno**, **G\_ERRBLUP**, **P**, **cores=1**)

# **m** is a  $\{0,1,2\}$  or  $\{0,2\}$  coded marker matrix with individuals in the rows and the markers in the columns.

# **Pheno** is a numeric vector of phenotypes.

# **G\_ERRBLUP** is the ERRBLUP relationship matrix which is one of the outputs of “Gall” function.

# **P** is a vector of all genotype combinations frequencies in the population which is one of the outputs of “Gall” function.

# **cores** is the number of cores and the default value is 1.

# The outputs of the “SNP\_Effect\_Var” function are two components, **Effect** as a vector of all estimated pairwise SNP interaction effects and **Effect.Var** as a vector of all estimated pairwise SNP interaction effects variances.

Then  $\Gamma_{sERR}$  is calculated for the top desired SNP interactions selected either based on the absolute effect estimates or their respective variance estimates

#### 6. Gtop (m, Estimations, k, cores=1)

# **m** is a {0,1,2} or {0,2} coded marker matrix with individuals in the rows and the markers in the columns.

# **Estimations** could be either a vector of all estimated pairwise SNP interaction effects (**Effect**) or a vector of all estimated pairwise SNP interaction effects variances (**Effect.Var**) which are the outputs of the “SNP\_Effect\_Var” function.

# **k** is the desired proportion of all SNP interactions to be selected.

# **cores** is the number of cores and the default value is 1.

# The output of the “Gtop” function is the sERRBLUP Relationship matrix for the “k” proportion of pairwise SNP interactions.

Finally, sERRBLUP genomic prediction of phenotypes can be done by

#### 7. sERRBLUP\_Stepwise(Pheno, G\_sERRBLUP)

# **Pheno** is a numeric vector of phenotypes.

# **G\_sERRBLUP** is the sERRBLUP relationship matrix (calculated either from SNP pairs selected based on the interaction effects estimates or their variances) for the “k” proportion of pairwise SNP interactions which is the output of “Gtop” function.

# The output of the “sERRBLUP” function is a vector of both phenotype estimations for the training set and phenotype predictions for the test set based on the sERRBLUP method.

In addition to the provided functions by which ERRBLUP and sERRBLUP phenotype prediction can be done step by step, there are two functions in the EpiGP package which provide ERRBLUP and sERRBLUP phenotype prediction in one step. The functions are as follows:

#### ERRBLUP(M, Pheno, cores = 1)

# **M** is the original marker matrix which could be {-1, 0, 1} or {0, 1} or character coded marker matrix.

# **Pheno** is a numeric vector of phenotypes.

# **cores** is the number of cores and the default value is 1.

The output of the function is a list of three components:

- 1- **Recodedmarkers** which is a {0, 1, 2} or {0, 2} coded marker matrix.
- 2- **Relationshipmatrix** which is a list of two components: ERRBLUP relationship matrix (**G**) and a vector of all genotype combinations frequencies in the population (**P**).
- 3- **Predictions** which is a numeric vector of both phenotype estimations of training set and phenotype predictions of test set based on ERRBLUP method.

**sERRBLUP(M, Pheno, k, cores = 1)**

# **M** is the original marker matrix which could be {-1, 0, 1} or {0, 1} or character coded marker matrix.

# **Pheno** is a numeric vector of phenotypes.

# **K** is the desired proportion of all SNP interactions to be selected.

# **cores** is the number of cores and the default value is 1.

The output of the function is a list of three components:

- 1- **Recodedmarkers** which is a {0, 1, 2} or {0, 2} coded marker matrix.
- 2- **Relationshipmatrix** which is a list of two components: ERRBLUP relationship matrix (**G**) and a vector of all genotype combinations frequencies in the population (**P**).
- 3- **Effect.Relationshipmatrix** which is the sERRBLUP Relationship matrix for the k percent of pairwise SNP interactions based on effect size estimations.
- 4- **Var.Relationshipmatrix** which is the sERRBLUP Relationship matrix for the k percent of pairwise SNP interactions based on effect size variance estimations.
- 5- **Effect** which is a numeric vector of all estimated pairwise SNP interaction effects.
- 6- **Effect.Variance** which is a numeric vector of all estimated pairwise SNP interaction effects variances.
- 7- **Effect.Predictions** a numeric vector of both phenotype estimations of training set and phenotype predictions of test set based on effect sizes.
- 8- **Var.Prediction** which is a numeric vector of both phenotype estimations of training set and phenotype predictions of test set based on effect sizes variances.

Based on 5-fold cross validation with 5 replicates we validated our method with the wheat data set genotype of CIMMYT inbred lines (Crossa *et al.*, 2010) which is included in the R-package BGLR (Pérez and de los Campos, 2014). The wheat lines were genotyped by 1279 Diversity Array Technology (DArT) markers, which are coded in binary form as being present (1) or absent (0). We simulated phenotypes for this dataset by randomly selection of 80 SNPs from marker matrix and generation of the phenotypic trait based on all 4 possible genotype combinations resulting in purely epistatic trait with the heritability of 1.

The following code is used for ERRBLUP and sERRBLUP in the wheat dataset markers and respective simulated phenotypes:

```

library(EpiGP)
library(BGLR)
data(wheat)
# First we recode the markers and named the lines in the marker matrix
m <- Recodemarkers(wheat.X)
# Then we calculated the ERRBLUP relationship matrix
G_all <- Gall (m, cores=1)
G_ERRBLUP <- G_all$G
P <- G_all$P
# Afterwards, we consider a subset of phenotypic values as a training set and
do the phenotype prediction based on ERRBLUP, we have done this by 5-fold cross
validation with 5 replicates. As an example here we randomly select 60 lines as
the test set and the remaining lines are considered as the training set as
follows
N <- length(Phenotype)
n <- 60
Testset <- sample(1:N,n)
Pheno <- Phenotype
pheno[test] <- NA
ERRBLUP_pred <- ERRBLUP_Stepwise(pheno, G_ERRBLUP)
# Then interaction effects and interaction effect variances are calculated as
follows
Estimations <- SNP_Effect_Var(m, pheno, G_ERRBLUP, P, cores=15)
t_hat <- Estimations$ Effect
sigma_hat <- Estimations$ Effect.Var
# Based on the desired proportion of interactions, the sERRBLUP relationship
matrix is calculated for both selections based on estimated effects and
estimated effect variances
k <- 10
Gtop_effect <- Gtop(m, t_hat, k, cores=1)
Gtop_var <- Gtop(m, sigma_hat, k, cores=1)
# Finally, sERRBLUP is performed based in both approaches as follows
sERRBLUP_effect <- sERRBLUP_Stepwise(pheno, Gtop_effect)
sERRBLUP_var <- sERRBLUP_Stepwise(pheno, Gtop_var)

```

Fig. 2.1 provides the comparison in predictive ability of GBLUP, ERRBLUP and sERRBLUP when up to 99.9% of SNP interactions are removed in the wheat dataset. Red dots demonstrate the sERRBLUP predicative ability when the SNP interaction selections are based on the estimated effects while green dots demonstrate the sERRBLUP predicative ability when the SNP interaction selections are based on the estimated effect variances. It is demonstrated that GBLUP has the minimum predictive ability (0.821), while the predictive ability obtained from ERRBLUP as a full epistatic model which incorporates all pairwise SNP interactions is 0.836. The maximum predictive ability (0.889) was observed for sERRBLUP when the top 1 percent of interactions are selected based on their absolute effect sizes. The sERRBLUP model selecting interactions based on their effect variances provides the best predictive ability (0.883) when 0.1 percent of the interaction effects are maintained in the model. In both sERRBLUP models there is a massive decrease in predictive ability if too many (more than 99.9 percent) SNP interactions are disregarded in the construction of the relationship matrix. In this simulated data, predictive ability of sERRBLUP decreases to 0.028 and 0.126 for the top 0.01 percent of interactions for selection based on effect sizes and their variances, respectively. This is a general pattern we observed in several data sets, however, the exact proportion of SNP interactions to account for in the model is data dependent and needs to be determined in all applications separately. EpiGP provides a function by which one can determine sERRBLUP predictive ability for different selected proportions of interactions and be able to compare their predictive abilities to choose the best proportion of interactions which provide the highest accuracy.

**sERRBLUP\_Proportions\_Test (M, Pheno, k, cores = 1)**

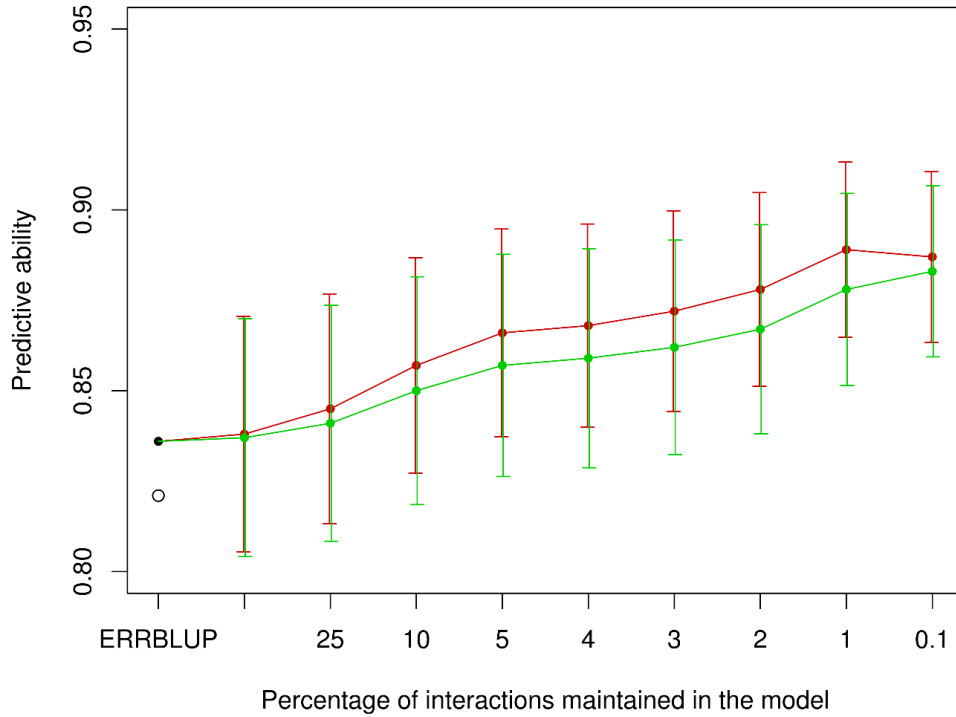
# **M** is the original marker matrix which could be {-1, 0, 1} or {0, 1} or character coded marker matrix.

# **Pheno** is a numeric vector of phenotypes.

# **K** is the desired proportion of all SNP interactions to be selected.

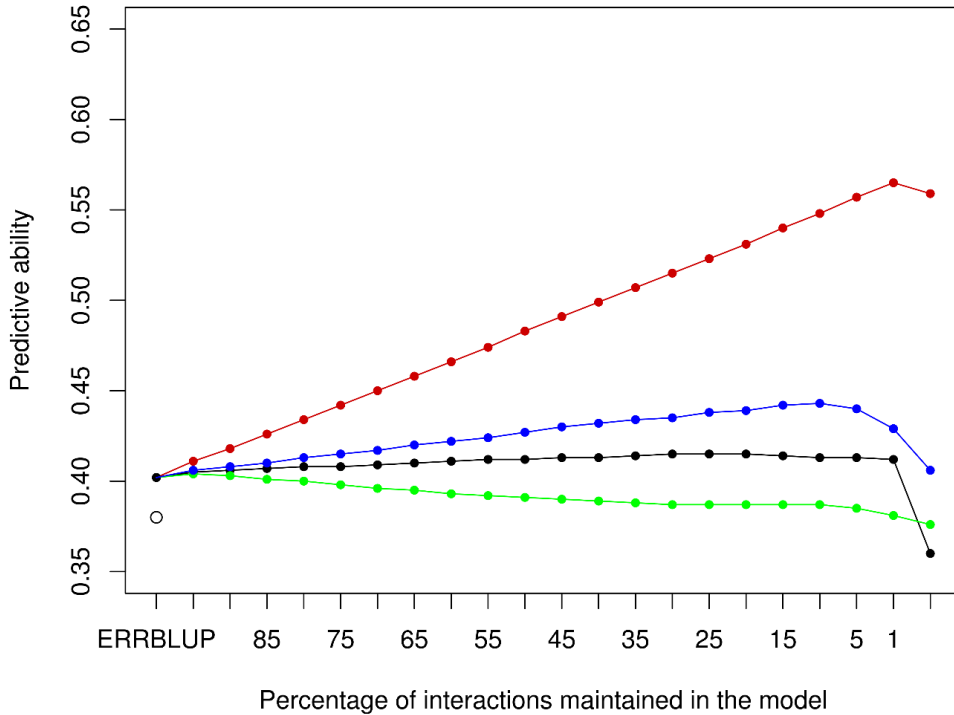
# **cores** is the number of cores and the default value is 1.

The output of the function is the data frame of three components: **Desired.Proportion** as the proportion of SNP interactions which is maintained in sERRBLUP model, **PA.Effcet** as the sERRBLUP predictive ability based on effect sizes selection and **PA.Var** as the sERRBLUP predictive ability based on effect variances selection.



**Fig. 2.1:** Comparison of predictive ability of GBLUP (open circle), ERRBLUP (black filled circle) and sERRBLUP for 50 to 0.1% of SNP interactions maintained in the model. Red and green dots represent the sERRBLUP predictive ability (with standard errors) when the SNP interaction selections are based on the estimated effects and estimated effects variances, respectively. The vertical arrows represent the predictions  $\pm$  their standard deviations.

Fig. 2.2 demonstrates predictive ability with sERRBLUP model when interactions are selected based on their absolute effect sizes for prediction within and across environment 3 of the wheat dataset. The green dots represent sERRBLUP predictive ability for prediction within environment 3 which shows no increase in predictive ability, while the black, blue and red dots represent the predictive ability by variable selection in environment 1, 2 and 4 and prediction of the phenotypes in environment 3 (Martini *et al.*, 2016). ERRBLUP and sERRBLUP which capture all possible genotype combinations, have higher predictive abilities for prediction across environment 3 compare to EGLUP model (0.390) which just captures {22} genotype combination and is affected by translation of the coding in marker matrix (Martini *et al.*, 2016, 2017). It is also demonstrated that GBLUP has the minimum predictive (0.380), while the predictive ability obtained from ERRBLUP is 0.402. The maximum predictive ability (0.565) was observed for sERRBLUP when the top 1 percent of interactions were selected based on their absolute effect sizes in relationship matrix determined from variable selection in environment 2.



**Fig. 2.2:** Comparison of predictive ability of GBLUP (open circle), ERRBLUP (black filled circle) and sERRBLUP for 95 to 0.1% of SNP interactions maintained in the model. Green dots represent the sERRBLUP predictive ability for prediction within environment 3, while black, blue and red dots represent the predictive ability with relationship matrices determined by variable selection in environments 1, 2 and 4.

It should also be noted that for purely additive markers the sERRBLUP model does not provide an increase in predictive ability.

## 2.4 Notes

1. Predictive ability obtained from sERRBLUP is higher than obtained from ERRBLUP and GBLUP

As in Martini *et al.* (2016) we found that prediction with ERRBLUP is higher than GBLUP, and that prediction with sERRBLUP systematically outperforms ERRBLUP. ERRBLUP can be seen as a special case of sERRBLUP (with  $\pi = 1$ ) and thus searching the parameter space of  $\pi$  for the optimum value will in most cases lead to an improved, but never to a reduced predictive ability of sERRBLUP compared to ERRBLUP.

2. Maximum predictive ability is obtained by selecting the optimum proportion of SNP interactions in sERRBLUP

The predictive ability of the sERRBLUP model is influenced by the selected subset of SNP interactions. As the full model is reduced by considering the top SNP interactions based on their estimated effects or estimated effect variances in the model, the predictive ability increases in sERRBLUP, while this can change to a decrease in predictive ability by a too strict selection of SNP interactions. The optimum proportion of interactions resulting in the maximum predictive ability in the considered simulated dataset was  $\pi = 0.01$  when selection was based on absolute effect sizes and  $\pi = 0.001$  when selection was based on effect variances. These proportions can be different in other data sets and need to be determined with an appropriate strategy based on cross validation.

3. sERRBLUP predictive ability is higher when SNP interactions are selected based on the absolute effect size rather than based on the estimated variance.

The predictive ability obtained from sERRBLUP model was found to be systematically higher when the same proportion of interactions were selected based on the absolute effect sizes compared to when interactions are selected based on estimated variances in wheat dataset. It should be noted that in the other datasets selecting the interactions based on estimated variances in sERRBLUP model can provide higher accuracy.

4. sERRBLUP computational load is less than in the model proposed by Martini (Martini *et al.*, 2016)

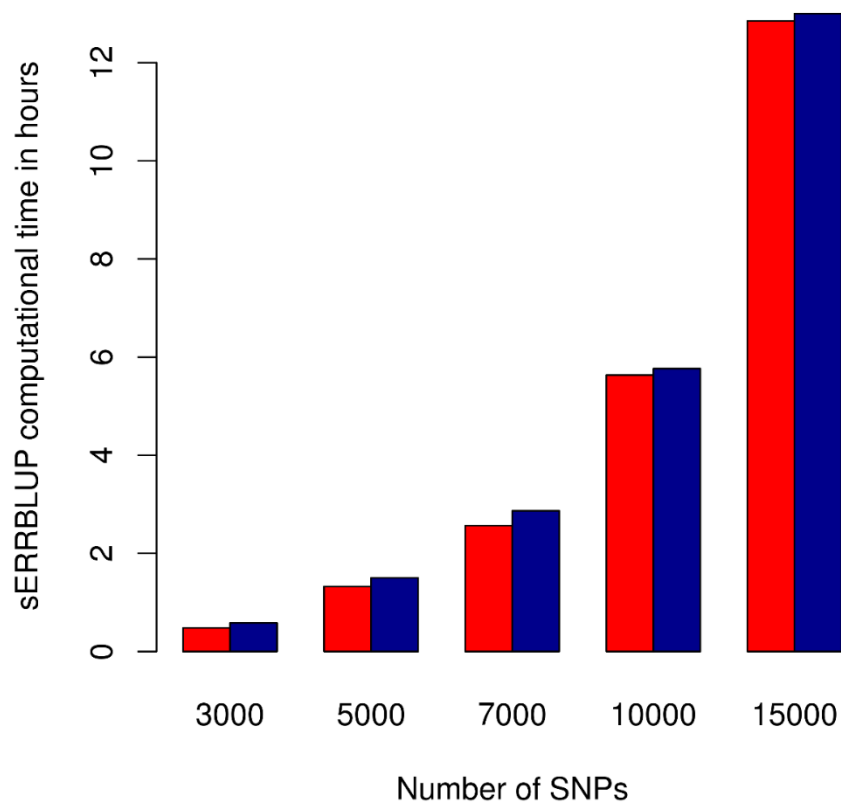
The computational time required for sERRBLUP model is independent of which proportion of interactions are included in the model. However, the computational time for the reduced epistasis model in Martini *et al.* (2016) model is dependent on the selected proportion of interactions, since in their method the full epistatic model is reduced in 5% steps, by which SNP interaction effects and corresponding relationship matrices are recalculated at each step in the remaining network. As it is demonstrated in Fig. 2.3 for different numbers of SNPs, computational time for sERRBLUP (red bars) is on average 10 percent less than Martini (Martini *et al.*, 2016) proposed method (blue bars) for the top 95 percent of interactions by using 15 cores. However, this difference in computational time get much higher by maintaining smaller proportions of interactions in the model. To illustrate this, the required computational time with the approach of Martini (Martini *et al.*, 2016) for maintaining the top 10 percent of interactions in the model is around 18 times more than what has been shown in Fig. 2.3 Computational time for sERRBLUP is on average around 3 times larger than the one for ERRBLUP. By increasing the number of SNPs sERRBLUP also requires a higher computational load, but still is feasible while the approach of Martini (Martini *et al.*, 2016) might not be feasible in a proper amount of time.

5. GBLUP, ERRBLUP or sERRBLUP

Among the three proposed genomic prediction models, sERRBLUP has the highest predictive ability. Although the sERRBLUP model provides the highest accuracy, its feasibility depends on the available the number of SNPs. In other words, if the number of SNPs grows to millions, this method also poses a considerable, eventually prohibitive computational load. Potential



strategies to overcome these limitations are to achieve a feature reduction by SNP pruning (Purcell *et al.*, 2007; Chang *et al.*, 2015) or the use of haplotype blocks (Pook *et al.*, 2019).



**Fig. 2.3:** Comparison of computational time for sERRBLUP by utilizing 15 cores (red bars) and the method of Martini (Martini *et al.*, 2016) with a reduced epistatic model (blue bars) for the top 95% of interactions maintenance in the model in different numbers of SNPs.



### **3 Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments**

This chapter contains the manuscript “Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments” which has been accepted in the journal TAG (Theoretical and Applied Genetics). In order to have uniform style in the thesis, the journal style is not used in this chapter.

This manuscript is the joint work of Elaheh Vojgani<sup>1\*</sup>, Torsten Pook<sup>1</sup>, Johannes W.R. Martini<sup>2</sup>, Armin C. Hölker<sup>3</sup>, Manfred Mayer<sup>3</sup>, Chris-Carolin Schön<sup>3</sup>, Henner Simianer<sup>1</sup> which is focused on developing epistasis models in the bivariate statistical framework and comparing the epistasis models in univariate and bivariate statistical settings for prediction across environments.

1: University of Goettingen, Center for Integrated Breeding Research, Animal Breeding and Genetics Group, Goettingen, Germany

2: International Maize and Wheat Improvement Center (CIMMYT), Texcoco, State of Mexico, Mexico

3: Plant Breeding, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

#### **Author contributions by EV**

EV analyzed the data, generated the pruned set of data, derived the results, wrote the initial manuscript and led the revision of the manuscript.

### 3.1 Key Message

The accuracy of genomic prediction of phenotypes can be increased by including the top ranked pairwise SNP interactions into the prediction model.

### 3.2 Abstract

We compared the predictive ability of various prediction models for a maize dataset derived from 910 doubled haploid lines from two European landraces (Kemater Landmais Gelb and Petkuser Ferdinand Rot), which were tested at six locations in Germany and Spain. The compared models were Genomic Best Linear Unbiased Prediction (GBLUP) as an additive model, Epistatic Random Regression BLUP (ERRBLUP) accounting for all pairwise SNP interactions, and selective Epistatic Random Regression BLUP (sERRBLUP) accounting for a selected subset of pairwise SNP interactions. These models have been compared in both univariate and bivariate statistical settings for predictions within and across environments. Our results indicate that modeling all pairwise SNP interactions into the univariate/bivariate model (ERRBLUP) is not superior in predictive ability to the respective additive model (GBLUP). However, incorporating only a selected subset of interactions with the highest effect variances in univariate/bivariate sERRBLUP can increase predictive ability significantly compared to the univariate/bivariate GBLUP. Overall, bivariate models consistently outperform univariate models in predictive ability. Across all studied traits, locations, and landraces, the increase in prediction accuracy from univariate GBLUP to univariate sERRBLUP ranged from 5.9 to 112.4 percent, with an average increase of 47 percent. For bivariate models, the change ranged from -0.3 to +27.9 percent comparing the bivariate sERRBLUP to the bivariate GBLUP, with an average increase of 11 percent. This considerable increase in predictive ability achieved by sERRBLUP may be of interest for “sparse testing” approaches in which only a subset of the lines/hybrids of interest is observed at each location.

**Keywords:** Genomic prediction, GBLUP, Multi-trait models, Epistasis, Interaction

### 3.3 Introduction

Genomic prediction of phenotypes has been widely explored for crops (Crosa *et al.*, 2010), livestock (Daetwyler *et al.*, 2013) and clinical research (de los Campos, Vazquez, *et al.*, 2013). Broad availability and cost effective generation of genomic data had a considerable impact on plant (Bernardo and Yu, 2007; de los Campos *et al.*, 2009; Crosa *et al.*, 2010, 2011; de Los Campos *et al.*, 2010; Pérez *et al.*, 2010) and animal breeding programs (de los Campos *et al.*, 2009; Hayes and Goddard, 2010; Daetwyler *et al.*, 2013). Genomic prediction relates a set of genome wide markers to the variability in the observed phenotypes and enables the prediction of phenotypes or genetic values of genotyped but unobserved material (Meuwissen *et al.*, 2001; Jones, 2012; Windhausen *et al.*, 2012). This approach has been positively evaluated in most major

crop and livestock species (Albrecht *et al.*, 2011; Daetwyler *et al.*, 2013; Desta and Ortiz, 2014) and is becoming a routine tool in commercial and public breeding programs (Stich and Ingheland, 2018). In plant breeding, phenotyping is one of the major current bottlenecks and the optimization or minimization of phenotyping costs within breeding programs is needed (Akdemir and Isidro-Sánchez, 2019). Therefore, the maximization of genomic prediction accuracy can be directly translated into reduced phenotyping costs (Akdemir and Isidro-Sánchez, 2019; Jarquin *et al.*, 2020).

Genomic selection and the corresponding prediction of breeding values is based on a covariance matrix describing the (additive) relationship between the considered individuals (Wolc *et al.*, 2011; Burgueño *et al.*, 2012). This matrix can be constructed from pedigree information, from marker information (VanRaden, 2007) or from a combination of pedigree and available genotypic information in a single step approach (Aguilar *et al.*, 2010; Legarra *et al.*, 2014). It has been broadly demonstrated that marker based relationship matrices enhance the reliability of breeding value estimation on average across traits and compared to pedigree based approaches (Meuwissen *et al.*, 2001; VanRaden, 2007; Hayes and Goddard, 2008; Crossa *et al.*, 2010). Since breeding values are additive by definition (Falconer and Mackay, 1996), the early development of prediction models exclusively accounted for the additive effects (Filho *et al.*, 2016).

Concerning additive models, genomic best linear unbiased prediction (GBLUP, Meuwissen *et al.*, 2001; VanRaden, 2007) is a widely-used linear mixed model (Da *et al.*, 2014; Rönnegård and Shen, 2016; Covarrubias-Pazarán *et al.*, 2018). Although various new approaches such as methods from the Bayesian alphabet (Gianola *et al.*, 2009) have been proposed, GBLUP remains the gold standard as new methods typically only perform marginally better, are less robust, require substantially more computing time and are more difficult to implement (Wang *et al.*, 2018). Daetwyler *et al.* (2010) showed that BayesB can yield higher accuracy than GBLUP for traits controlled by a small number of quantitative trait nucleotides, emphasizing that the genetic architecture of the trait has an important impact on which method may predict better (Wimmer *et al.*, 2013; Momen *et al.*, 2018). Moreover, the training set size was shown to play a role. For instance, human height prediction using BayesB and BayesC methods in a small reference population (<6,000 individuals) had no advantage over GBLUP. Only when increasing the size of the reference population (>6,000 individuals), these methods outperformed GBLUP (Karaman *et al.*, 2016).

Understanding how genetic variation causes phenotypic variation in quantitative traits is still a major challenge of contemporary biology. It has been proven that epistasis as a statistical interaction between two or more loci (Falconer and Mackay, 1996) contributes substantially to the genetic variation of quantitative traits (Wright, 1931; Carlborg and Haley, 2004; Hill *et al.*, 2008; Huang *et al.*, 2012; Mackay, 2014). On the one hand, models which incorporate epistasis have the potential to increase predictive ability (de Los Campos *et al.*, 2010; Hu *et al.*, 2011; Wang *et al.*, 2012; Mackay, 2014). On the other hand, accounting for epistasis by modeling interactions explicitly was considered to be computationally challenging (Mackay, 2014). In this context, the extended genomic best linear unbiased prediction (EG-BLUP), as an epistasis marker effect model

(Jiang and Reif, 2015; Martini *et al.*, 2016) and reproducing kernel Hilbert space regression (RKHS), as a semi-parametric model (Gianola *et al.*, 2006; Gianola and van Kaam, 2008; de Los Campos *et al.*, 2010) based on Gaussian kernel (Jiang and Reif, 2015) were proposed to reduce the computational load by constructing marker-based epistatic relationship matrices (Jiang and Reif, 2015; Martini *et al.*, 2016). RKHS has shown to be as good as (Jiang and Reif, 2015) or better than EG-BLUP (Martini *et al.*, 2017). While EG-BLUP is potentially beneficial for genomic prediction, its performance depends on the marker coding (Martini *et al.*, 2017, 2019). Moreover, it has been shown that the superiority of epistasis models over the additive GBLUP in terms of predictive ability may vanish when the number of markers increases (Schrauf *et al.*, 2020). Also, the Hadamard products of the additive genomic relationship matrices provide only an approximation for the interaction effect model based on interactions between different loci (Martini *et al.*, 2020), and more correcting factors are required for interactions of higher degree (Jiang and Reif, 2020).

Another downside of epistasis models is that, due to the high number of interactions, a large number of unimportant variables can be introduced into the model (Rönnegård and Shen, 2016). This ‘noise’ might prevent a gain in predictive ability. In this regard, Martini *et al.* (2016) showed that selecting just a subset of the largest epistatic interaction effects has the potential to improve predictive ability. Therefore, reducing the full epistasis model to a model based on a subnetwork of ‘most relevant’ pairwise SNP interactions may be beneficial for prediction performance (Martini *et al.*, 2016).

In addition to the extension from additive effect models to models including epistatic interactions, genomic prediction models can be extended from univariate models to multivariate models. Univariate models consider each trait separately, while multivariate models treat several traits simultaneously with the objective to exploit the genetic correlation between them to increase predictive ability. Multivariate models which have been first proposed for the prediction of genetic values by Henderson and Quaas (1976) were shown to be potentially beneficial for prediction accuracy when the correlation between traits is strong (He *et al.*, 2016; Covarrubias-Pazaran *et al.*, 2018; Schulthess *et al.*, 2018; Velazco *et al.*, 2019). A situation of dealing with multiple environments can also be considered in the framework of a multivariate model by simply considering a trait-in-environment combination as another correlated trait. This is considered as the multi-environment model which is usually employed to assess  $G \times E$  interaction (Montesinos-López *et al.*, 2016; Hassen *et al.*, 2018) and captures the differences in genotypes’ performances from one environment to the other as one of the breeders’ major challenges in plant breeding (Kang and Gorman, 1989). Prediction accuracy could be potentially enhanced through borrowing information across environments by utilizing multi-environment models (Burgueño *et al.*, 2012). In addition to multi-environment models, Martini *et al.* (2016) showed that the predictive ability of EG-BLUP as a univariate model can be increased in one environment by variable selection in the other environment under the assumption of a relevant correlation of phenotypes in different environments. This, however, was only demonstrated with a data set of limited size and especially a limited set of markers and, thus, marker interactions.

In the present study, we use a data set of doubled haploid lines derived from two European landraces, to investigate how beneficial the use of subnetworks of interactions in the proposed sERRBLUP framework can be. This was compared in the context of univariate and bivariate models. We assess the optimum proportion of SNP interactions to be kept in the model in the variable selection step. The development of efficient selection strategies which could mitigate costly and time consuming phenotyping of a large number of selection candidates in multiple environments has been a particular focus of research in plant breeding (Jarquin *et al.*, 2020). A successful application of our models may reduce the cost of phenotyping by reducing the number of test locations per line.

## 3.4 Materials and Methods

### 3.4.1 Data used for analysis

We used a set of 501 / 409 doubled haploid lines of the European maize landraces Kemater Landmais Gelb / Petkuser Ferdinand Rot genotyped with 501,124 markers using the Affymetrix® Axiom Maize Genotyping Array (Unterseer *et al.*, 2014), out of which 471 and 402 lines were phenotyped for Kemater (KE) and Petkuser (PE), respectively. The performance of the lines has been evaluated by ten separate  $10 \times 10$  lattice designs in four German locations and five separate  $10 \times 10$  lattice designs in two Spanish locations with two replicates. For more details see Hölker *et al.* (2019).

The lines were phenotyped in 2017 for a series of traits in six different environments which were Bernburg (BBG, Germany), Einbeck (EIN, Germany), Oberer Lindenhof (OLI, Germany), Rogenstein (ROG, Germany), Golada (GOL, Spain) and Tomeza (TOM, Spain).

The descriptions of the phenotypic traits, comprising early vigour and mean plant height of three plants of the plot at three growth stages (EV\_V3, EV\_V4, EV\_V6, PH\_V4, PH\_V6, PH\_final), days from sowing until female flowering (FF) and root lodging (RL) are given in the supplementary (Table S. 3.1), together with the number of phenotyped lines per location, phenotypic means, standard deviations, and maximum and minimum values. To correct for spatial structure and population effects, Best Linear Unbiased Estimations (BLUEs) were used as input for all considered prediction models. The interested reader is referred to Hölker *et al.* (2019) for details on the correction procedure and the detailed description of the considered traits. E.g., the trait “growth stage V4” indicates the growth stage at which four leaf collars are fully developed (Abendroth *et al.*, 2011). In our study, we chose PH\_V4 as the main trait for evaluating and illustrating our methods, since it is a relevant metric quantitative trait for early plant development which is suitable for testing our methods. The phenotypic correlations of PH\_V4 across all environments are provided in Table 3.1.

Among the phenotypic traits, root lodging (RL) and female flowering (FF) were not phenotyped in all the environments: RL was only scored in BBG, ROG, OLI and EIN, and FF was phenotyped in all environments except GOL.

**Table 3.1:** Phenotypic correlation across all environments for the trait PH\_V4 in KE (blue numbers above diagonal) and PE (red numbers below diagonal) which are highly significant ( $p$ -values < 0.001).

Location	BBG	EIN	OLI	ROG	GOL	TOM
BBG	-	0.82	0.66	0.67	0.69	0.58
EIN	0.78	-	0.71	0.77	0.75	0.66
OLI	0.60	0.66	-	0.71	0.65	0.50
ROG	0.62	0.69	0.65	-	0.70	0.58
GOL	0.55	0.59	0.46	0.51	-	0.69
TOM	0.57	0.68	0.57	0.58	0.54	-

### 3.4.2 Quality control, coding and imputing

As we would not expect any heterozygous calls in DH material, all heterozygous calls were set to missing. Genotype calls were coded according to the allele counts of the B73 AGPv4 reference sequence (Jiao *et al.*, 2017) (0 = homozygous for the reference allele, 2 = homozygous for the alternative allele). Imputation of missing values was performed separately for each landrace, using BEAGLE version 4.0 with parameters buildwindow=50, nsamples=50 (Browning and Browning, 2007; Pook *et al.*, 2020). For the remaining heterozygous calls, the DS (dosage) information of the BEAGLE output was used and genotyped with DS <1 were set to 0 and DS >= 1 to 2.

### 3.4.3 Linkage disequilibrium pruning

Linkage disequilibrium based SNP pruning with PLINK v1.07 was used to generate a subset of SNPs which are in approximate linkage equilibrium with each other. The parameters: indep 50 5 2 were used, in which 50 is the window size in SNPs, 5 is the number of SNPs to shift the window at each step and 2 is the variance inflation factor  $VIF = 1/(1 - r^2)$ , where  $r^2$  is the squared correlation between single SNPs and linear combinations of all SNPs in the window. All variants in the 50 SNP window which had a  $VIF > 2$  were removed. Then, the window was shifted 5 SNPs forward and the procedure was repeated (Purcell *et al.* 2007; Chang *et al.* 2015).



In our study, LD pruning was done separately for each landrace, resulting in data panels containing 25'437 SNPs for KE and 30'212 SNPs for PE.

### 3.4.4 Univariate statistical models for phenotype prediction

We used three different statistical models to predict phenotypes, which are all based on a linear mixed model (Henderson 1975). We assume that we have in total  $n$  lines which are genotyped, and phenotypes are available for a subset of  $n_1$  lines. These  $n_1$  lines are used to train the model and missing phenotypes for the remaining  $n_2 = n - n_1$  lines are predicted by using the genotypes of these lines. The basic univariate model is

$$\mathbf{y} = \mathbf{1}\mu + (\mathbf{I} \ \mathbf{O})\mathbf{g} + \boldsymbol{\epsilon},$$

where  $\mathbf{y}$  is an  $n_1 \times 1$  vector of phenotypes,  $\mathbf{1}$  is an  $n_1 \times 1$  vector with all entries equal to 1,  $\mu$  is a scalar fixed effect,  $\mathbf{I}$  is an identity matrix of dimension  $n_1 \times n_1$  and  $\mathbf{O}$  is a matrix of dimension  $n_1 \times n_2$  of zeros. The design matrix  $(\mathbf{I} \ \mathbf{O})$  is the  $n_1 \times (n_1 + n_2)$  matrix resulting from the concatenation of  $\mathbf{I}$  and  $\mathbf{O}$ . Moreover,  $\mathbf{g} \sim N(0, \boldsymbol{\Gamma}\sigma_g^2)$  is an  $n \times 1$  vector of random genetic effects, and  $\boldsymbol{\epsilon} \sim N(0, \mathbf{I}\sigma_\epsilon^2)$  is a random error vector, where  $\boldsymbol{\Gamma}$  and  $\mathbf{I}$  are the respective dispersion matrices and  $\sigma_g^2$  and  $\sigma_\epsilon^2$  are the corresponding variance components.

With this model, the population mean and the genetic effects  $\mathbf{g}$  for all lines, including those without phenotypes, are estimated using

$$\begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{g}}_1 \\ \hat{\mathbf{g}}_2 \end{bmatrix} = \begin{bmatrix} n_1 & \mathbf{1}' & \mathbf{0} \\ \mathbf{1} & \mathbf{I} + \lambda\boldsymbol{\Gamma}^{11} & \lambda\boldsymbol{\Gamma}^{12} \\ \mathbf{0} & \lambda\boldsymbol{\Gamma}^{21} & \lambda\boldsymbol{\Gamma}^{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}'\mathbf{y} \\ \mathbf{y} \\ \mathbf{0} \end{bmatrix}, \quad (\text{eq. 1})$$

where  $\lambda = \sigma_\epsilon^2 / \sigma_g^2$ ,  $\boldsymbol{\Gamma}^{-1} = \begin{bmatrix} \boldsymbol{\Gamma}^{11} & \boldsymbol{\Gamma}^{12} \\ \boldsymbol{\Gamma}^{21} & \boldsymbol{\Gamma}^{22} \end{bmatrix}$  and  $\mathbf{g} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix}$  and the indices pertain to the subset of individuals with (index 1) or without (index 2) phenotypes, respectively.

With these estimates, the phenotypes for the set of unphenotyped individuals can be predicted as  $\hat{\mathbf{y}}_2 = \mathbf{1}_2\hat{\mu} + \hat{\mathbf{g}}_2$ , where  $\hat{\mathbf{y}}_2$  is the  $n_2 \times 1$  vector of predicted phenotypes and  $\mathbf{1}_2$  is an  $n_2 \times 1$  vector of ones.

For  $n = n_1$  and  $n_2 = 0$  the solution of eq. 1 provides estimates of genetic effects when all lines are phenotyped and genotyped.

### 3.4.5 Bivariate statistical models for phenotype prediction

Besides univariate models, we also used bivariate models, where the two variables represent the same trait measured in two different environments.

The basic bivariate model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\mu} + \mathbf{Z}\mathbf{g} + \mathbf{e}$$

or, in more detail,

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{1}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix}, \quad (\text{eq. 2})$$

where,  $\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$  is the phenotype vector of length  $m = m_1 + m_2$  for environment 1 ( $m_1$ ) and 2 ( $m_2$ ),  $\mathbf{1}_1$  and  $\mathbf{1}_2$  are respectively  $m_1 \times 1$  and  $m_2 \times 1$  vectors with all entries equal to 1,  $\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$  is the vector of population means for environment 1 and 2,  $\mathbf{I}_1$  and  $\mathbf{I}_2$  are identity matrices of dimension  $m_1 \times m_1$  and  $m_2 \times m_2$ , respectively assigning genomic values to phenotypes. Moreover,  $\begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix}$  is the vector of random genomic values which is assumed to have a multivariate normal distribution with mean zero and variance  $\mathbf{G} = \mathbf{H} \otimes \boldsymbol{\Gamma}$ , where  $\mathbf{H} = \begin{bmatrix} \sigma_{g_1}^2 & \sigma_{g_{12}} \\ \sigma_{g_{12}} & \sigma_{g_2}^2 \end{bmatrix}$ ,  $\boldsymbol{\Gamma}$  is the dispersion matrix of genetic effects and  $\otimes$  is the Kronecker product.  $\begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix}$  is the vector of random errors which is assumed to have a multivariate normal distribution with mean zero and variance  $\mathbf{R} = \mathbf{R}_0 \otimes \mathbf{I}$ , where  $\mathbf{R}_0 = \begin{bmatrix} \sigma_{e_1}^2 & \sigma_{e_{12}} \\ \sigma_{e_{12}} & \sigma_{e_2}^2 \end{bmatrix}$ .  $\sigma_{g_i}^2$  and  $\sigma_{e_i}^2$  represent the genetic and residual variance of environment  $i = 1, 2$ , and  $\sigma_{g_{12}}$  and  $\sigma_{e_{12}}$  are the genetic and residual covariance between the environment 1 and 2 (Guo *et al.*, 2014). In this model, the phenotypes have to be ordered in the same way in both environments. In case the number of observations in environment 1 and environment 2 is not identical (i.e. in general terms  $m_1 \neq m_2$ ) or different lines are considered in the model, the incidence matrices have to be adapted accordingly.

With this model, the vector of environment specific population means and the vector of genetic effects for all lines are estimated using the standard mixed model equations

$$\begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}$$

In analogy to the procedure described in the univariate setting, we consider a setting in which the last  $l$  phenotypes for environment 2 are masked and predicted from all observations in environment 1 and the first  $k = m_2 - l$  non-masked observations in environment 2.

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_k \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{1}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_k \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_l \end{bmatrix} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_{2k} \\ \mathbf{g}_{2l} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_k \\ \mathbf{0} \end{bmatrix}$$

From the solutions obtained with this model, the phenotypes for the set of unphenotyped individuals in environment 2 can be predicted as  $\hat{\mathbf{y}}_l = \mathbf{1}_l \hat{\mu}_2 + \hat{\mathbf{g}}_{2l}$ , where  $\hat{\mathbf{y}}_l$  is the  $l \times 1$  vector of predicted phenotypes and  $\mathbf{1}_l$  is an  $l \times 1$  vector of ones.

The three models compared in this study only differ in the choice of the dispersion matrix  $\mathbf{\Gamma}$  of the genetic effects.

### 3.4.6 Model 1: Genomic Best Linear Unbiased Prediction (GBLUP)

In this additive model, we use as  $\mathbf{\Gamma}$  the genomic relationship matrix which is calculated according to VanRaden (2008) as

$$\mathbf{\Gamma}_{VR} = \frac{(\mathbf{M} - \mathbf{P})(\mathbf{M} - \mathbf{P})'}{2 \cdot \sum_{i=1}^m (p_i(1 - p_i))'}$$

where  $\mathbf{M}$  is the  $n \times m$  marker matrix which gives  $m$  marker values for  $n$  lines under the assumption of having  $n$  genotyped lines in total.  $\mathbf{P}$  is a matrix of equal dimension as  $\mathbf{M}$  with  $2 \cdot p_i$  in the  $i^{th}$  column, and  $p_i$  is the allele frequency of the minor allele of SNP  $i$ .

### 3.4.7 Model 2: Epistatic Random Regression BLUP (ERRBLUP)

This model accounts for all possible SNP interactions in the prediction model. With  $m$  markers and fully inbred lines, we have two possible genotypes at a single locus, i.e. 0 or 2 when coded as the counts of the minor allele. For each pair of loci, we have four different possible genotype combinations: {00, 02, 20, 22}. The total number of pairs of loci is  $\frac{m \times (m+1)}{2}$  if we allow for interaction of a locus with itself. Since for each of these pairs we have four possible genotype combinations, the total number of combinations to be considered as dummy variables is

$$m^* = 4 \times \frac{m \times (m+1)}{2} = 2m \times (m + 1).$$

We define a marker combination matrix  $\mathbf{M}^*$  of dimension  $n \times m^*$  whose element  $i, j$  is 1 if genotype combination  $j$  is present in individual  $i$  and is 0 otherwise. We further define for column  $i$  of this matrix the average value  $p_i^*$ , giving the frequency of the respective genotype combination in the population, and a matrix  $\mathbf{P}^*$  being of equal dimension as  $\mathbf{M}^*$  with  $p_i^*$  in the  $i^{th}$  column.

Then, the relationship matrix based on all SNP interactions was calculated according to VanRaden (2008) as

$$\mathbf{\Gamma}_{ERR} = \frac{(\mathbf{M}^* - \mathbf{P}^*)(\mathbf{M}^* - \mathbf{P}^*)'}{\sum_{i=1}^{m^*} (p_i^*(1 - p_i^*))}$$

and this matrix was used in ERRBLUP as dispersion matrix for the genetic effects, which now are based on epistatic interaction effects. It should be noted that including the interaction of each locus with itself replaces the additive effect, so that it is not necessary to use a model that separately accounts for additive and epistatic effects. This model had been introduced earlier as “categorical epistasis model” (Martini *et al.*, 2017).

### 3.4.8 Model 3: selective Epistatic Random Regression BLUP (sERRBLUP)

sERRBLUP is based on the same approach as ERRBLUP, but here the  $\mathbf{\Gamma}$ -matrix is constructed from a selected subset of genotype interactions. We decided to use those interactions with the highest estimated marker effects variances. Selection based on highest absolute effects (as used by Martini *et al.* (2016) in the framework of the EGBLUP epistasis model) was also considered, but lead to similar to slightly worse results. For this, it was necessary to backsolve interaction effects  $\hat{\mathbf{t}}$  and effects variances  $\hat{\sigma}^2$  from the ERRBLUP model using (Mrode, 2014)

$$\hat{\mathbf{t}} = \frac{\widehat{\sigma}_g^2}{\sum_{i=1}^{m^*} (p_i^*(1-p_i^*))} (\mathbf{M}^* - \mathbf{P}^*)' \left( \widehat{\sigma}_g^2 \mathbf{\Gamma}_{ERR} + \widehat{\sigma}_\epsilon^2 \mathbf{I} \right)^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}),$$

$$\hat{\sigma}^2 = (\hat{\mathbf{t}} \circ \hat{\mathbf{t}}) 2\mathbf{P}^*(\mathbf{1} - \mathbf{P}^*),$$

with  $\circ$  denoting the Hadamard product.

After estimating SNP interaction effects in  $\hat{\mathbf{t}}$  and effects variances in  $\hat{\sigma}^2$ , we selected those interactions whose absolute estimated effects or effect variances were in the top  $\pi = 0.05, 0.01, 0.001, 0.0001, 0.00001$  or  $0.000001$  proportion of all interactions, respectively. These proportions were chosen since it was observed in preliminary analyses that they cover the most relevant range. For each of these subsets, we generated reduced matrices  $\mathbf{M}_\pi^*$  and  $\mathbf{P}_\pi^*$  of dimension  $n \times \pi m^*$ , containing only those columns of  $\mathbf{M}^*$  and  $\mathbf{P}^*$  pertaining to the selected subset of genotype interactions, and then set up the dispersion matrix in analogy to VanRaden (2008) as

$$\mathbf{\Gamma}_{sERR} = \frac{(\mathbf{M}_\pi^* - \mathbf{P}_\pi^*)(\mathbf{M}_\pi^* - \mathbf{P}_\pi^*)'}{\sum_{i=1}^{\pi m^*} (p_{\pi i}^*(1 - p_{\pi i}^*))},$$

where  $p_{\pi i}^*$  are the mean frequencies of the selected genotype combinations.

Note here that even for the univariate model, information from another environment is used for the prediction, namely for variable selection and the definition of  $\mathbf{\Gamma}_{sERR}$ . However, having used the information from another environment to define the subset of interactions and to derive the relationship matrix  $\mathbf{\Gamma}_{sERR}$ , the actual prediction is within the considered environment from the training to the test set.

We used the miraculix package (Schlather, 2020) to efficiently calculate  $\mathbf{\Gamma}_{ERR}$ ,  $\hat{\mathbf{t}}$  and  $\mathbf{\Gamma}_{sERR}$ .

### 3.4.9 Assessment of predictive ability via 5-fold random cross validation with 5 replicates

In a 5-fold cross validation, the original sample is randomly partitioned into five subsamples of equal size. Out of the five subsamples, each subsample is subsequently considered as the test set for validating the model, and the remaining four subsamples are considered as training data. The training set is used to predict the test set. By this, all observations are used for both training and testing and each observation is only used once for testing (Utz *et al.*, 2000). We repeated the cross-validation procedure 5 times, using random partitions of the original sample. The results of the 25 repetitions were then averaged (Erbe *et al.*, 2010). We used the Pearson correlation between the predicted genetic value and the observed phenotype in the test set as the measure for predictive ability. In our study, predictive ability was assessed for PE and KE for all phenotypic traits separately. In addition, the trait’s prediction accuracy was calculated by dividing the obtained predictive ability by the square-root of the respective trait’s heritability (Dekkers, 2007). The numbers of KE and PE lines which are available for all combinations of environments are summarized in Table 3.2. For some traits these numbers can be smaller or even zero for some environment combinations.

**Table 3.2:** Number of KE (blue numbers above diagonal) and PE (red numbers below diagonal) phenotyped lines in each pair of environments for trait PH\_V4.

Location	BBG	EIN	OLI	ROG	GOL	TOM
BBG	393/461	461	441	461	200	200
EIN	393	393/462	441	461	201	201
OLI	390	390	390/441	441	182	181
ROG	390	390	389	390/461	200	200
GOL	195	195	195	195	204/211	209
TOM	195	195	195	195	204	204/210

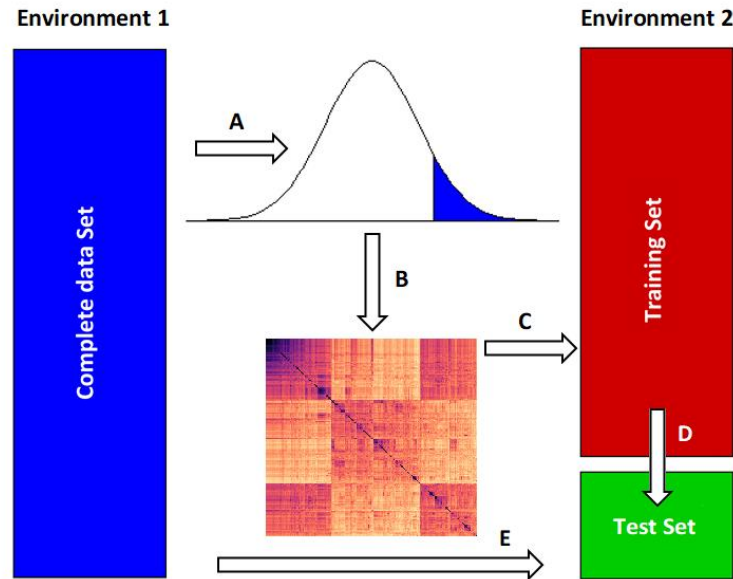
We evaluated our univariate and bivariate models as follows:

### 3.4.10 Assessment of GBLUP, ERRBLUP and sERRBLUP predictive abilities

The univariate GBLUP and ERRBLUP within environments were evaluated by training the model in the same environment as the test set was sampled from.

The basic strategy for univariate and bivariate sERRBLUP across environments is illustrated in Fig. 3.1: first, all pairwise SNP interaction effects and their variances are estimated from all data in environment 1 and effects are ordered either by absolute effect size or effect variance (A). Next, an epistatic relationship matrix for all lines is constructed from the top ranked subset of interaction effects (B). Then, this matrix is used in environment 2 (C) to predict phenotypes of the test set (green) from the respective training set (red) (D). This approach henceforth is termed ‘sERRBLUP across environments’. In the case of bivariate sERRBLUP both the full data panel from environment 1 and the training set from environment 2 are used in a bivariate prediction model.

The basic strategy for bivariate GBLUP and ERRBLUP can also be illustrated in Fig. 3.1 when the model is trained jointly on the complete dataset of environment 1 (E) and the training set of environment 2 (D). The test set of environment 2 is then predicted, using as dispersion matrix for the genetic effects either  $\Gamma_{VR}$  or  $\Gamma_{ERR}$ .



**Fig. 3.1:** Basic scheme of uni- and bivariate sERRBLUP across environments. All pairwise SNP interaction effects and their variances are estimated from all data in environment 1, and effects are ordered either by absolute effect size or effect variance (A). Then, an epistatic relationship matrix for all lines is constructed from the top ranked subset of interaction effects (B) which in the univariate model is used in environment 2 (C) to predict phenotypes of the test set (green) from the respective training set (red, D). In the bivariate model, this information is combined with the complete data from environment 1 (blue, E) to predict the test set.

### 3.4.11 Use of multiple environments jointly

In addition to considering each environment separately, we used the average of all environments, except the current target environment, as an additional environment. This was considered for univariate sERRBLUP and all bivariate models.

### 3.4.12 Estimation of variance and covariance components

Since we aimed at estimating variance components in each replicate of the cross-validation from the training data, but variance component estimation with ASREML has a certain risk of non-convergence in particular in models with a high number of parameters such as the models proposed here. Therefore, we needed to specify a strategy to deal with such cases in an automated manner. In univariate analyses, variance components were estimated using EMMREML (Akdemir and Godfrey, 2015) in each run of a 5-fold cross validation based on the training set. In bivariate analyses, the variance components were estimated using ASReML-R (Butler *et al.*, 2018). In the bivariate ERRBLUP and sERRBLUP models, the genetic and residual variance and covariance were estimated first from the full data set in a bivariate ASReML-R model for each combination of environments in each trait. If the estimation of variance components didn't converge after 100 iterations, then the computation was stopped and the genetic and residual variance and covariance estimates at the last iteration (100) were extracted. These estimates were defined as the initial starting values of the bivariate ASReML-R model in each run of a 5-fold cross validation, followed by a re-estimation of the variance and covariance components based on the training set in the cross validation. If the estimation of variance components did not converge at 50 iterations in each fold, the pre-estimated variance and covariance components based on the full dataset, which was defined as the initial start values of the model, were used as fixed values, so that the breeding values were estimated based on these pre-estimated parameters. It was verified from converged estimates that variance and covariance components estimated from the training set deviated only little from the variances and covariances from the full set (see Fig. S. 3.1). Also, the mean result obtained from just the converged replicates and the mean results of all replicates including the ones where variance and covariance components were fixed were rather similar (Fig. S. 3.2), only when the majority (>20) of replicates failed to converge, substantial random fluctuation was observed. Thus, we argue that this strategy appears justifiable, but still the number of cases where estimates did not converge in 5-fold cross validation with 5 replicates and the combinations whose pre estimation of variance components also did not converge in 100 iterations are detailed in the supplementary (Table S. 3.2 - Table S. 3.9).

## 3.5 Results

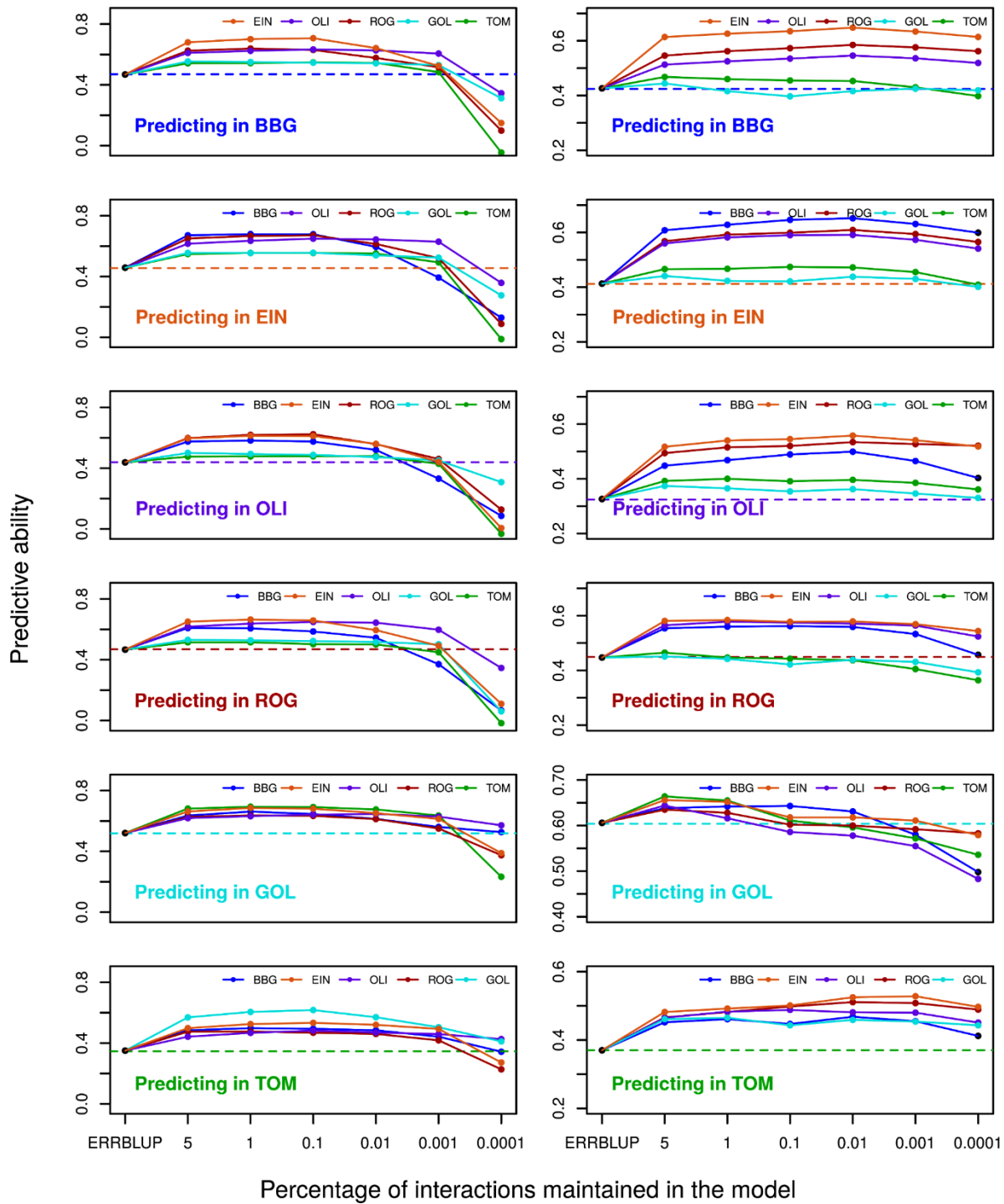
Predictive abilities of univariate sERRBLUP across environments compared to univariate ERRBLUP and univariate GBLUP within environments for the trait PH\_V4 are shown in Fig. 3.2 for KE and PE. Univariate GBLUP within the environment is used as a reference and is compared to results obtained with univariate ERRBLUP within environments and univariate sERRBLUP when the top 5, 1, 0.1, 0.01, 0.001 and 0.0001 percent of pairwise SNP interactions are maintained in the model. Fig. 3.2 shows that the predictive abilities of univariate GBLUP and univariate ERRBLUP within the environment are almost identical (the highest deviation observed was 0.004). A considerable increase in predictive ability was observed when the top 1 or 0.1 percent of SNP interactions, selected based on their effect variances, were kept in the univariate

sERRBLUP model. A more stringent selection, i.e. by considering only the top 0.01, 0.001 and 0.0001 percent of SNP interactions in the model, often led to a reduction in predictive ability, such that for the most stringent selection of 0.001 and 0.0001 percent, the predictive ability was sometimes even below the univariate GBLUP reference. This pattern is observed across all environments and is more pronounced in KE than PE. Results for the other traits are given in the Supplementary (Fig. S. 3.3 - Fig. S. 3.9). In this study, estimated effect variances were identified as the best selection criteria in sERRBLUP, since sERRBLUP predictive abilities were observed to be more robust when the selection of pairwise SNP interaction was based on the effect variances compared to absolute effect sizes, especially when the top 0.001 and 0.0001 percent of interactions are maintained in the model (Fig. S. 3.10 - Fig. S. 3.11). In addition, the maximum predictive ability obtained from univariate sERRBLUP are almost identical when selecting SNP interactions based on absolute effect sizes or effect variances for both KE and PE (Fig. S. 3.12).

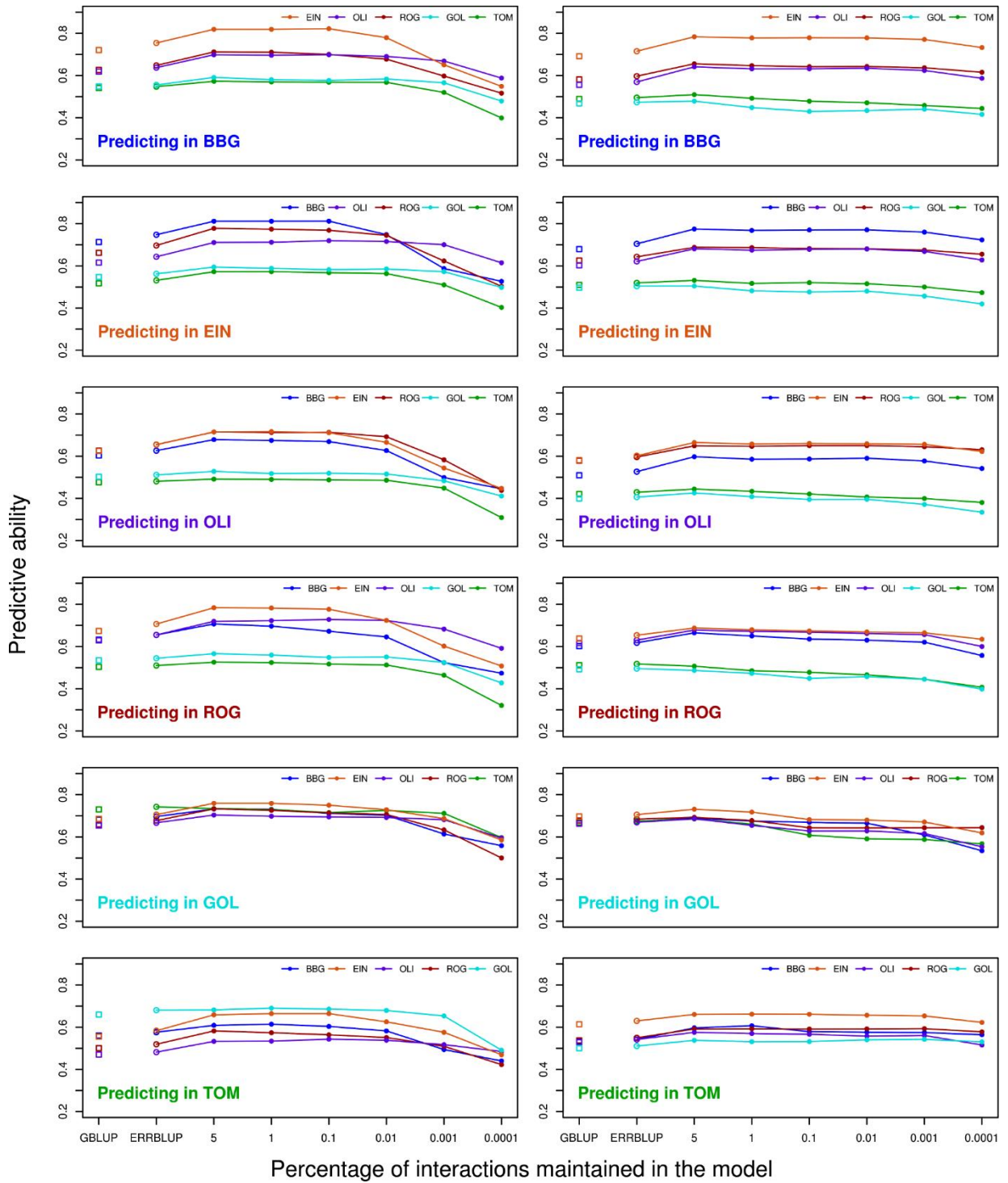
In the context of univariate models, we also investigated the predictive ability of univariate sERRBLUP when the variable selection was based on the training set from the same environment as the test set. This was exemplarily done within Bernburg for the trait PH\_V4 (Fig. S. 3.13), illustrating that the predictive ability obtained from univariate sERRBLUP is marginally higher than univariate GBLUP only when the top 0.01 percent of interactions are kept in the model. When the selection of effects is too strict, with only 0.001 percent of interactions used, the predictive ability of univariate sERRBLUP within Bernburg is smaller than the one obtained with GBLUP, especially if the selection is based on effect sizes.

The predictive abilities of bivariate GBLUP, ERRBLUP and sERRBLUP when SNP interactions were selected based on estimated effect variances are compared for trait PH\_V4 in KE and PE in Fig. 3.3. Fig. 3.3 shows that the bivariate ERRBLUP increases the predictive ability slightly compared to bivariate GBLUP with the maximum absolute increase of 0.03 in KE and 0.02 in PE across all environments' combinations. A considerable increase in predictive ability is obtained in bivariate sERRBLUP mostly when the top 5 or 1 percent of interactions are maintained in the model. However, the bivariate sERRBLUP predictive abilities decrease dramatically for too stringent selection of pairwise SNP interactions such as 0.01, 0.001 or 0.0001 percent. Moreover, the reduction in predictive ability with too stringent factor selection is more severe for KE than for PE. This pattern is observed for the majority of environments for both landraces and the results for other traits are shown in the supplementary (Fig. S. 3.14 - Fig. S. 3.20).





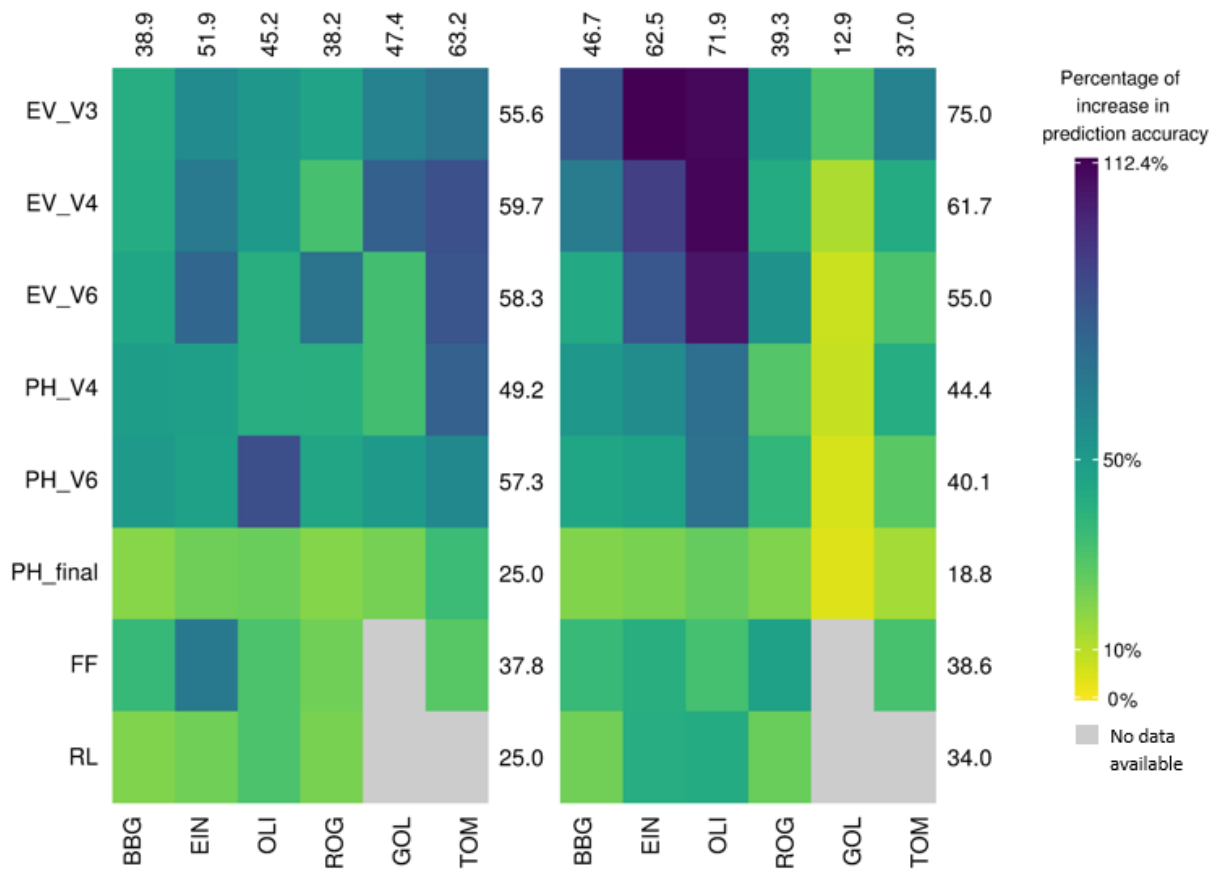
**Fig. 3.2:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments (solid colored lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH\_V4. In each panel, the solid lines' color indicates the environment in which the relationship matrices were determined by variable selection.



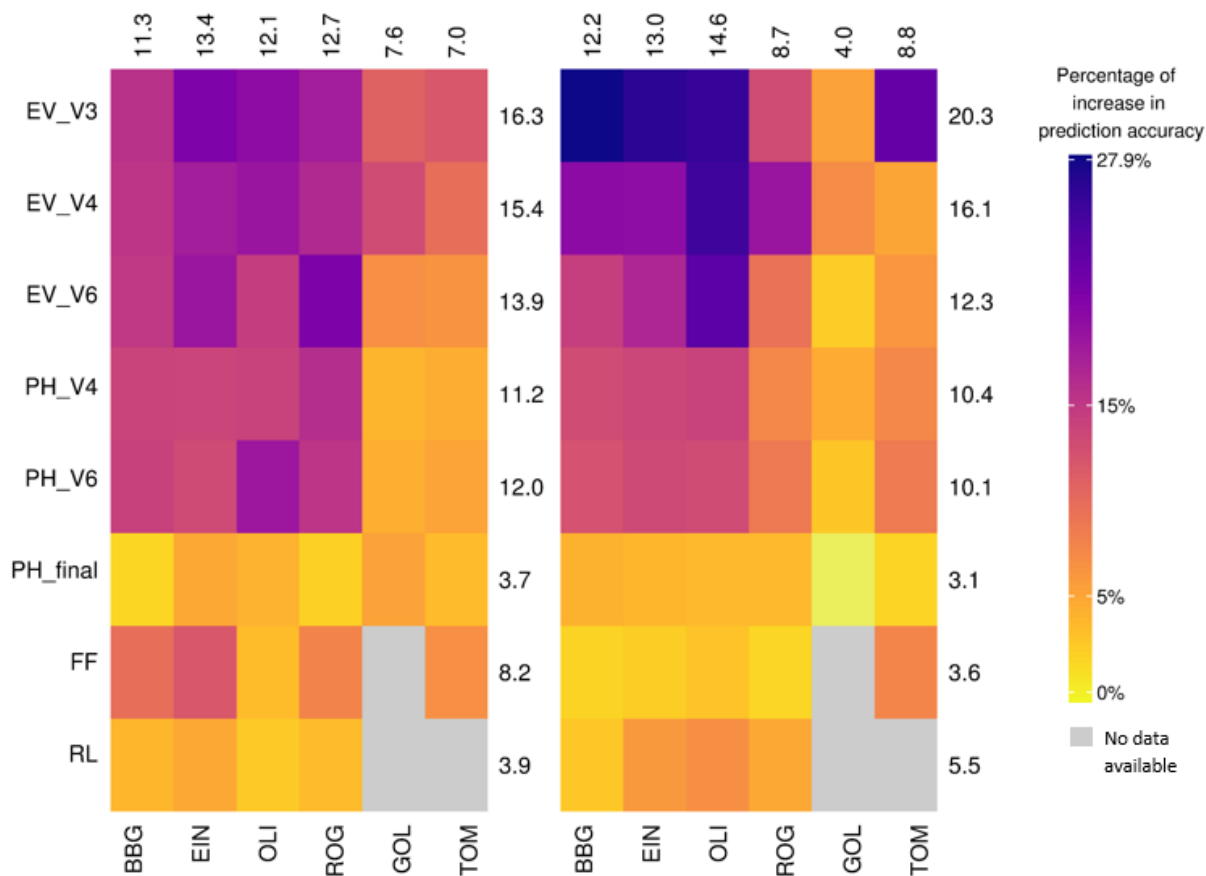
**Fig. 3.3:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH\_V4. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.

The relative increase in prediction accuracy of the best univariate sERRBLUP across environments compared to univariate GBLUP within environments for all traits and all locations is shown in form of a heat map in Fig. 3.4 for both landraces. The maximum relative increase in prediction accuracy among all traits and all environments in KE is 85.6 percent (PH\_V6 in OLI) and in PE it is 112.4 percent (EV\_V3 in EIN). Those highest increases in accuracy were found in traits and environment combinations where the univariate GBLUP prediction accuracy was particularly low. An increase is observed in each studied trait by location combination, with the smallest increase in both landraces for PH\_final in BBG (20.1 percent in KE) or in GOL (5.9 percent in PE). In general, both plots in Fig. 3.4 demonstrate that for the majority of traits and environments, there is more than a 30 percent increase in prediction accuracy from univariate GBLUP within environments to the best univariate sERRBLUP across environments. The average increase across all combinations in KE is 47.1 percent and in PE is 46.7 percent. Note that this increase is somewhat inflated as a single GBLUP accuracy is compared against the best prediction from a set of various models (environment / selection proportions). However, even when using a set environment and a fixed proportion of interactions, there are still substantial gain. Exemplary, EIN with a proportion of 0.1 still lead to an increase of 43.1 percent in KE and 36.9 percent in PE (Fig. S. 3.21). The choice of EIN was made as it had the highest number of phenotyped lines (Table S. 3.1), while 0.1 in general led to stable models. Results using any other location or reasonable choice of the share of included interactions were very similar. The absolute increase in prediction accuracy is also shown as a heat map in Supplementary Fig. S. 3.22, which indicates the average absolute increase of 0.204 in KE and 0.181 in PE.

Fig. 3.5 also shows the relative increase in prediction accuracy from the best bivariate GBLUP to the best bivariate sERRBLUP for all traits and all locations. The maximum increase in prediction accuracy among all traits and all environments is 21.1 percent (EV\_V6 in ROG) in KE and 27.9 percent (EV\_V3 in BBG) in PE. There is an increase across all studied traits in all environments except for the trait PH\_final in PE which shows a relative decrease of 0.3 percent. The minimum increase in prediction accuracy in KE was also observed for PH\_final (1.7 percent). In general, Fig. 3.5 shows that the relative increase in prediction accuracy from the best bivariate GBLUP to the best bivariate sERRBLUP is more than 7 percent for the majority of trait by location combinations in both landraces with an average increase of 10.9 percent in KE and 10.5 in PE across all combinations. The absolute increase in prediction accuracy of bivariate models is also shown as a heat map in supplementary (Fig. S. 3.23) indicating an average absolute increase of 0.1 across all traits, environment combinations, and landraces.



**Fig. 3.4:** Percentage of increase in prediction accuracy from univariate GBLUP within environments to the maximum prediction accuracy of univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances in KE (left side) and in PE (right side). The average percentage of increase in prediction accuracy for each trait and environments are displayed in rows and columns, respectively.

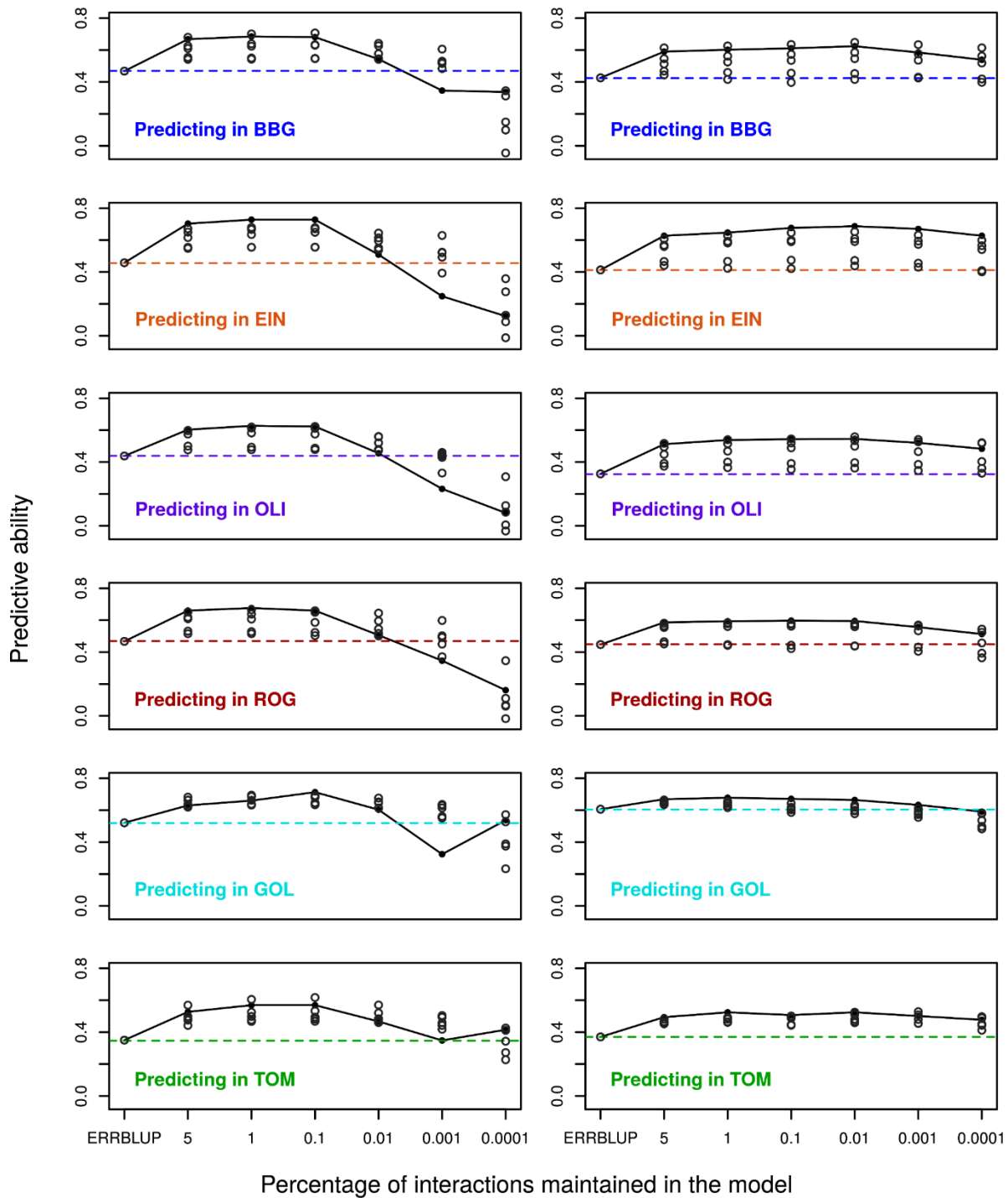


**Fig. 3.5:** Percentage of increase in prediction accuracy from the maximum bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP when the SNP interaction selections are based on estimated effects variances in KE (left side) and in PE (right side). The average percentage of increase in prediction accuracy for each trait and environments are displayed in rows and columns, respectively.

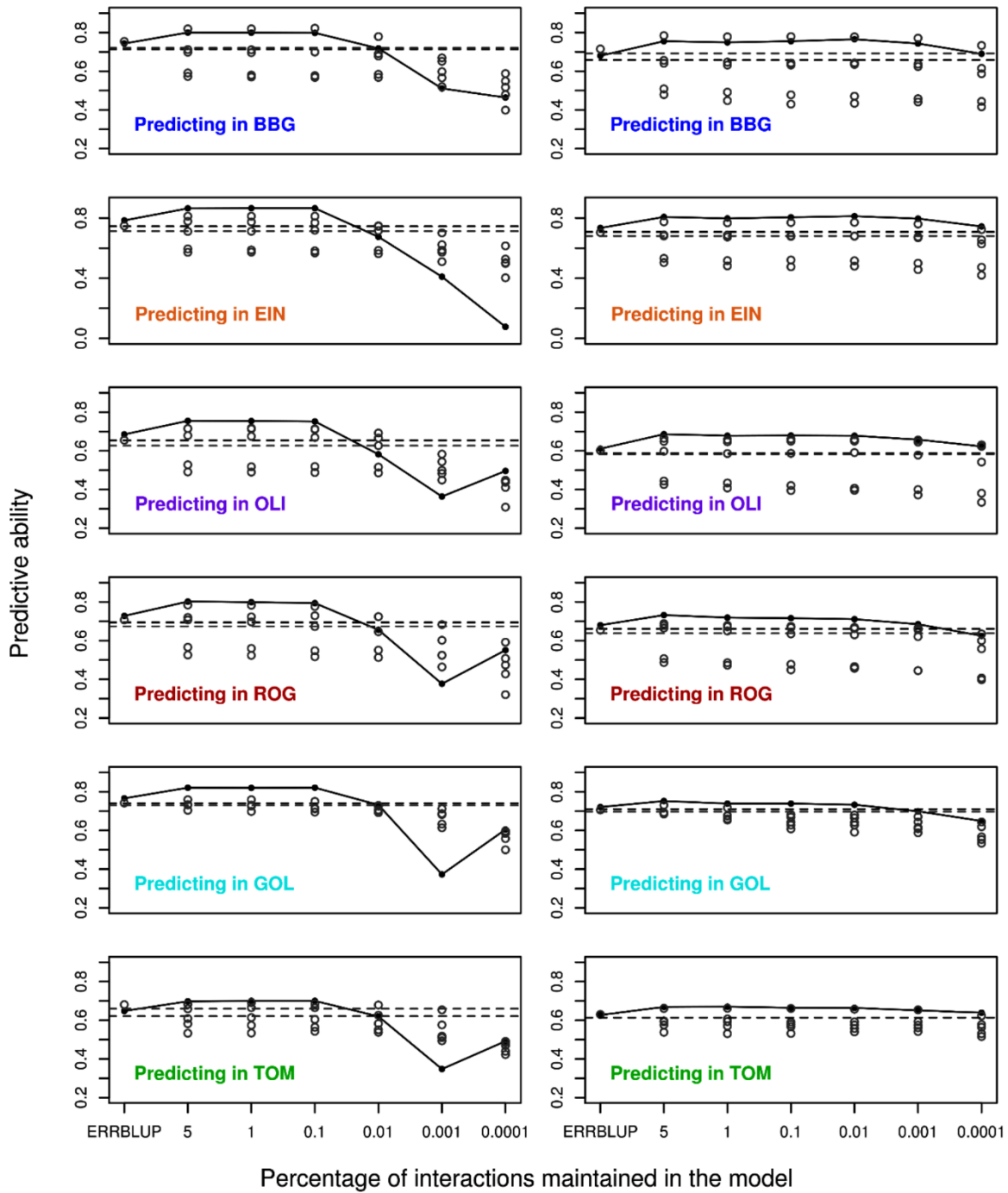
In addition to assessing the predictive ability of univariate sERRBLUP based on a single environment, Fig. 3.6 displays the comparison between the predictive ability obtained from univariate GBLUP and univariate ERRBLUP within environments, and univariate sERRBLUP across multiple environments jointly for trait PH\_V4 in KE and PE. It is demonstrated that univariate sERRBLUP has a higher predictive ability than univariate GBLUP when interactions are selected based on all the other five environments jointly. The preliminary analysis also reveals the robustness of the selection strategy based on the effects variance compared to selection strategy based on the absolute effects sizes in univariate sERRBLUP across multiple environments jointly for KE (Fig. S. 3.24), while for PE it does not show a significant difference for the interaction selection strategy (Fig. S. 3.25). Fig. 3.6 demonstrates that the predictive ability of univariate sERRBLUP across multiple environments jointly is as good as or better than using a single environment with few exceptions when the selection of effects is not too strict. With less than 0.1

percent of interactions used, predictive abilities deteriorate (especially so in KE) and selection from combined environments turns out to be worse than selection from single environments.

Fig. 3.7 illustrates the comparison between the predictive ability of bivariate GBLUP, ERRBLUP and sERRBLUP across multiple environments jointly and the maximum predictive ability of bivariate GBLUP and ERRBLUP and all the predictive abilities of sERRBLUP when a single environment is considered as an additional environment for the trait PH\_V4 in both KE and PE. The results indicate that bivariate sERRBLUP across multiple environments jointly increases the predictive ability compared to bivariate GBLUP and ERRBLUP across multiple environments jointly. In most cases, bivariate GBLUP, ERRBLUP and sERRBLUP across multiple environments jointly performs as good as or better than when using a single environment.



**Fig. 3.6:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (gray open circle), univariate sERRBLUP using a single environment for selecting the SNP interactions (gray open circles) and univariate sERRBLUP using all 5 environments jointly (filled black circles and solid line) for the SNP interaction selection based on estimated effects variances for trait PH\_V4 in KE (left side) and PE (right side).



**Fig. 3.7:** Predictive ability for bivariate GBLUP (black dashed horizontal line), bivariate ERRBLUP and bivariate sERRBLUP (filled black circles) for the SNP interaction selection based on estimated effects variances using all 5 environments jointly for trait PH-V4 in KE (left side) and PE (right side). In each panel, gray horizontal line and first gray open circles refer to maximum bivariate GBLUP and maximum bivariate ERRBLUP, and the gray open circles at the top 5, 1, 0.1, 0.01, 0.001, 0.0001 quantiles refer to bivariate sERRBLUP using a single environment as an additional environment.



### 3.6 Discussion

The accuracy of genomic prediction when incorporating epistasis interactions in the model compared to prediction models with only main effects has been widely discussed over the last years. In particular, it was found that accounting for epistasis can increase predictive ability (Carlborg and Haley, 2004; Hu *et al.*, 2011; Huang *et al.*, 2012; Wang *et al.*, 2012; Mackay, 2014; Jiang and Reif, 2015; Ober *et al.*, 2015; Rönnegård and Shen, 2016).

A major concern in utilizing epistasis models has been the high computational load (Mackay, 2014) which has been reduced for the full model including all interactions by utilizing marker based epistasis relationship matrices derived from Hadamard products of additive genomic relationship matrices (Jiang and Reif, 2015; Ober *et al.*, 2015; Martini *et al.*, 2016). The key advantage of this approach is that the number of random effects in the model is reduced from the number of SNP interactions to the number of genotypes. While the approaches of Jiang and Reif (2015) and Martini *et al.* (2016), only capture the interactions whose products differ from zero (i.e.  $\{22\}$  genotype combinations for 0, 2 coded markers), our approach captures all possible genotype combinations ( $\{00\}$ ,  $\{02\}$ ,  $\{20\}$ , and  $\{22\}$ ). Further, these epistasis relationship matrices and interaction effects were computed by bit-wise computations via the R-package miraculix (Schlather, 2020), which carries out matrix multiplications about 15 times faster than regular matrix multiplications on genotype data in EpiGP R-package (Vojgani *et al.*, 2021). In the analyzed datasets containing up to 30'212 SNPs (and thus 456'397'578 interactions), the computing time required to set up the sERRBLUP relationship matrix was about 810 minutes out of which around 330 minutes were required to estimate all pairwise SNP interaction effects and 480 minutes were required to set up the sERRBLUP relationship matrix for selected proportion of interactions by utilizing the R-package miraculix with 15 cores on a server cluster with Intel E5-2650 (2X12 core 2.2GHz) processors. Computing times for sERRBLUP scale approximately quadratic in the number of markers considered. The released EpiGP R-package (Vojgani *et al.*, 2021), which is available at <https://github.com/evoigani/EpiGP>, has been utilized for ERRBLUP and sERRBLUP genomic prediction of phenotypes.

Our proposed epistasis model eventually can generate a considerable prohibitive computational load if the number of SNPs grows to hundreds of thousands (Vojgani, *et al.*, 2019). The computing time for sERRBLUP exhibits quadratic growth with increasing number of SNPs. A potential strategy to overcome these limitations is to achieve a feature reduction by SNP pruning, as was implemented in our maize dataset (Purcell *et al.*, 2007; Chang *et al.*, 2015). Another option to obtain an even stronger variable reduction than pruning might be the use of haplotype blocks (Pook *et al.*, 2019). Although sERRBLUP model can be computationally challenging by increasing the number of SNPs, its predictive ability is constantly higher than the models such as RKHS, which reduces the computational time considerably (Table S. 3.10).

In this study, we showed that the predictive ability obtained by use of GBLUP and a full epistasis model with all pairwise SNP interactions included (ERRBLUP) was almost identical. In contrast, it was shown that the use of sERRBLUP increases predictive ability when only the most relevant

SNP interactions are taken into account, regardless of the choice of the training environment, which is likely a result of enriching for true causal variant combinations among the list of all variant combinations used to construct the genetic covariance matrix. In our study, the maximum predictive ability with sERRBLUP was obtained by incorporating the top 5, 1 or 0.1 percent of pairwise SNP interactions, while a too strict selection of SNP interactions such as the top 0.01, 0.001 and 0.0001 percent often reduced the predictive ability. A similar loss in predictive ability with a too strict selection of interactions to be included in the model was also observed by Ober et al. (2015). The difference in interaction selection can be explained by the fact that the absolute number of interaction effects in the model is more important than the percentage of interaction effects. To illustrate this, the absolute numbers of interactions maintained in the model for the top 0.001 and 0.0001 percent of interactions in KE are respectively 3'235 and 323, which is less than the number of additive effects in KE (25'437) where the obtained sERRBLUP predictive ability is lower than GBLUP predictive ability. In addition, the possible differences in linkage can also lead to different redundancy patterns of interactions. Here we also saw the only major systematic difference between the two selection criteria: when SNP interactions were selected based on the magnitude of their estimated (absolute) effects, the loss in predictive ability when selecting too few interactions was much more severe than when SNP interactions were selected based on the variance associated with them. This phenomenon has been more prevalent in KE than in PE (Fig. S. 3.10 - Fig. S. 3.11), and is valid in both scenarios, using information either from a single environment or from the average of all other environments (Fig. S. 3.24 - Fig. S. 3.25). A potential reason for this is that the few interactions that remain in the model are highly linked and thus no proper representation of the overall population structure is possible anymore. This effect was even more pronounced when selecting based on effect sizes. Thus, we recommend the use of effect variances as a selection criterion in sERRBLUP applications since this should be conceptually more robust.

The bivariate models exhibited a considerably higher predictive ability than univariate models. In consequence, the bivariate GBLUP performed slightly better than the best univariate sERRBLUP in most cases (Fig. S. 3.26). Across all studied traits, the increase in prediction accuracy from GBLUP to sERRBLUP displays a similar pattern in both univariate and bivariate models. It has to be noted that this increase in predictive ability is exclusively caused by the modelling of epistasis in a bivariate statistical setting, while it is caused by both modelling of epistasis and borrowing information across environments through variable selection in the univariate statistical setting.

In general, it is expected that the predictive ability for phenotypes should be higher with higher heritability. In this study, the correlation between the heritability of all traits, which have been calculated on an entry-mean basis within each landraces (Hallauer *et al.*, 2010) over all environments, was 0.296 with univariate GBLUP within environments and 0.543 with maximum univariate sERRBLUP across environments (Fig. S. 3.27). Corresponding correlations were higher in the bivariate statistical setting of the respective models, with an increase in the respective correlation from maximum bivariate GBLUP (0.537) to maximum bivariate sERRBLUP (0.647) (Fig. S. 3.28).

When comparing sERRBLUP to a traditional GxE model (Kang and Gorman, 1989), the modelling approach is quite different. In sERRBLUP, the selection of marker-by-marker interactions is done based on a second environment. However, for the final estimation of the actual effect size, the data from the same environment is used. Thus, effect sizes can substantially differ between environments. In contrast to this, traditional GxE model will assign effects to specific marker-by-environment combinations. As included interactions between different environments in sERRBLUP are different, it is not possible to put concrete GxE effects on specific markers or marker-by-environment interactions, which would be the essence of traditional GxE models. As prediction performances are increasing quite substantially by the use of sERRBLUP, this still can be seen as an indication that effect regions are similar between environment (although effect sizes might differ).

Our results indicate that a higher number of phenotyped lines (in particular overlapping between environments) and including information from a more similar second environments were beneficial for prediction. E.g., when the two Spanish locations GOL or TOM were used as the second environment to predict a German environment, prediction accuracies were lower as these environments have substantially different climate and for some traits lower overlap between phenotyped lines. On the other hand, the best prediction results for GOL were obtained when using TOM as second environment and vice versa.

In both univariate and bivariate models, it was shown that the obtained predictive ability across multiple environments jointly was mostly equivalent or higher than the maximum predictive ability obtained based on a single environment. Thus, using an average across all other environments should be a robust alternative which in most cases will yield a result that is as good as or even better than the best single environment.

Overall, our results demonstrate that bivariate models can outperform univariate models and epistatic interactions can substantially increase the predictive ability. In the context of univariate models, it was shown that selecting a suitable subset of interactions based on other environments where phenotypic data of the full set of lines are available can substantially increase the predictive ability. As the ideal share  $\pi$  of interactions to be included in sERRBLUP is not known in practice, one could consider to run a testing scheme with an additional validation set for the identification of a suitable  $\pi$ . As results were quite robust as long as a reasonable fraction (between 5 and 0.1 percent) of interactions were included in the model and this introduces further computational load, this should however usually not be required.

The presented approach can substantially improve the phenotype prediction accuracy in another environment by ‘borrowing’ information on effect regions from another variable. In our case, this other variable were phenotypes of the same trait grown in different environments. However, one could also imagine using data from another growing season or even from a highly correlated second trait. This can be useful in sparse testing designs, e.g. where not all lines are grown in all environments. The suggested approach can be used to ‘impute’ missing phenotypes with a much increased accuracy compared to conventional approaches.



## 3.7 Supplementary material

### 3.7.1 Supplementary Tables

**Table S. 3.1:** Phenotypic traits descriptions, locations, number of lines, mean, minimum, maximum and standard deviation of phenotypic traits in each location for KE (blue numbers) and PE (red numbers).

Trait	Definition	Location	Number of Phenotyped lines	Mean	Minimum	Maximum	Standard deviation
<b>EV_V3</b>	Early vigour at V3 stage scored on scale from 1 (very poor early vigour) to 9 (very high early vigour)	<b>BBG</b>	459\393	4.10\4.68	0.78\1.00	7.28\7.55	1.28\1.18
		<b>EIN</b>	462\393	4.11\4.70	0.86\1.03	9.00\9.03	1.31\1.18
		<b>OLI</b>	440\389	5.31\6.17	1.22\3.30	8.05\8.74	1.15\0.86
		<b>ROG</b>	460\390	5.35\5.84	1.71\2.90	7.90\7.92	0.95\0.75
		<b>GOL</b>	210\204	6.31\6.67	4.07\5.49	8.49\7.98	0.69\0.51
		<b>TOM</b>	210\204	5.51\6.15	1.93\3.84	7.34\8.45	0.99\0.67
<b>EV_V4</b>	Early vigour at V4 stage scored on scale from 1 (very poor early vigour) to 9 (very high early vigour)	<b>BBG</b>	459\393	3.85\4.65	0.67\0.93	8.29\8.49	1.48\1.48
		<b>EIN</b>	462\393	4.24\4.82	0.94\1.52	7.07\7.46	1.11\0.98
		<b>OLI</b>	439\389	5.27\6.07	0.80\2.99	7.52\8.36	1.08\0.75
		<b>ROG</b>	459\390	5.44\5.85	2.65\2.88	7.86\7.94	0.92\0.78
		<b>GOL</b>	210\204	5.71\5.98	3.37\3.91	7.89\7.89	0.81\0.83
		<b>TOM</b>	210\204	5.26\5.75	2.59\3.92	6.89\7.35	0.83\0.61
<b>EV_V6</b>	Early vigour at V6 stage scored on scale from 1 (very poor early vigour) to 9 (very high early vigour)	<b>BBG</b>	459\393	3.92\4.64	0.74\0.84	8.75\8.22	1.39\1.41
		<b>EIN</b>	462\393	5.03\5.54	0.97\1.51	8.05\8.39	1.24\1.06
		<b>OLI</b>	437\388	5.30\6.07	0.54\3.56	7.17\8.09	0.96\0.74
		<b>ROG</b>	461\390	5.55\5.91	1.02\2.52	8.07\7.76	0.95\0.77
		<b>GOL</b>	210\204	6.24\6.24	3.90\3.81	8.45\7.94	0.85\0.85
		<b>TOM</b>	210\204	5.58\5.86	2.96\3.90	7.66\7.91	0.92\0.68

Trait	Definition	Location	Number of Phenotyped lines	Mean	Minimum	Maximum	Standard deviation
<b>PH_V4</b>	Mean plant height of three plants of the plot at V4 stage in cm	<b>BBG</b>	461\393	35.86\41.48	9.27\16.38	60.40\62.85	8.43\7.93
		<b>EIN</b>	462\393	34.49\38.73	6.90\20.43	53.14\57.94	7.24\6.17
		<b>OLI</b>	440\388	18.43\22.55	7.35\11.89	31.11\35.75	3.93\3.87
		<b>ROG</b>	460\390	25.50\28.10	9.23\13.63	42.29\41.54	4.60\4.53
		<b>GOL</b>	210\204	62.88\68.98	34.30\38.39	88.24\95.30	9.79\10.96
		<b>TOM</b>	210\204	41.60\47.45	11.98\25.37	63.89\72.12	8.71\8.27
<b>PH_V6</b>	Mean plant height of three plants of the plot at V6 stage in cm	<b>BBG</b>	461\393	61.75\69.08	19.41\30.36	93.84\100.39	11.80\11.12
		<b>EIN</b>	462\393	62.40\69.36	21.41\36.53	95.54\98.80	11.89\9.62
		<b>OLI</b>	440\390	36.74\45.35	8.34\14.78	58.40\72.48	8.69\8.53
		<b>ROG</b>	459\390	61.46\68.91	32.17\30.35	89.74\94.77	9.34\9.52
		<b>GOL</b>	210\204	94.21\98.30	37.28\54.75	127.54\130.51	15.05\15.29
		<b>TOM</b>	210\204	83.86\92.35	48.46\57.81	119.07\124.98	14.41\12.79
<b>PH_final</b>	Final plant height after flowering in cm	<b>BBG</b>	461\393	142.65\120.60	95.85\59.78	210.08\179.03	20.53\19.11
		<b>EIN</b>	462\393	159.18\141.35	100.84\69.01	228.96\211.14	21.57\21.10
		<b>OLI</b>	432\387	118.17\112.46	58.74\58.55	175.81\173.15	21.95\20.48
		<b>ROG</b>	461\390	137.04\122.25	74.25\63.56	211.14\201.92	22.32\20.56
		<b>GOL</b>	209\204	115.68\102.69	49.27\30.21	167.58\149.14	21.73\23.59
		<b>TOM</b>	210\204	157.99\144.61	81.92\79.28	245.00\195.36	24.82\18.95
<b>FF</b>	Days after sowing until female flowering (days until 50% of the plot showed silks)	<b>BBG</b>	461\393	82.10\82.08	69.45\69.78	95.74\92.04	4.31\4.18
		<b>EIN</b>	462\393	82.55\81.78	70.36\68.86	102.02\101.50	5.23\5.17
		<b>OLI</b>	346\347	83.32\82.41	72.60\69.46	92.13\91.54	3.76\3.68
		<b>ROG</b>	458\389	73.06\71.91	62.45\59.10	91.22\88.03	4.82\4.47
		<b>TOM</b>	209\203	76.88\74.16	63.93\62.13	93.28\92.17	5.58\4.64
<b>RL</b>	Root lodging score from 1 to 9 (1 belonged to no lodging and 9 belonged to severe lodging)	<b>BBG</b>	461\392	5.02\3.03	0.59\0.03	9.58\9.22	2.78\2.39
		<b>EIN</b>	462\393	3.48\2.23	0.63\0.76	9.21\8.08	2.29\1.54
		<b>OLI</b>	439\388	2.59\1.80	0.59\0.52	9.15\7.65	1.64\1.19
		<b>ROG</b>	460\390	2.39\1.50	0.96\0.95	9.01\8.50	2.21\1.13

**Table S. 3.2:** The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait PH-V4 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.

<b>Predicted Environment</b>	<b>Additional Environment</b>	<b>Top 5</b>	<b>Top 1</b>	<b>Top 0.1</b>	<b>Top 0.01</b>	<b>Top 0.001</b>
<b>BBG</b>	<b>EIN</b>	100%	100%	100%	100%	100%*
<b>BBG</b>	<b>GOL</b>	0%*	8%*	12%	100%	100%
<b>EIN</b>	<b>BBG</b>	100%	100%	100%	100%	100%*
<b>EIN</b>	<b>GOL</b>	100%*	100%*	28%*	100%	100%
<b>OLI</b>	<b>GOL</b>	16%*	16%*	100%*	100%	100%
<b>ROG</b>	<b>BBG</b>	100%	100%	100%	100%	100%*
<b>ROG</b>	<b>GOL</b>	100%*	100%*	100%*	100%	100%
<b>TOM</b>	<b>BBG</b>	100%*	100%*	100%*	0%*	100%*
<b>TOM</b>	<b>EIN</b>	100%	100%*	100%*	96%	100%*
<b>TOM</b>	<b>GOL</b>	100%*	100%*	100%*	64%	100%*

**Table S. 3.3:** The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait EV\_V3 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The stars represent the non-convergence of pre estimated variance components based on the full set.

Predicted Environment	Additional Environment	Top 5	Top 1	Top 0.1	Top 0.01	Top 0.001
BBG	EIN	100%	100%	100%	100%	96%*
BBG	GOL	96%	100%	100%	100%	100%
BBG	TOM	100%	100%	100%	100%	88%
EIN	ROG	100%	100%	92%	96%	92%
EIN	BBG	100%	100%	100%	100%	88%*
EIN	TOM	100%	100%	96%	96%	96%
OLI	ROG	100%	100%	100%	92%	96%
OLI	EIN	100%	100%	100%	92%	0%* / 96%
OLI	BBG	100%	100%	96%	96% / 88%	100%* / 76%
OLI	GOL	96%	96%	88%	96%	92%
OLI	TOM	92%	96%	72%	72%*	60%*
ROG	EIN	100%	100%	96%	52%*	0%* / 76%
ROG	BBG	100%	100%	100%	96% / 92%	0%*
ROG	GOL	100%	100%	96%	92%	96%
ROG	TOM	100%	100%	100%	84%	100%
GOL	ROG	100%	100%	100%	92%	100%
GOL	EIN	100%	100%	100%	100%	0%*
GOL	BBG	100%	100%	96%	96%	80%* / 96%
GOL	OLI	96%	92%	96%	92%	88%
GOL	TOM	100%	100%	100%	88% / 80%	80% / 96%
TOM	EIN	96%	92%	64%	44%*	0%*
TOM	BBG	100%	100%	100%	96%	80%*
TOM	OLI	96%	100%	100%	100%	92%
TOM	GOL	0%*	96%	100%	80%	96% / 96%



**Table S. 3.4:** The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait EV\_V4 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.

<b>Predicted Environment</b>	<b>Additional Environment</b>	<b>Top 5</b>	<b>Top 1</b>	<b>Top 0.1</b>	<b>Top 0.01</b>	<b>Top 0.001</b>
<b>BBG</b>	<b>EIN</b>	100%	100%	100%	100%	12%*
<b>BBG</b>	<b>GOL</b>	100%	100%	96%	68%	56%
<b>EIN</b>	<b>OLI</b>	100%	96%	100%	100%	100%
<b>OLI</b>	<b>EIN</b>	88%	88%	96% / 16%*	96% / 4%*	8%* / 8%*
<b>OLI</b>	<b>BBG</b>	100%	100%	96%	100%	100%
<b>OLI</b>	<b>GOL</b>	100%	100%	100%	96%	100%
<b>OLI</b>	<b>TOM</b>	100%	100%	100%	96%	96%
<b>ROG</b>	<b>EIN</b>	100%	100%	100%	100%	0%*
<b>GOL</b>	<b>EIN</b>	100%	100%	100%	100%	0%*
<b>GOL</b>	<b>TOM</b>	100%	96%	100%	100%	100%
<b>TOM</b>	<b>ROG</b>	100%	96%	100%	92%	100%
<b>TOM</b>	<b>EIN</b>	100%	96%	88%	100%	100%
<b>TOM</b>	<b>BBG</b>	100%	96%	25%*	40%*	80%
<b>TOM</b>	<b>GOL</b>	80%	80%	96% / 48%*	96%	80%

**Table S. 3.5:** The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait EV\_V6 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.

Predicted Environment	Additional Environment	Top 5	Top 1	Top 0.1	Top 0.01	Top 0.001
BBG	ROG	100%	100%	100%	100%	80%*
BBG	EIN	100%	100%	96%	92%	0%*
BBG	GOL	0%* / 25%*	0%* / 48%*	24%* / 24%*	0%*	100%
EIN	ROG	100%	100%	96%	80%* / 96%	52%* / 96%
EIN	GOL	4%* / 0%*	4%* / 88%*	48%* / 0%*	32%	100%
OLI	EIN	100%	100%	88%	88%	4%* / 84%
OLI	ROG	100%	100%	100%	100%	0%*
OLI	GOL	25%* / 0%*	72% / 96%	24%* / 4%*	12%*	100%
ROG	EIN	100%	100%	92%	80%	40%* / 88%
ROG	BBG	96%	96%	100%	96% / 84%*	96% / 52%*
ROG	OLI	100%	100%	96%	100%	96%
ROG	GOL	25% / 0%*	0%* / 96%	0%* / 36%	0%	100%
GOL	ROG	100%	100%	100%	100%	96%*
GOL	EIN	100%	100%	100%	100%	80%*
TOM	ROG	100%	96%	24%*	60%* / 96%	0%* / 72%
TOM	EIN	100%	100%	72%	72%	0%*
TOM	BBG	100%	76%	76% / 96%	84%	84%
TOM	OLI	100%	100%	100%	100%	96%
TOM	GOL	0%* / 0%	0%* / 84%	0%* / 0%	16%* / 0%*	25%* / 60%*

**Table S. 3.6:** The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait PH\_V6 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The stars represent the non-convergence of pre estimated variance components based on the full set.

Predicted Environment	Additional Environment	Top 5	Top 1	Top 0.1	Top 0.01	Top 0.001
BBG	ROG	100%	100%	100%	96%	80%*
BBG	EIN	92%	100%	100%	100%	100%
BBG	OLI	100%	100%	100%	100%	36%
BBG	GOL	100%	100%	96%	96%	92%*
EIN	ROG	100%	100%	100%	100%	0%*
EIN	BBG	100%	100%	80%	100%	32%*
EIN	OLI	100%	100%	100%	100%	96%
EIN	GOL	100%	100%	100%	100%	100%*
OLI	ROG	100%	100%	96%	68%	16%*
OLI	EIN	100%	100%	100%	100%	92%
OLI	BBG	100%	100%	100%	100%	56%*
OLI	GOL	100%	100%	100%	96%	0%*
ROG	EIN	100%	100%	100%	100%	92%
ROG	BBG	100%	100%	100%	96%	84%*
ROG	OLI	100%	100%	100%	100%	0%*
ROG	GOL	100%	100%	100%	100%	8%*
GOL	ROG	100%	100%	100%	100%	96%*
GOL	BBG	100%	100%	100%	100%	0%* / 96%
GOL	OLI	100%	100%	100%	100%	0%*
TOM	ROG	100%	100%	92%	88%	40%*
TOM	EIN	100%	100%	100%	100%	100%
TOM	BBG	100%	100%	32%	92%	0%*
TOM	OLI	100%	100%	100%	96%	4%*
TOM	GOL	52%*	68%* / 88%	72%* / 92%	92%	16%* / 96%

**Table S. 3.7:** The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait PH\_final for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.

Predicted Environment	Additional Environment	Top 5	Top 1	Top 0.1	Top 0.01	Top 0.001
BBG	ROG	52%* / 0%*	0%* / 0%*	0%* / 0%*	0%*	100%
BBG	EIN	100%	100%	100%	100%	96%
BBG	TOM	36%*	4%*	0%*	92%	100%
EIN	ROG	0%* / 25%	0%* / 0%*	0%* / 0%*	96% / 0%*	92%
EIN	BBG	8%*	4%*	0%*	96%	96%
EIN	OLI	100%	100%	100%	100%	92%
EIN	TOM	64%	4%*	0%*	100%	100%
OLI	ROG	0%* / 4%*	0%* / 0%*	0%* / 0%*	0%*	96%
OLI	EIN	100%	96%	100%	100%	100%
OLI	BBG	4%*	0%*	0%*	100%	96%
OLI	GOL	100%	100%	100%	100%	100%
OLI	TOM	28%*	12%*	0%*	96% / 84%*	92%
ROG	EIN	100%	100%	100%	100%	96%
ROG	BBG	40%*	0%*	0%*	100%	100%
ROG	TOM	24%*	8%*	4%*	88%	100%
GOL	ROG	8%* / 0%*	4%* / 4%*	8%* / 4%*	0%*	100%
GOL	BBG	8%*	4%*	4%*	100%	100%
GOL	TOM	0%*	0%*	4%*	88%	100%
TOM	ROG	0%* / 0%*	0%* / 0%*	0%* / 0%*	36%* / 0%*	48%* / 96%
TOM	EIN	100%	100%	84%	76%	40%* / 96%
TOM	BBG	8%*	0%*	0%*	88%	52%*
TOM	OLI	100%	100%	92%	92%	96%
TOM	GOL	100%	100%	96%	100%	96%

**Table S. 3.8:** The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait FF for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.

Predicted Environment	Additional Environment	Top 5	Top 1	Top 0.1	Top 0.01	Top 0.001
BBG	EIN	4%*	4%*	0%*	92% / 32%*	80%
BBG	ROG	0%*	0%*	0%* / 0%*	80% / 0%*	68%
EIN	BBG	96% / 36%*	92% / 96%	92% / 0%*	88%* / 0%*	88% / 96%
EIN	ROG	0%*	0%*	0%* / 0%*	96% / 0%*	96% / 84%
OLI	ROG	0%*	92% / 0%*	16%* / 0%*	10 / 0%*	92% / 76%
OLI	BBG	100%	84%	0%*	4%*	96% / 32%*
OLI	EIN	64%*	20%	0%*	52%*	44%*
ROG	EIN	0%*	0%*	0%*	80%	52%*
ROG	BBG	40%	100%	0%*	0%*	96%
TOM	EIN	28%*	20%*	24%*	40%*	72% / 12%*
TOM	BBG	72%	100%	4%*	8%*	96%
TOM	OLI	100%	100%	96%	96%	96%
TOM	ROG	12%*	96% / 0%	0%* / 4%*	20%* / 4%*	4%* / 40%*

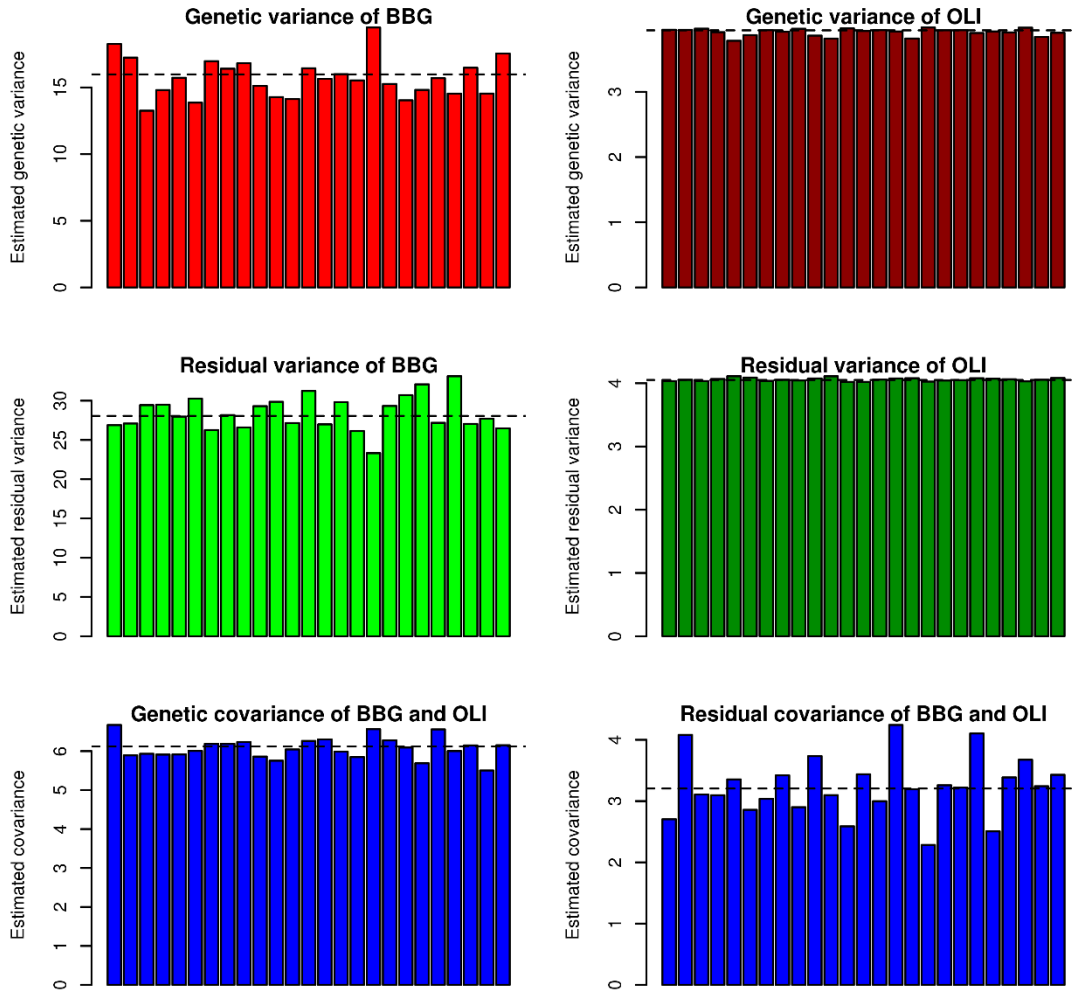
**Table S. 3.9:** The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait RL for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.

Predicted Environment	Additional Environment	Top 5	Top 1	Top 0.1	Top 0.01	Top 0.001
<b>BBG</b>	<b>EIN</b>	100%	100%	100%	100%	96%
<b>EIN</b>	<b>ROG</b>	100%	100%	96%	96%	96%
<b>EIN</b>	<b>BBG</b>	100%	72%	60%*	20%*	96% / 16%*
<b>EIN</b>	<b>OLI</b>	92%	92%	92%	96%	68%
<b>ROG</b>	<b>EIN</b>	100%	100%	96%	88%	68%
<b>ROG</b>	<b>BBG</b>	100%	100%	100%	96%	96%
<b>ROG</b>	<b>OLI</b>	100%	100%	96%	100%	100%

**Table S. 3.10:** The predictive ability of RKHS, univariate GBLUP within environments, univariate ERRBLUP within environments and maximum univariate sERRBLUP across environments (blue numbers) and the maximum predictive ability of bivariate GBLUP, bivariate ERRBLUP and bivariate sERRBLUP (red numbers) in KE and PE for the trait PH\_V4.

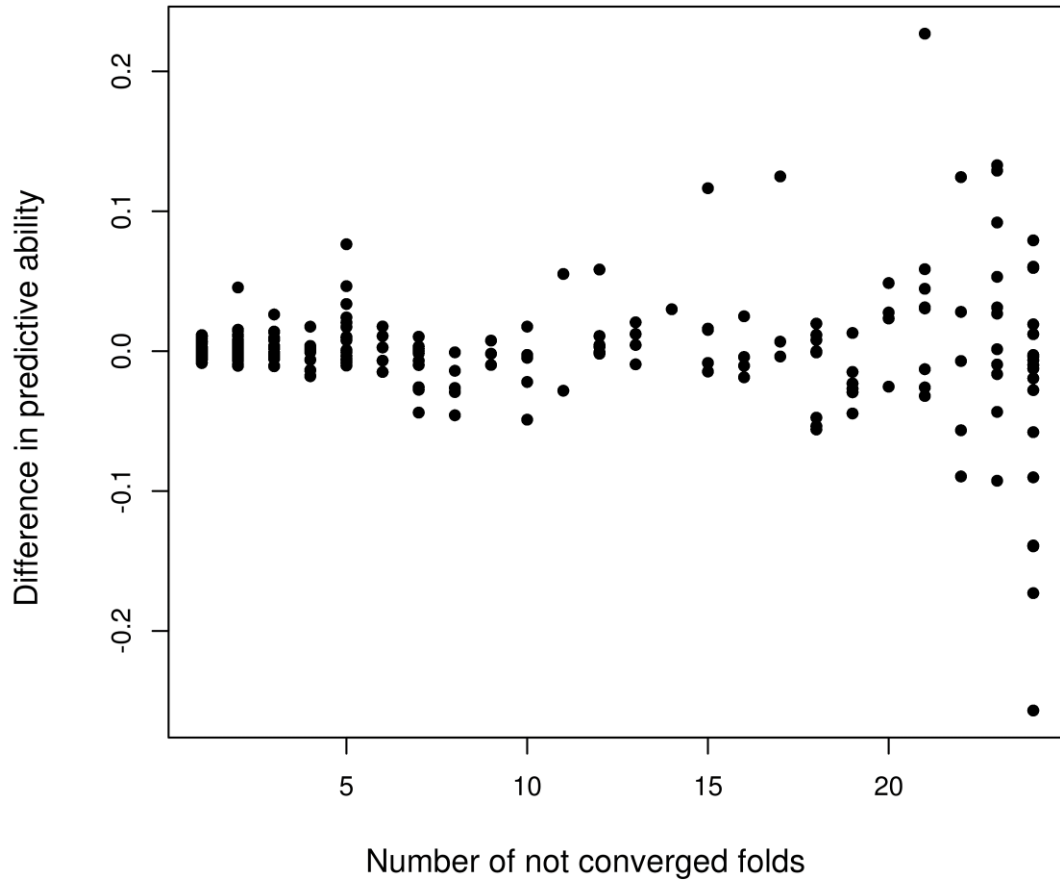
Landrace	Locations	RKHS	GBLUP	ERRBLUP	sERRBLUP
<b>KE</b>	<b>BBG</b>	0.479	0.470/0.721	0.468/0.754	0.707/0.822
	<b>EIN</b>	0.479	0.455/0.713	0.458/0.748	0.679/0.812
	<b>OLI</b>	0.435	0.439/0.627	0.438/0.655	0.623/0.717
	<b>ROG</b>	0.460	0.469/0.674	0.467/0.707	0.665/0.784
	<b>GOL</b>	0.543	0.519/0.730	0.521/0.742	0.694/0.759
	<b>TOM</b>	0.353	0.346/0.660	0.350/0.681	0.617/0.691
	<b>PE</b>	<b>BBG</b>	0.423	0.424/0.692	0.426/0.715
<b>EIN</b>		0.420	0.412/0.680	0.413/0.704	0.652/0.775
<b>OLI</b>		0.334	0.324/0.582	0.326/0.602	0.558/0.665
<b>ROG</b>		0.422	0.449/0.638	0.447/0.654	0.584/0.688
<b>GOL</b>		0.590	0.604/0.697	0.606/0.706	0.664/0.731
<b>TOM</b>		0.342	0.370/0.614	0.370/0.630	0.528/0.662

### 3.7.2 Supplementary Figures



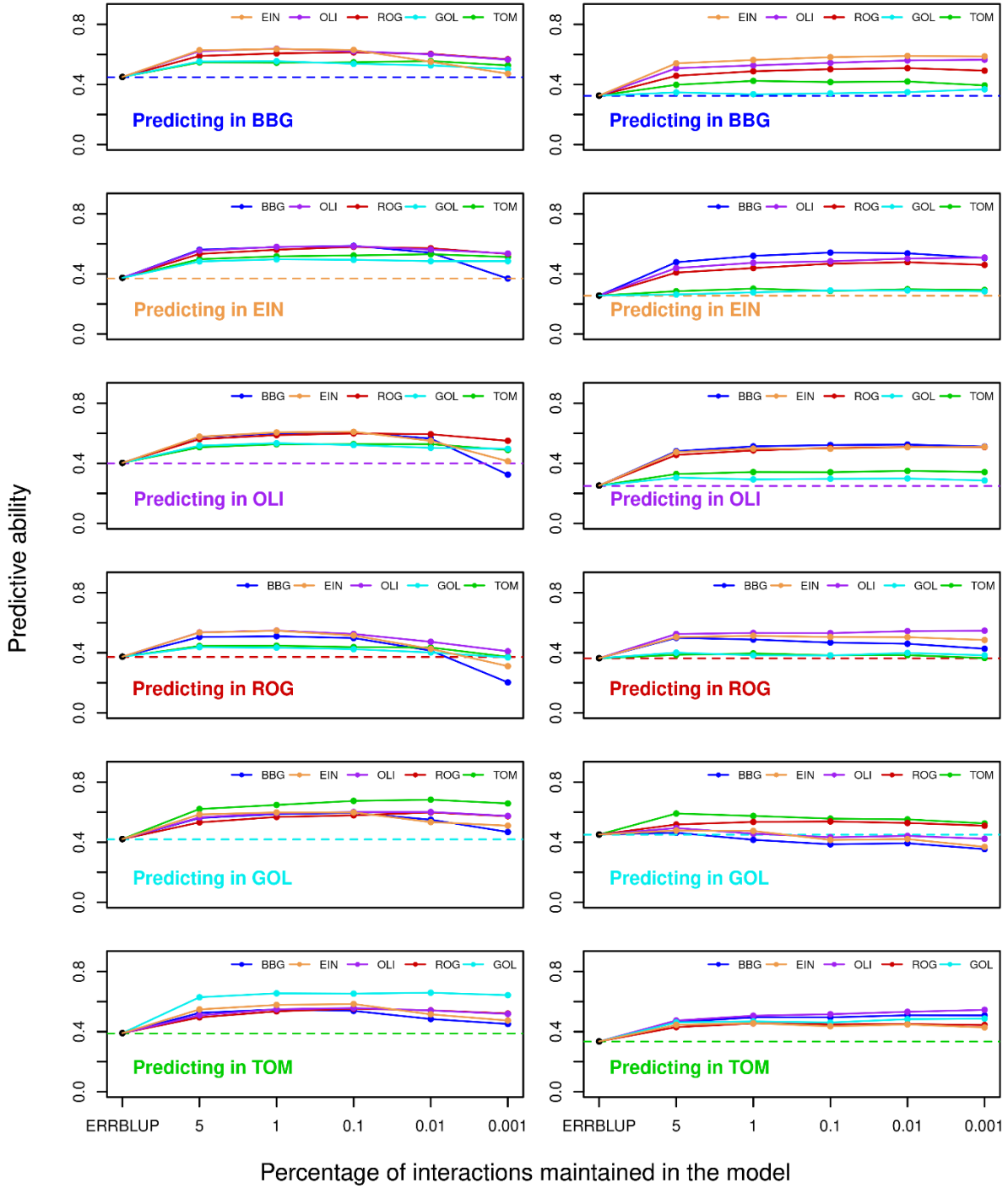
5-fold cross validation with 5 replicates

**Fig. S. 3.1:** Comparison of pre estimated genetic and residual variances and covariances of converged bivariate sERRBLUP model (top 5%) based on the full dataset (dashed horizontal lines) and estimated genetic and residual variances and covariances of converged bivariate sERRBLUP (top 5%) based on training set in each run of 5-fold cross validation with 5 replicates (colored bars) for predicting BBG when the additional environment is OLI in KE for trait PH\_V4.

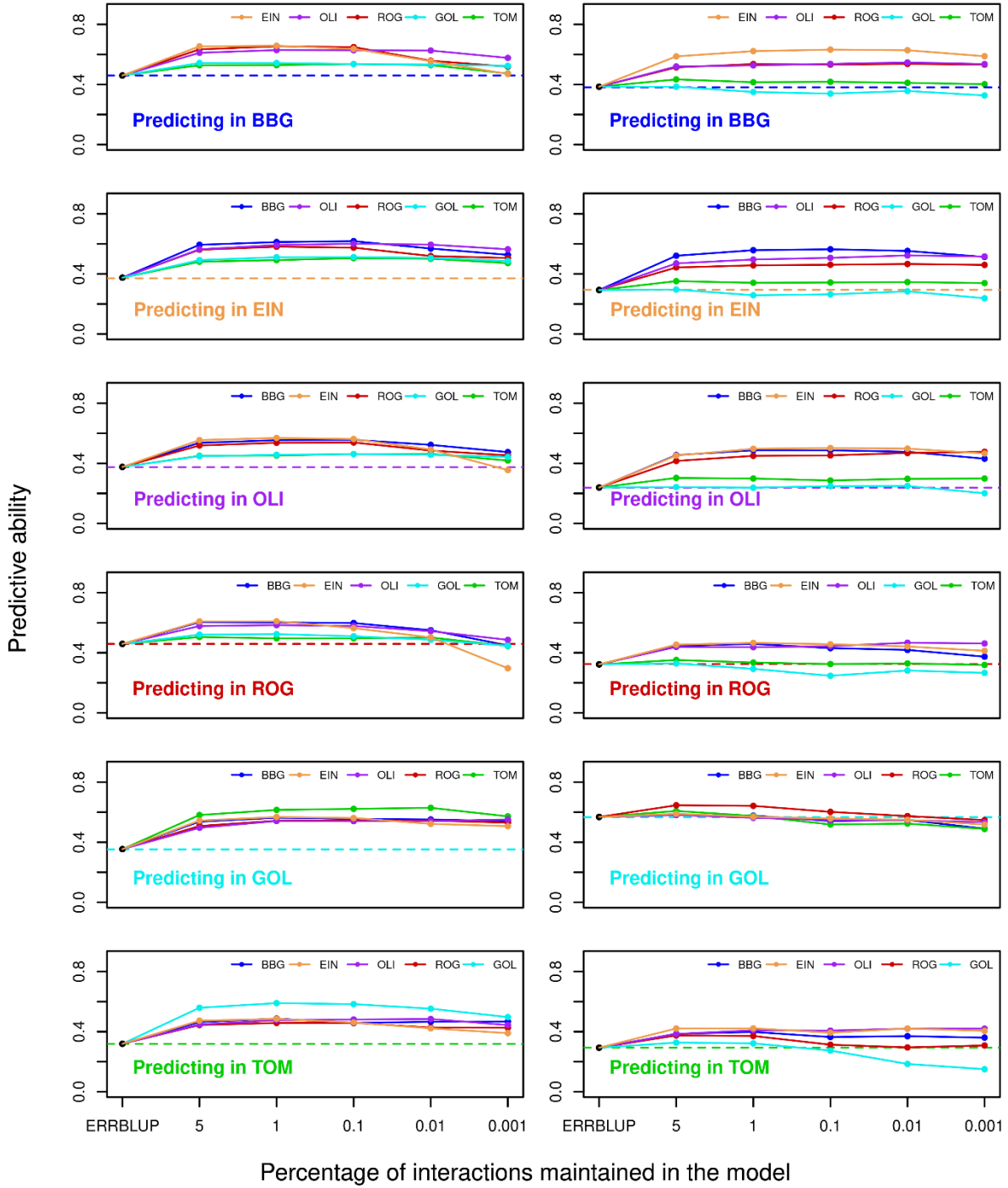


**Fig. S. 3.2:** The difference between the mean predictive ability of only the converged folds and the mean predictive ability of all folds in 5-fold cross validation with 5 replicates vs. the number of the folds (1 to 24) which did not converge across all traits in all combinations for both KE and PE.

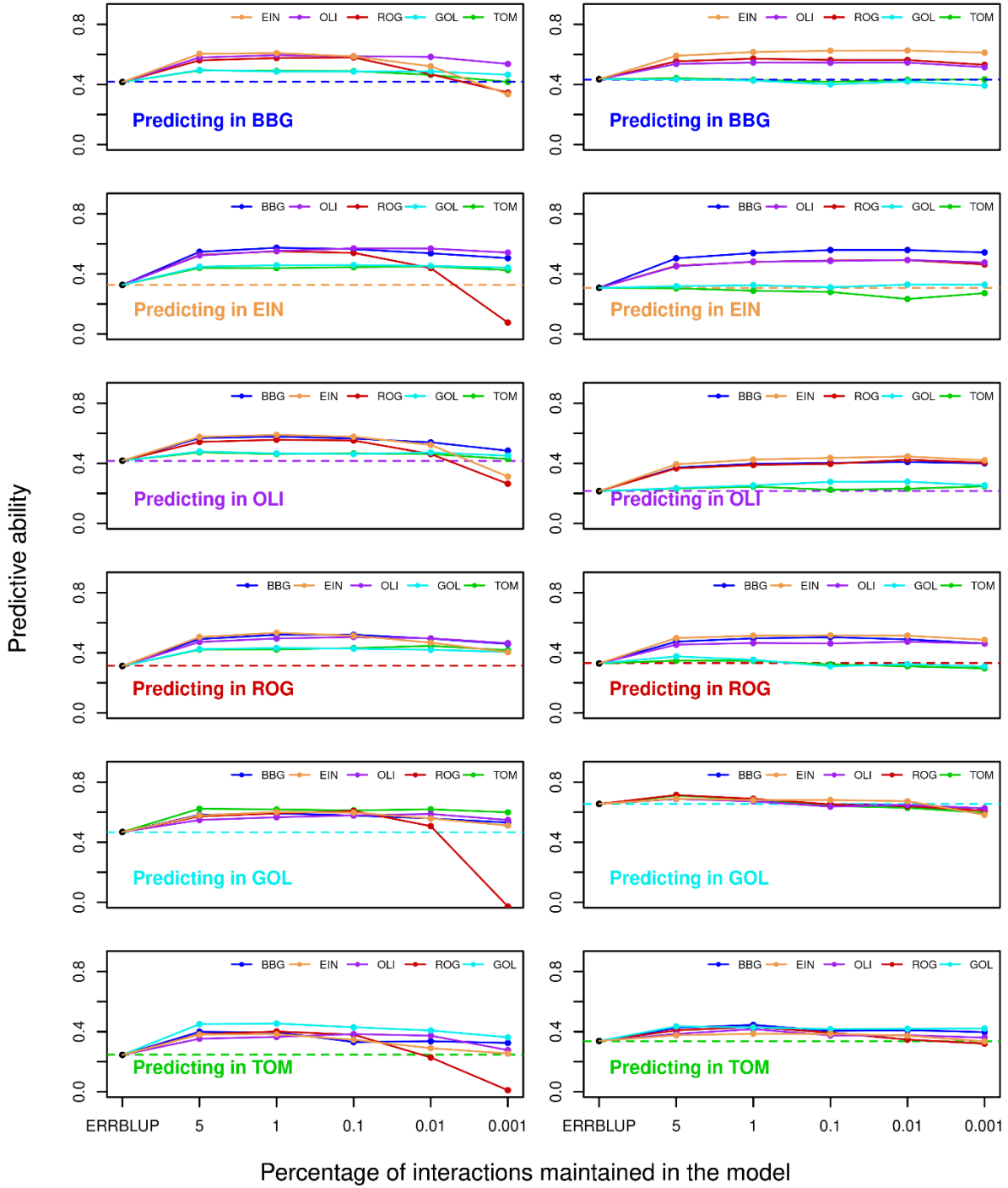




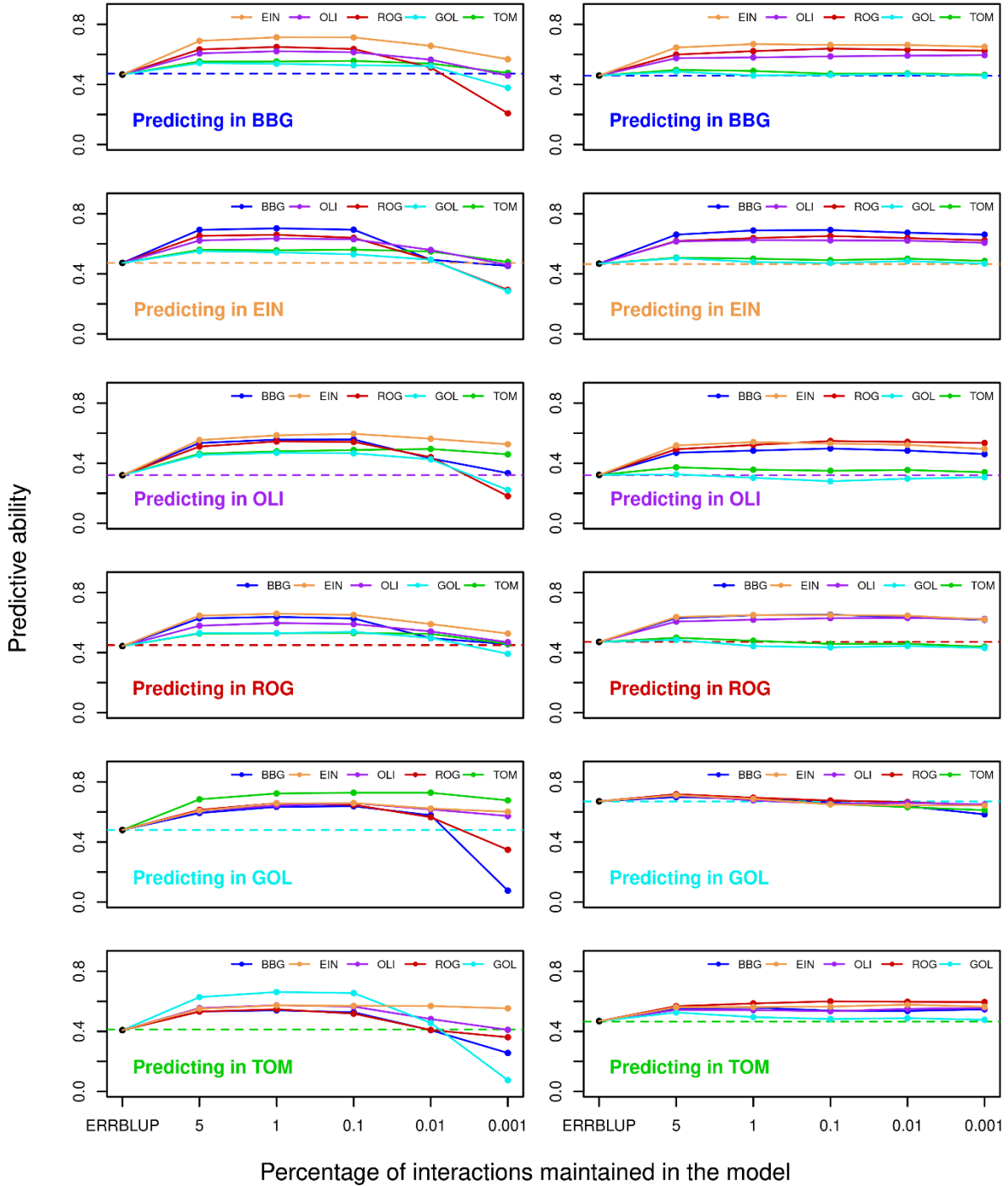
**Fig. S. 3.3:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait EV\_V3 in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection.



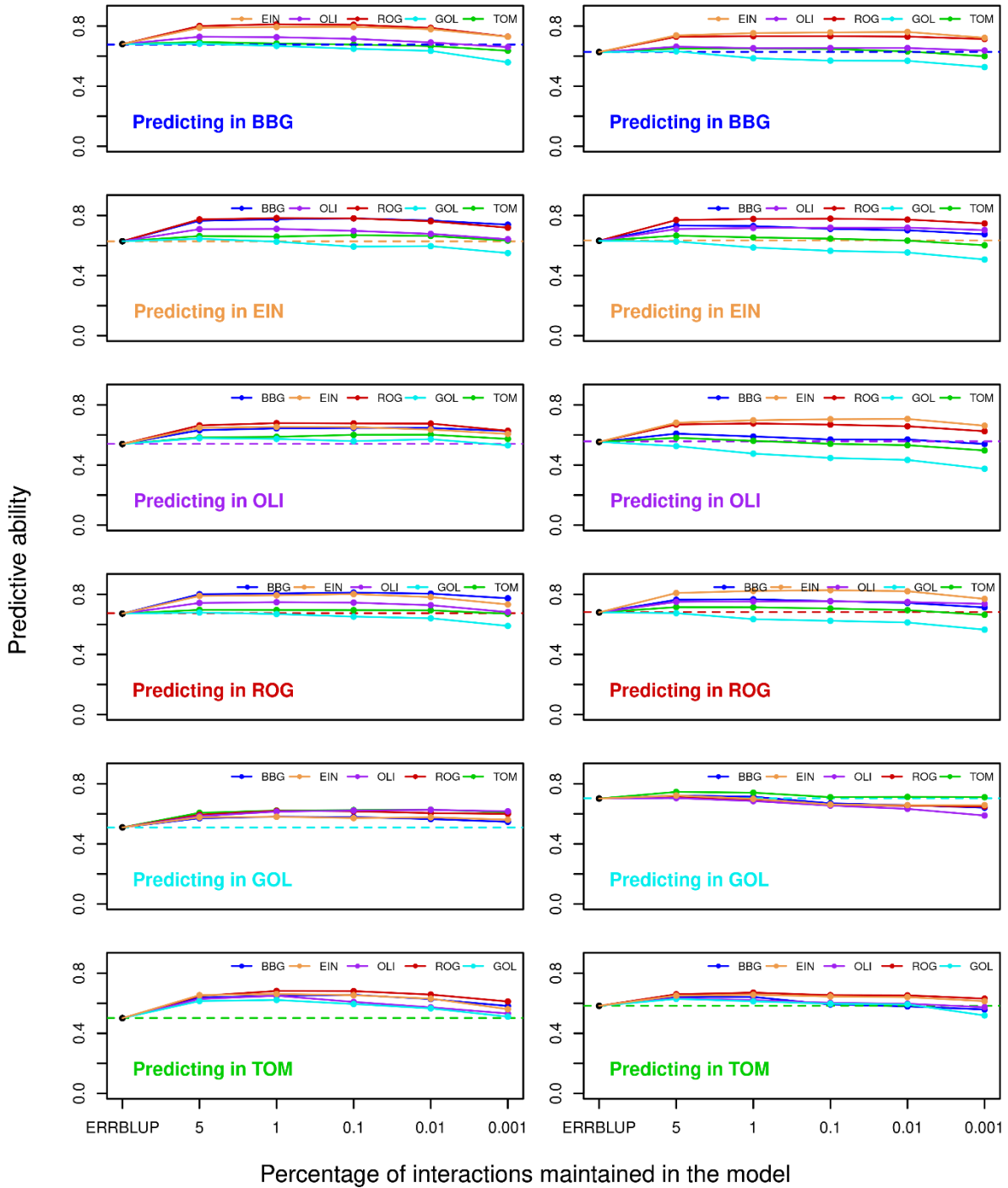
**Fig. S. 3.4:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait EV\_V4 in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection.



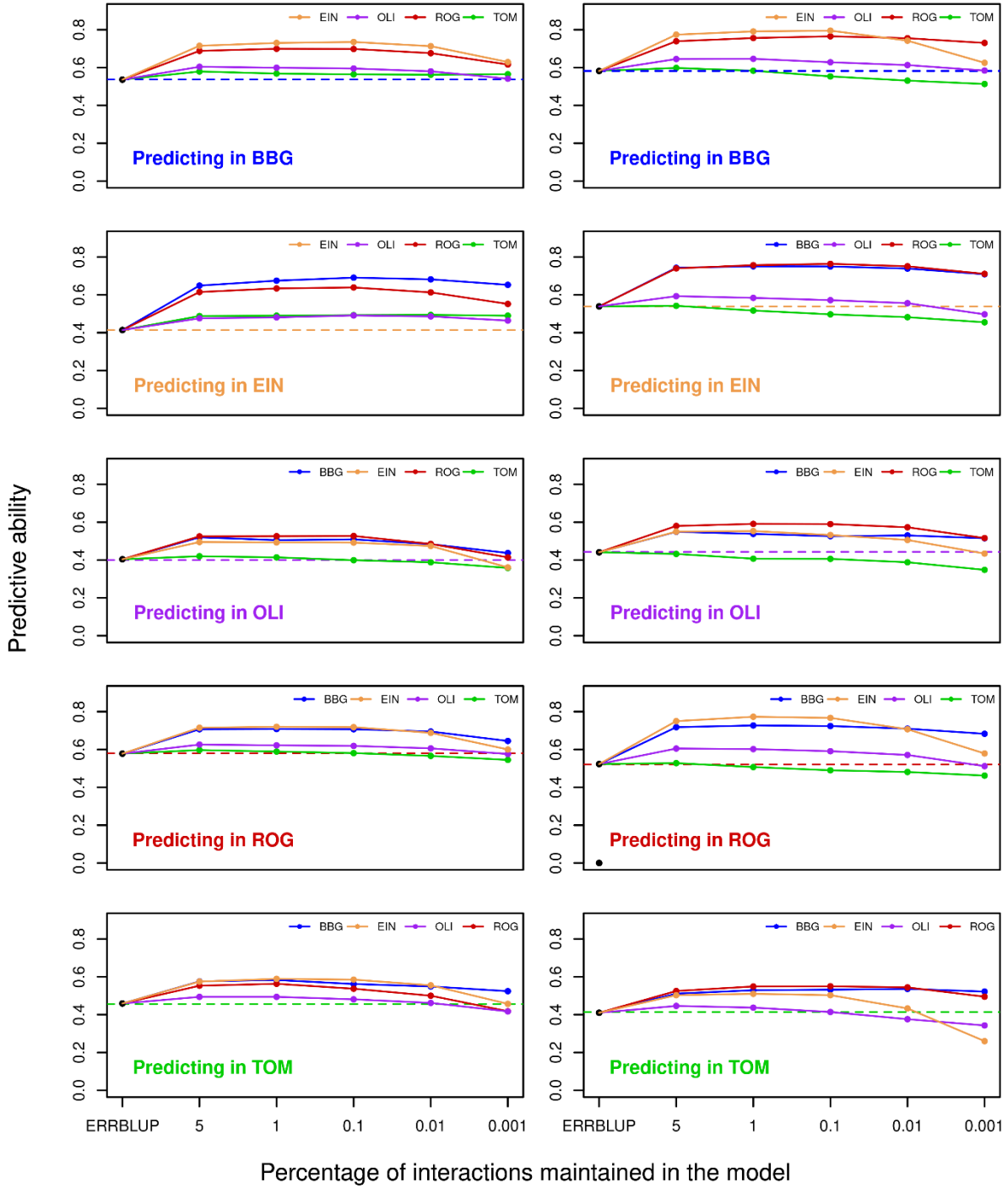
**Fig. S. 3.5:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait EV\_V6 in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection.



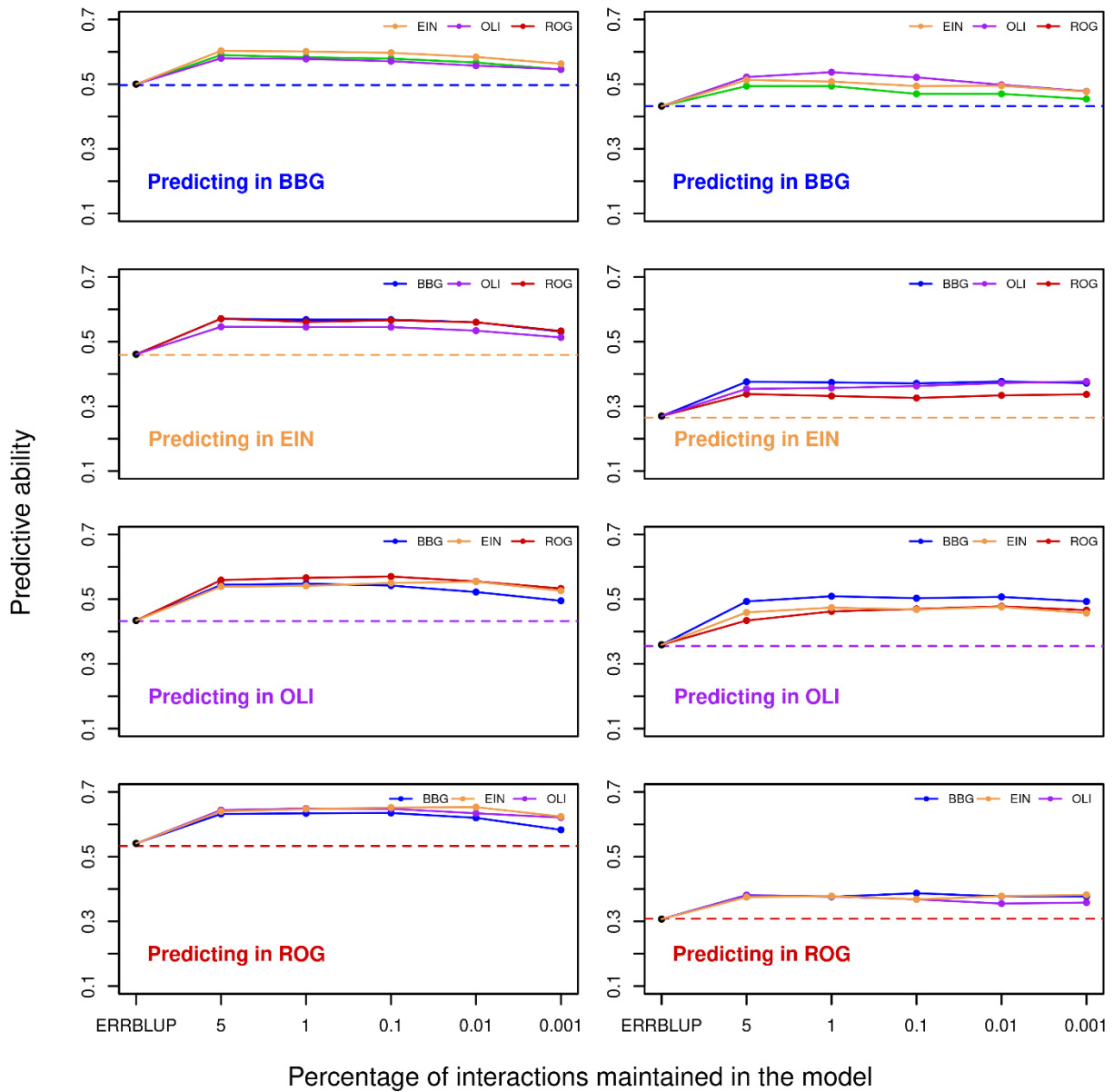
**Fig. S. 3.6:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait PH-V6 in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection.



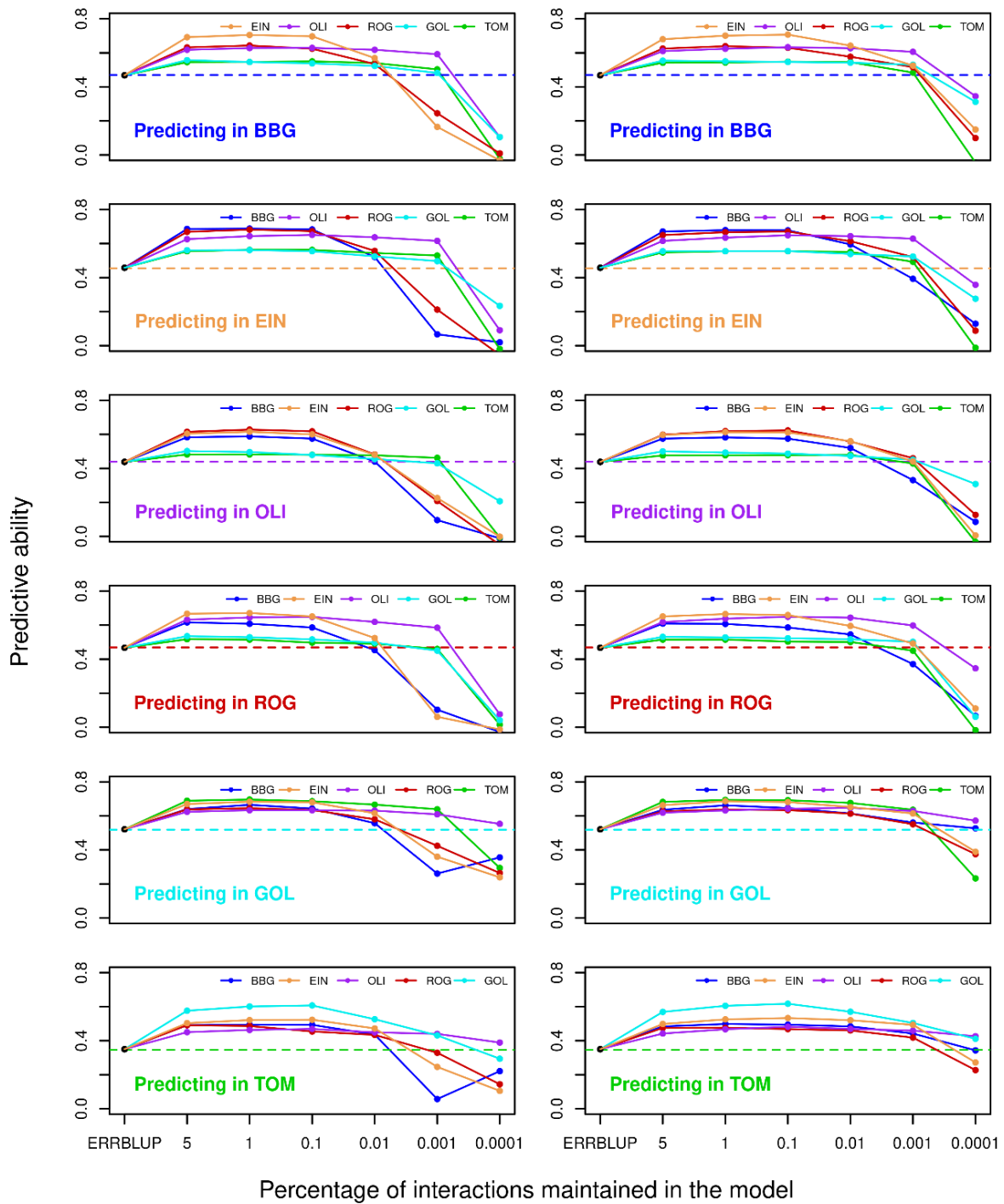
**Fig. S. 3.7:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait PH-final in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection.



**Fig. S. 3.8:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait FF in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection.

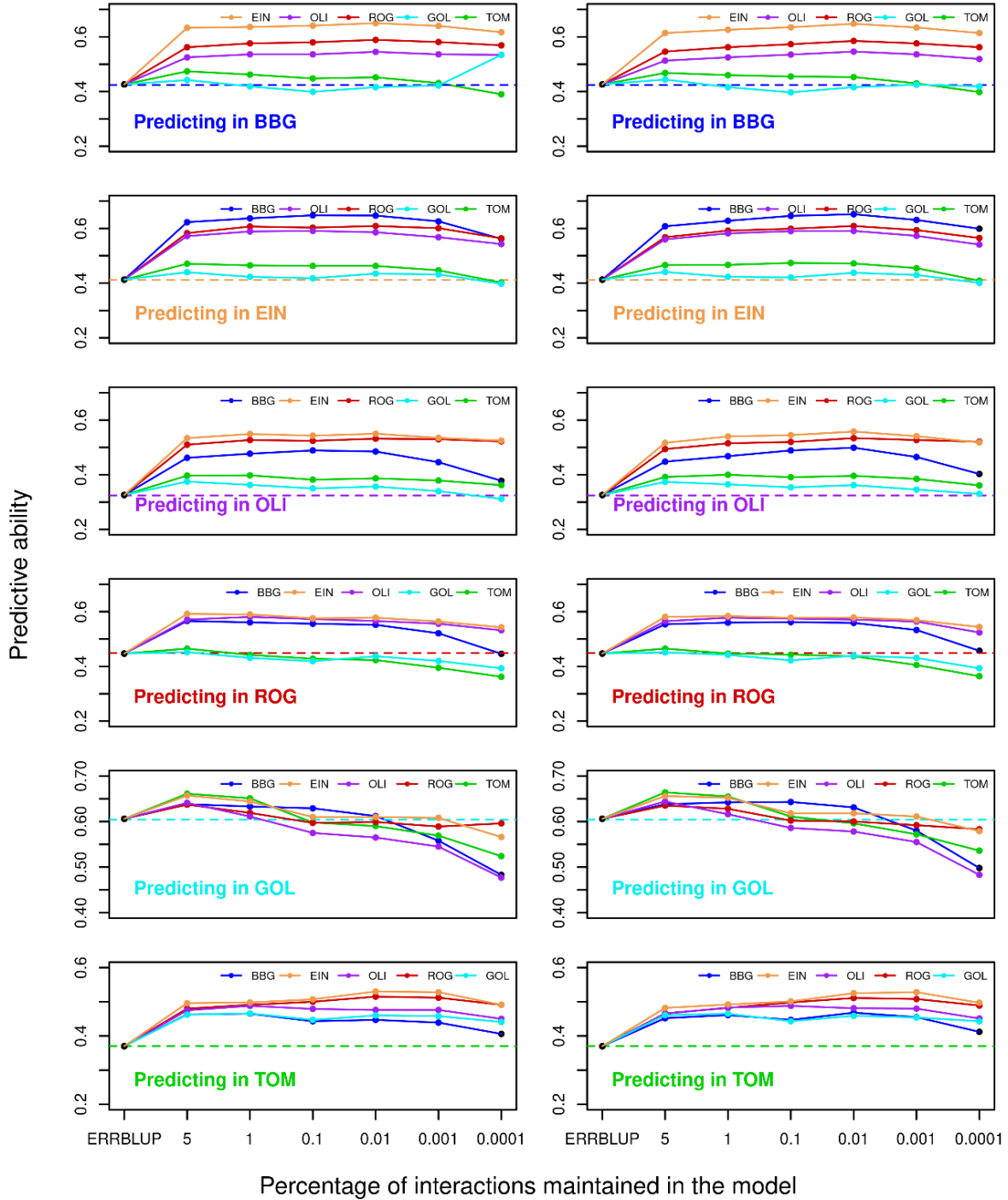


**Fig. S. 3.9:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait RL in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection.

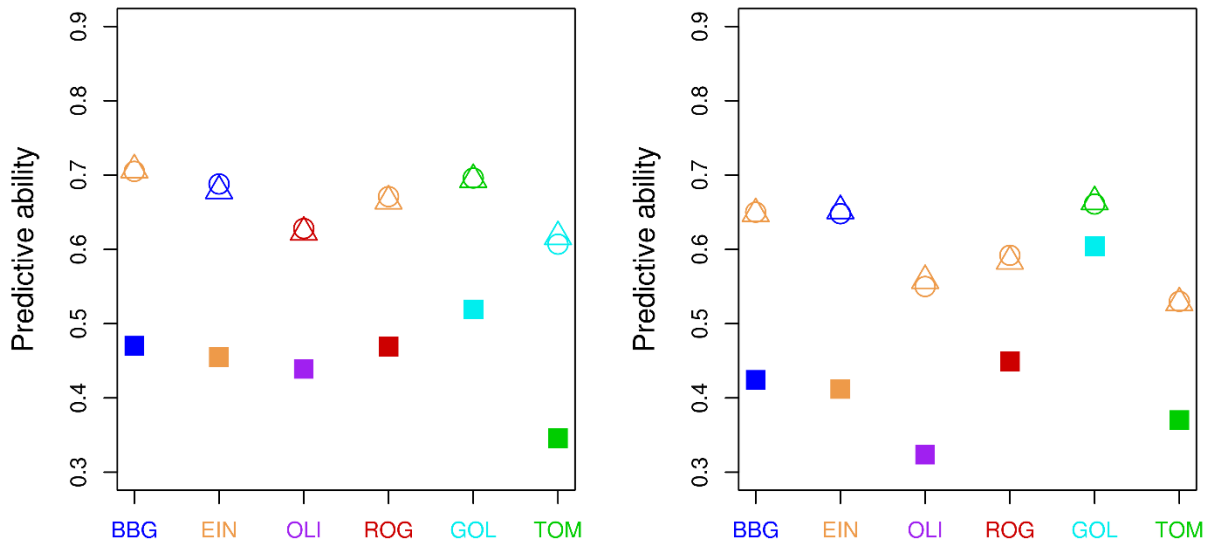


**Fig. S. 3.10:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments (solid colored lines) when SNP interaction selections are based on estimated effects sizes (left side) and estimated effects variances (right side) for trait PH-V4 in KE. In each panel, the solid lines' color indicates the environment in which the relationship matrices were determined by variable selection.

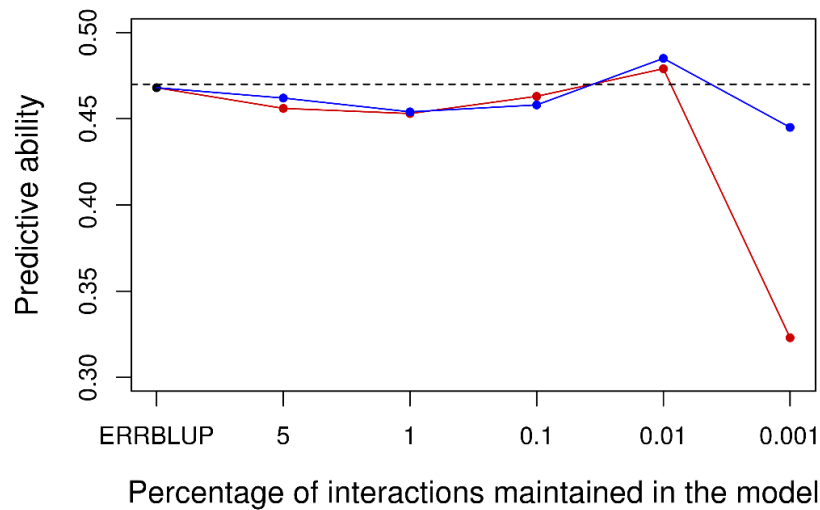




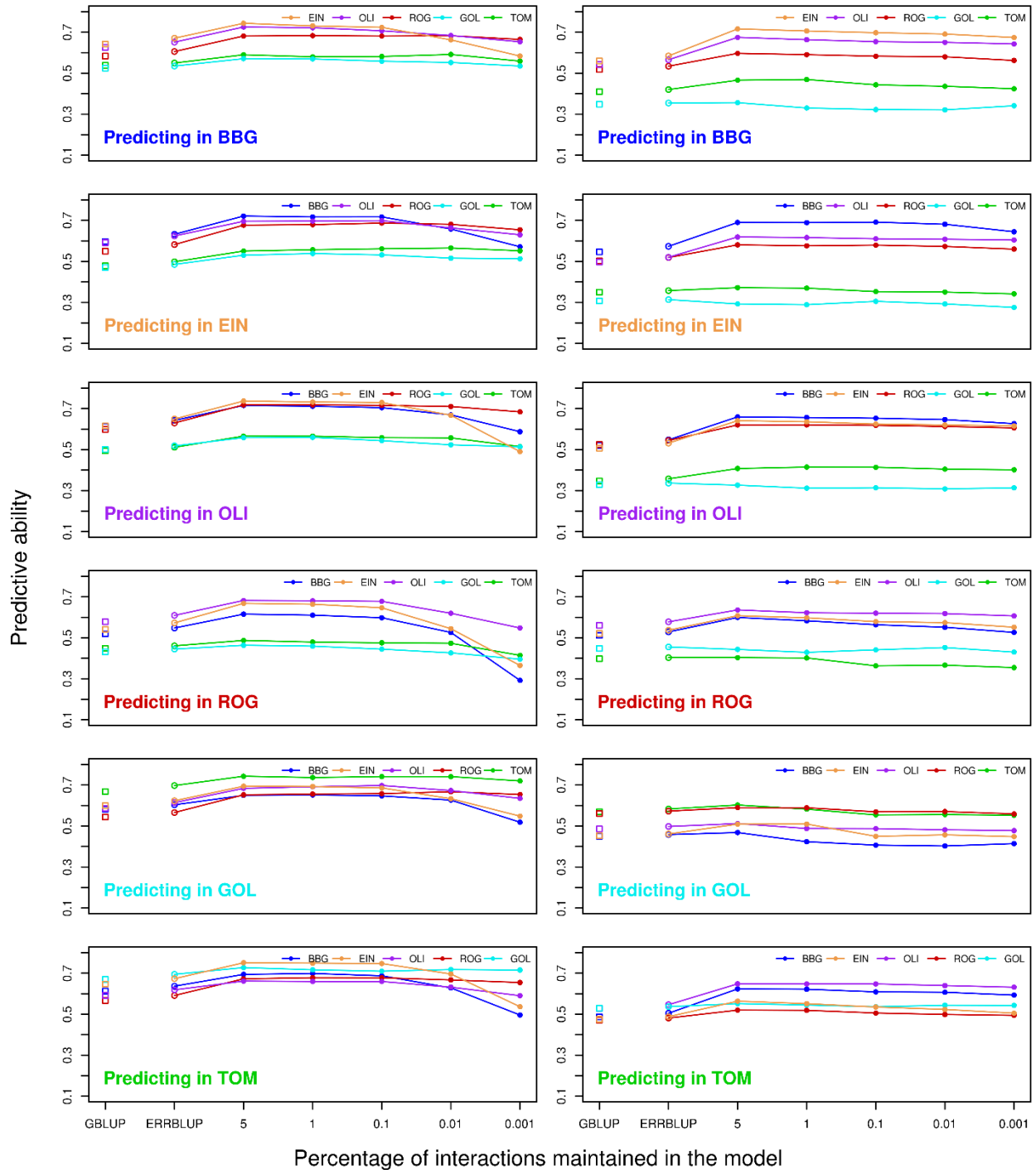
**Fig. S. 3.11:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments (solid colored lines) when SNP interaction selections are based on estimated effects sizes (left side) and estimated effects variances (right side) for trait PH-V4 in PE. In each panel, the solid lines' color indicates the environment in which the relationship matrices were determined by variable selection.



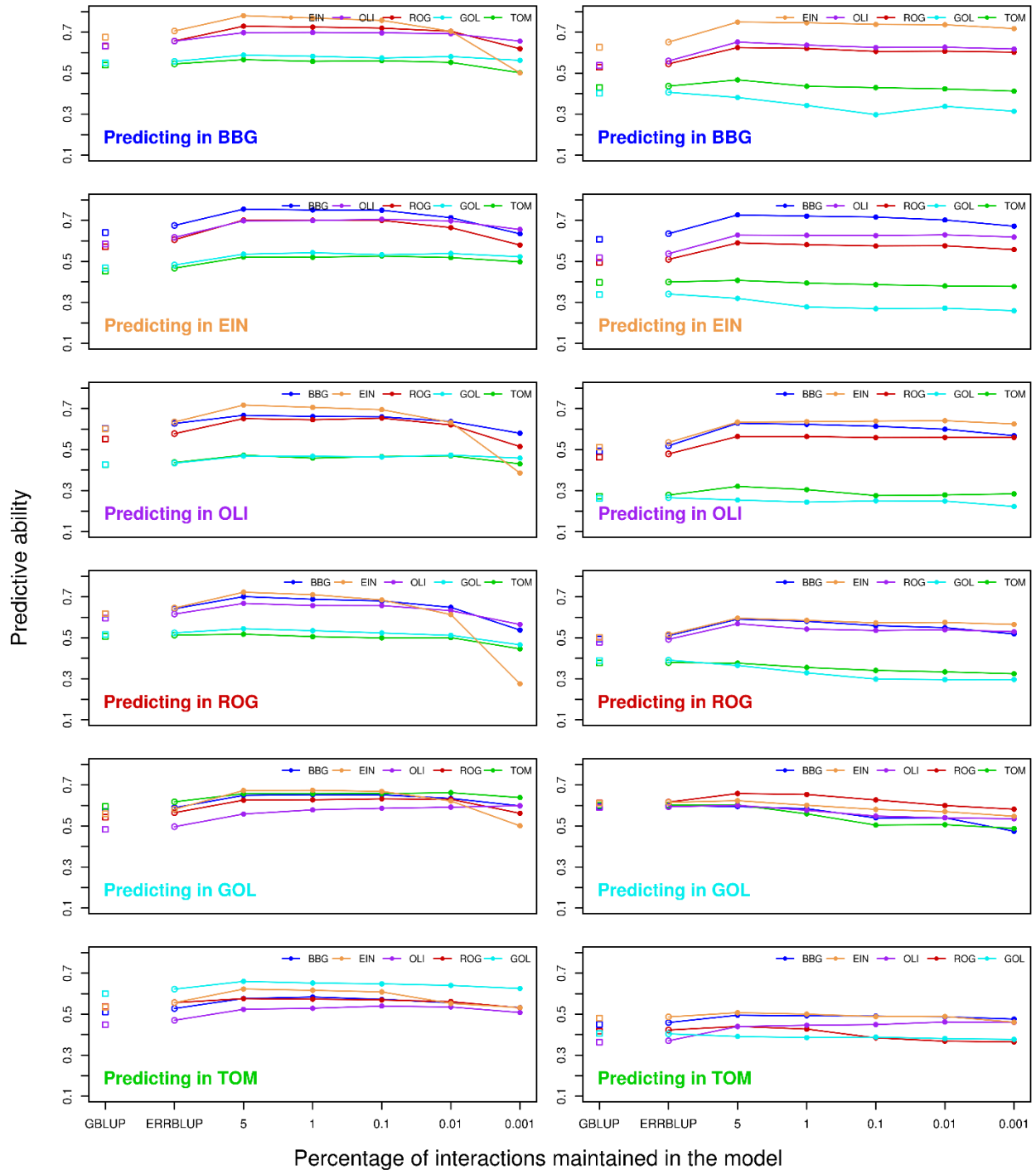
**Fig. S. 3.12:** Comparison of predictive ability of univariate GBLUP within environments (filled squares) and the maximum predictive ability of univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects sizes (circles) and estimated effects variances (triangles) for trait PH-V4 in KE (left side plot) and in PE (right side plot). The colors dark blue, orange, purple, red, light blue and green represent the environments BBG, EIN, OLI, ROG, GOL and TOM, respectively. The circles' and triangles' colors indicate the environment which had the maximum predictive ability for this respective target environment.



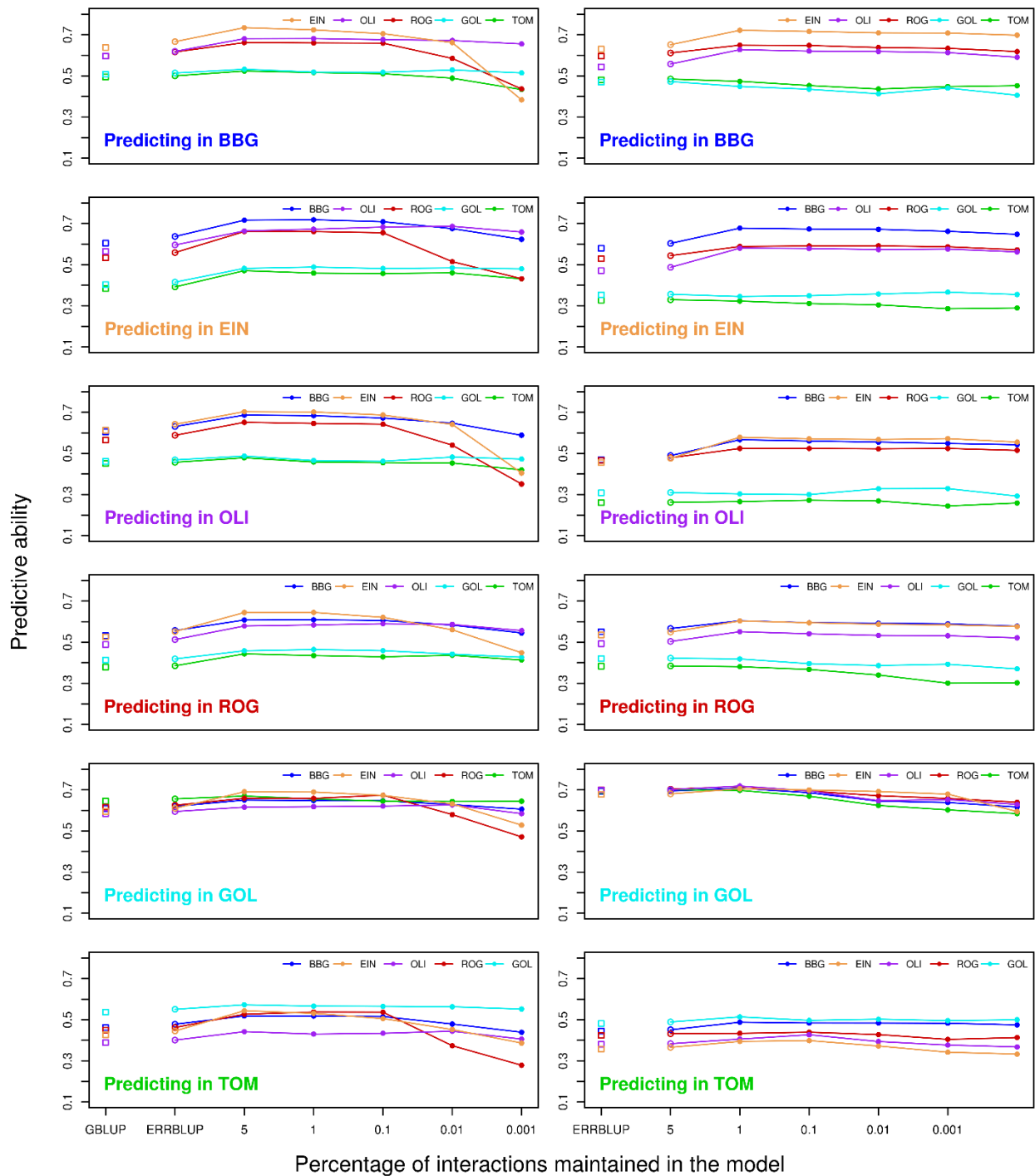
**Fig. S. 3.13:** Predictive ability for univariate GBLUP within Bernburg (dashed horizontal line), univariate ERRBLUP within Bernburg (black filled circle) and univariate sERRBLUP when the SNP interaction selections are based on estimated effects variances (blue solid line) and estimated effect sizes (red solid line) within Bernburg for trait PH-V4 in KE.



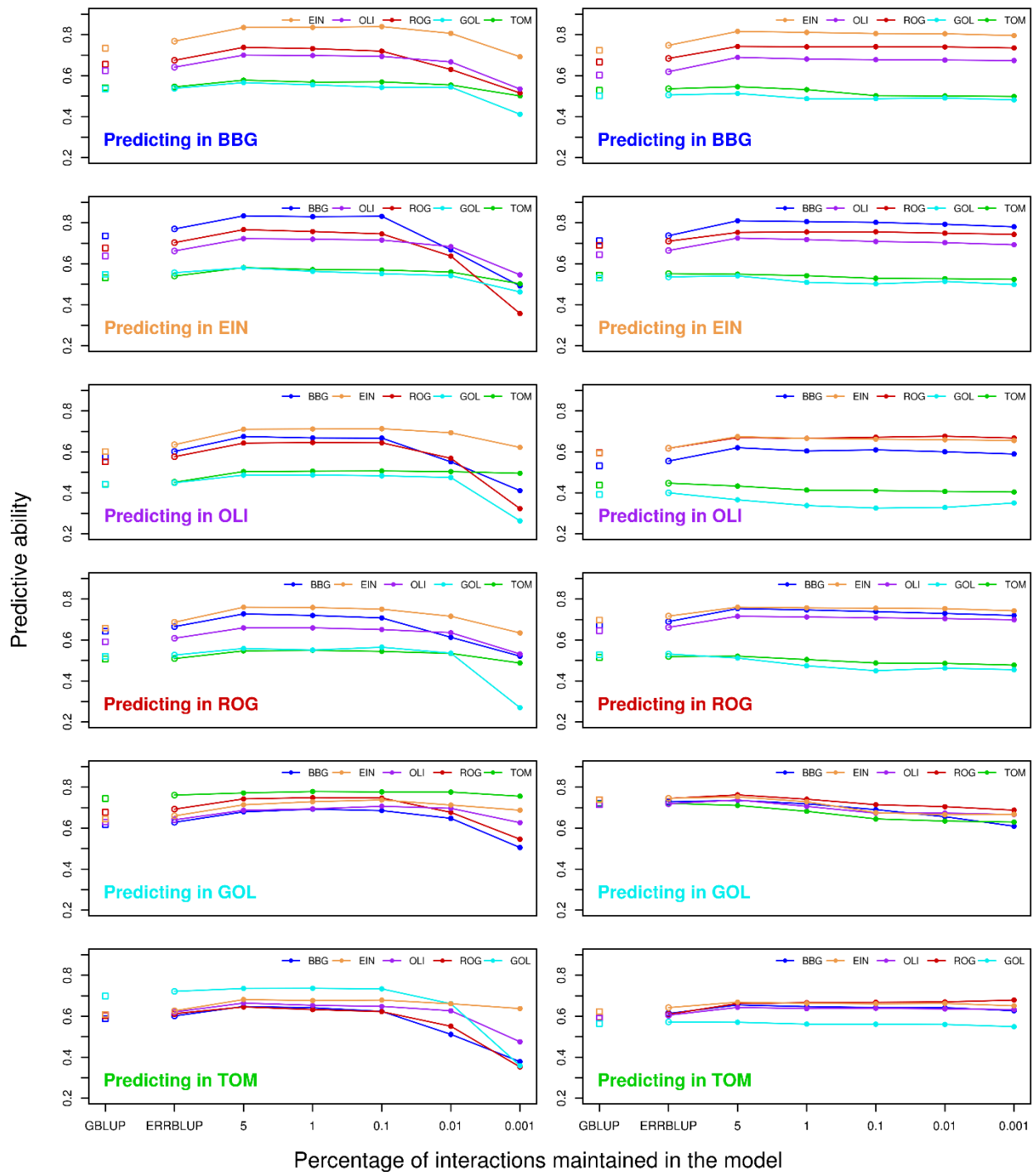
**Fig. S. 3.14:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait EV-V3. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.



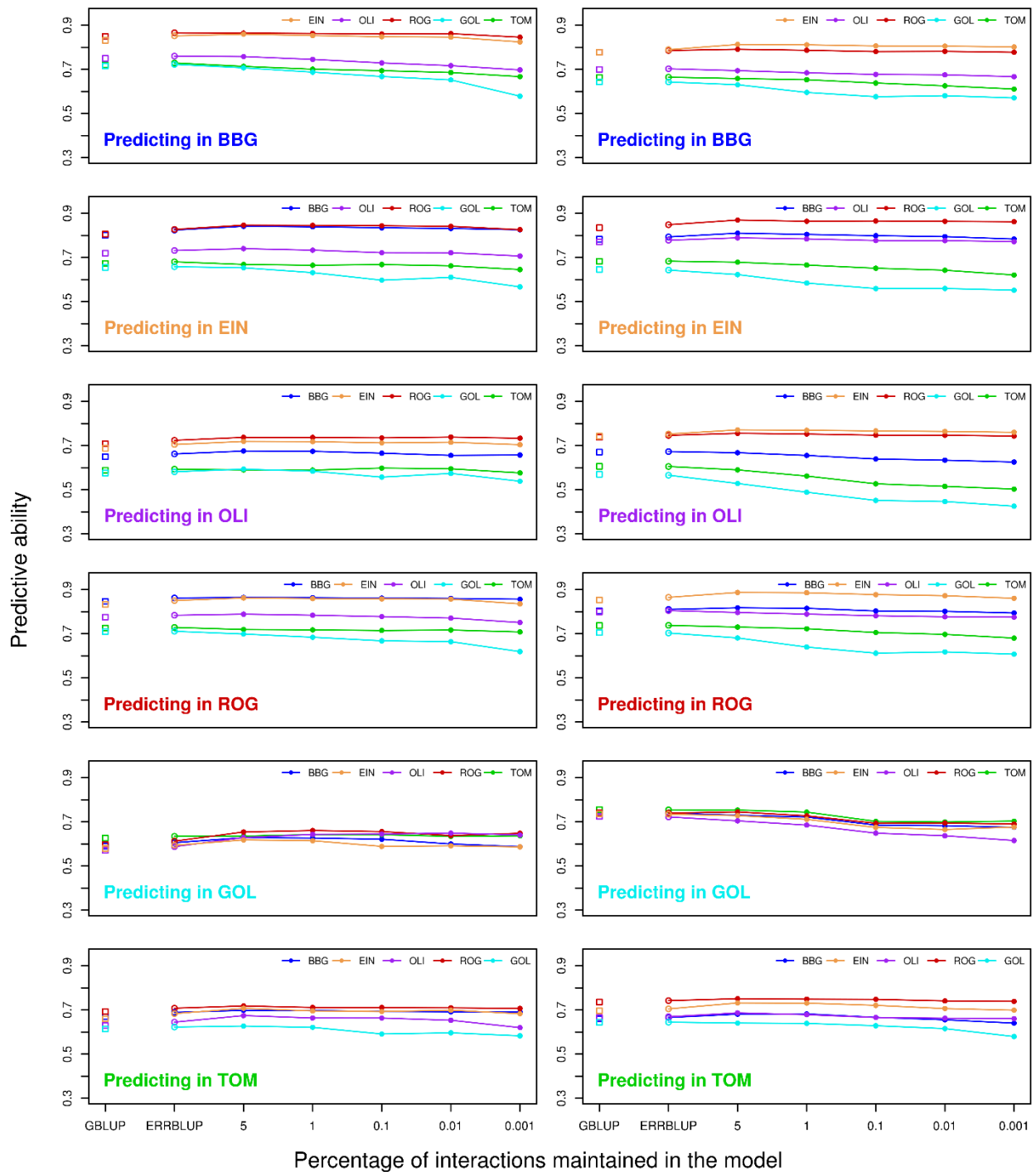
**Fig. S. 3.15:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait EV-V4. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.



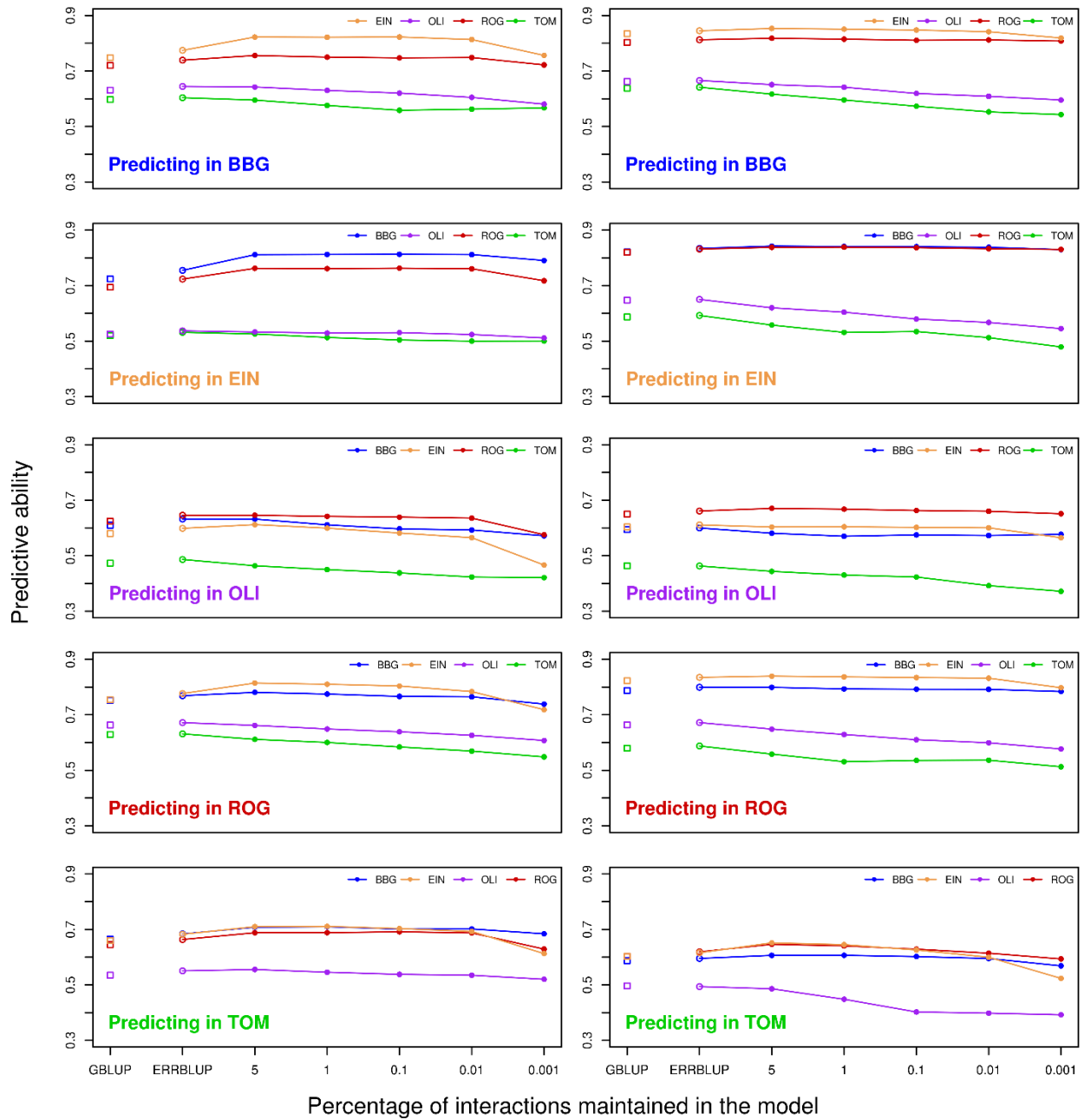
**Fig. S. 3.16:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait EV-V6. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.



**Fig. S. 3.17:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH-V6. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.

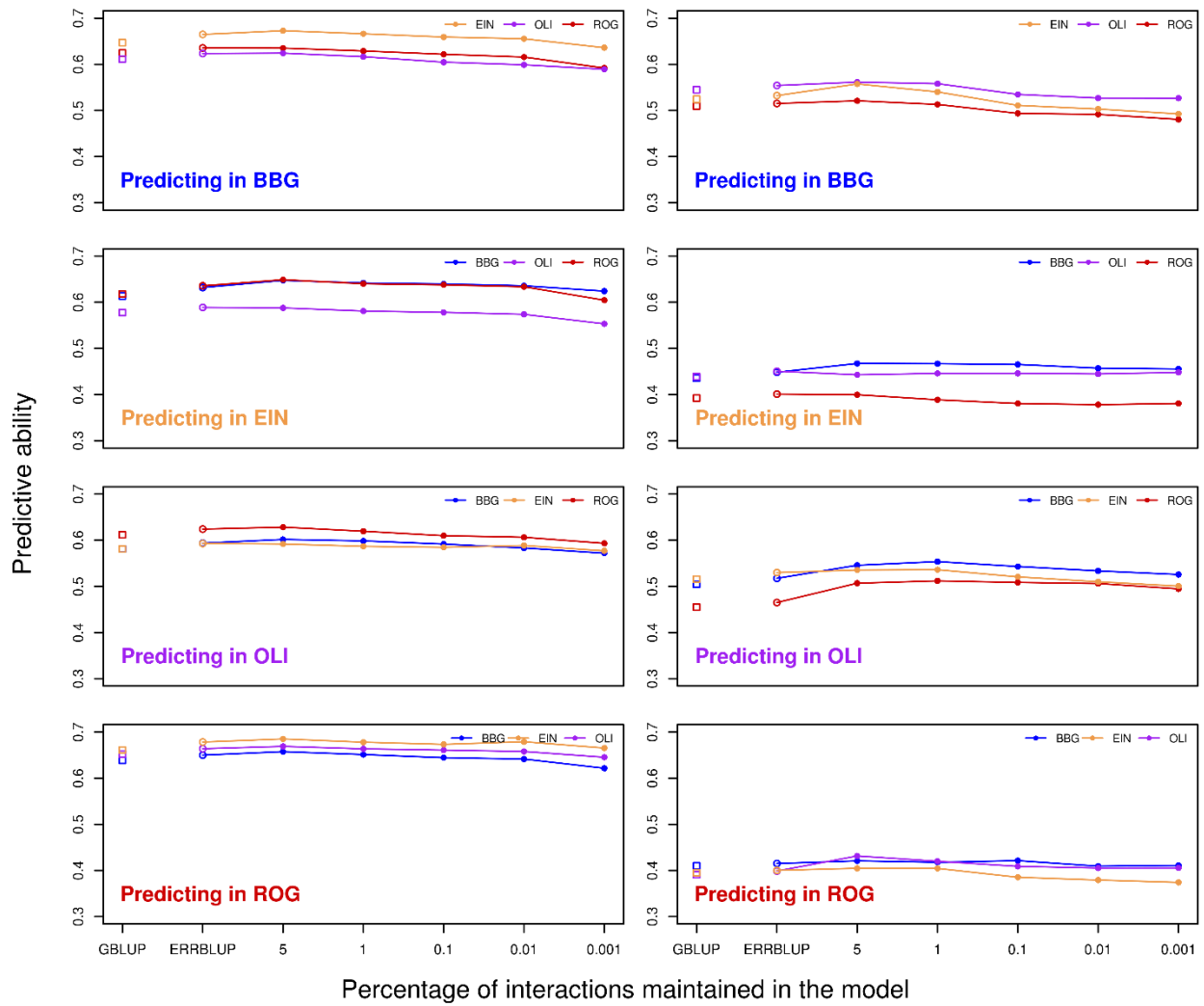


**Fig. S. 3.18:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH-final. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.

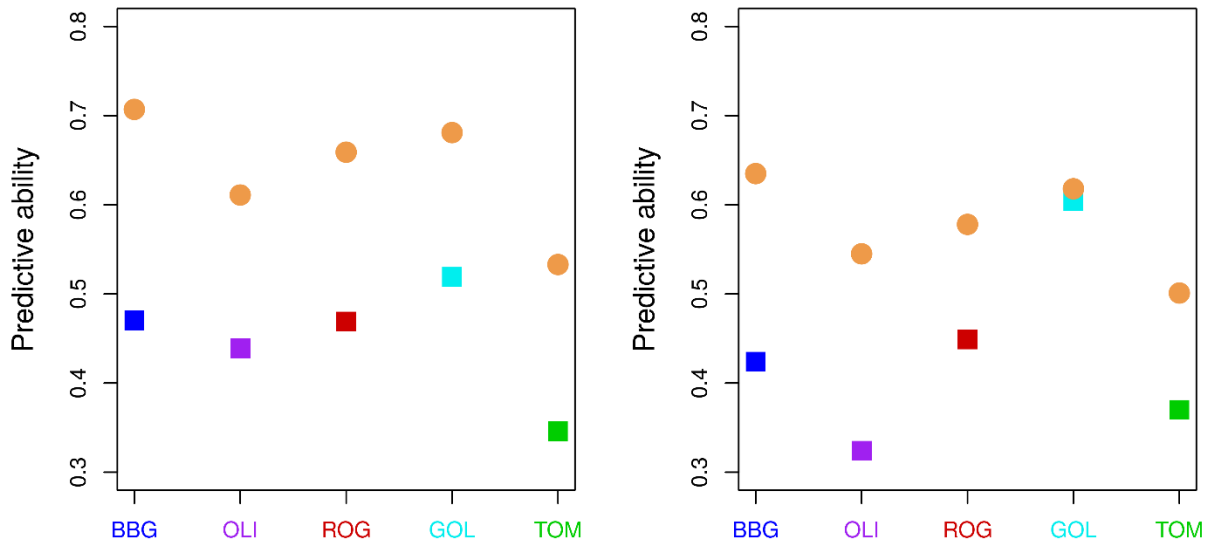


**Fig. S. 3.19:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait FF. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.

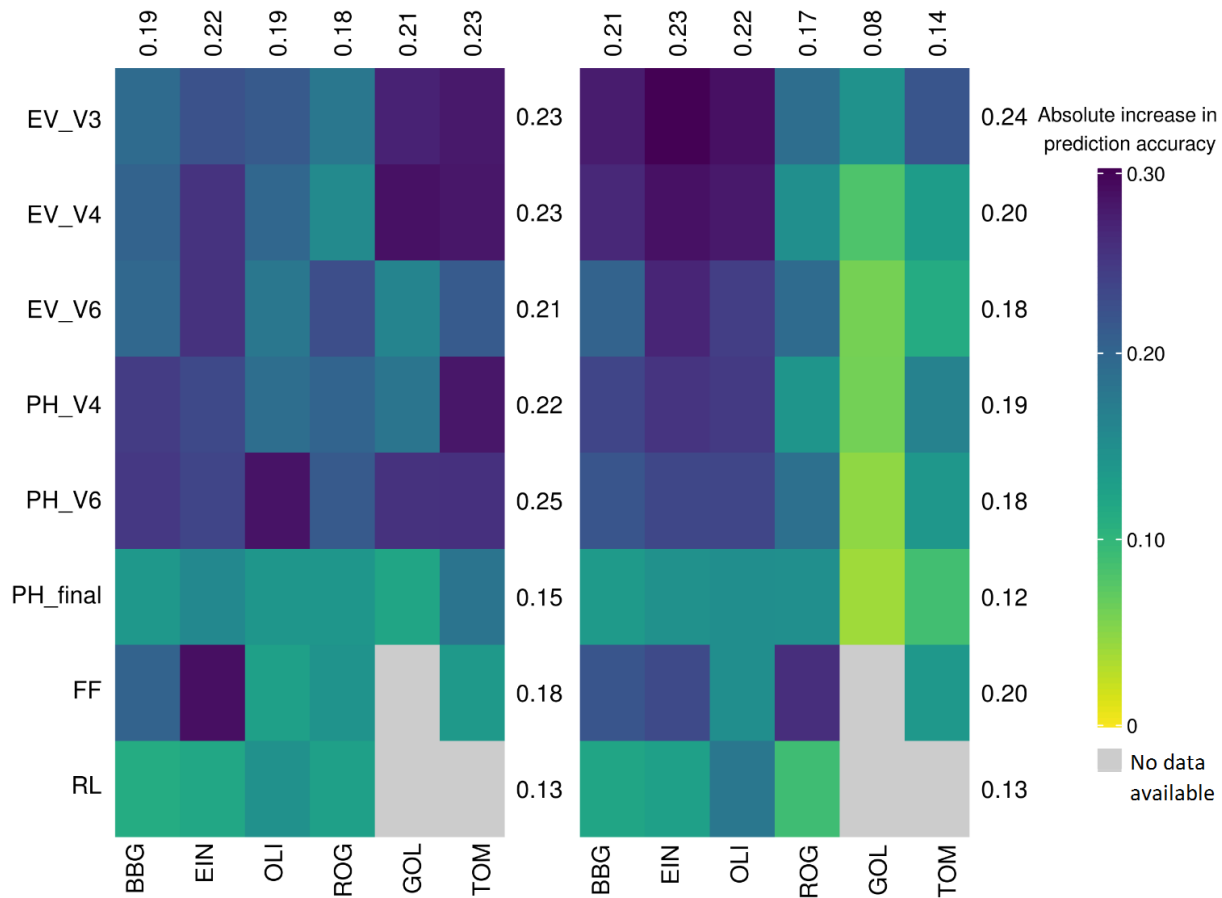




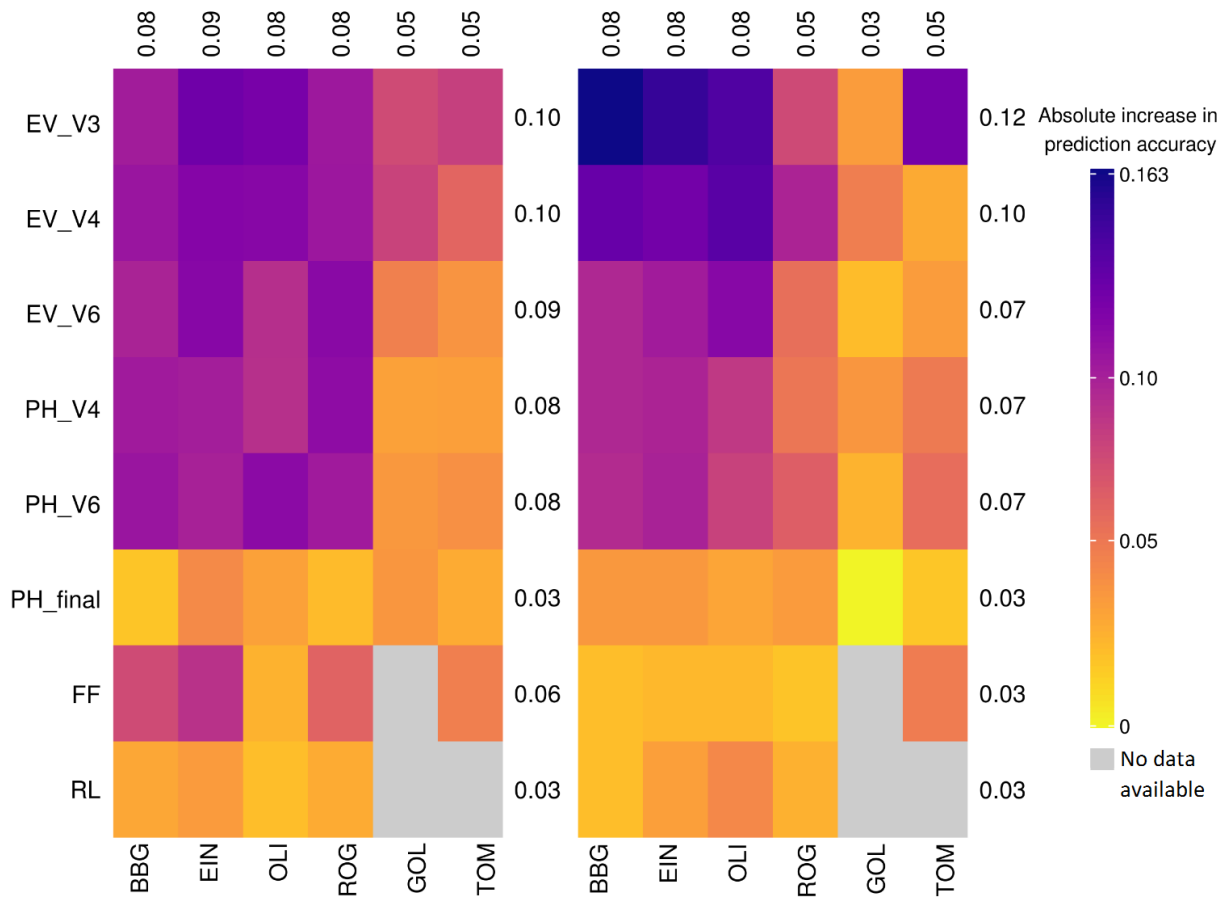
**Fig. S. 3.20:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait RL. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.



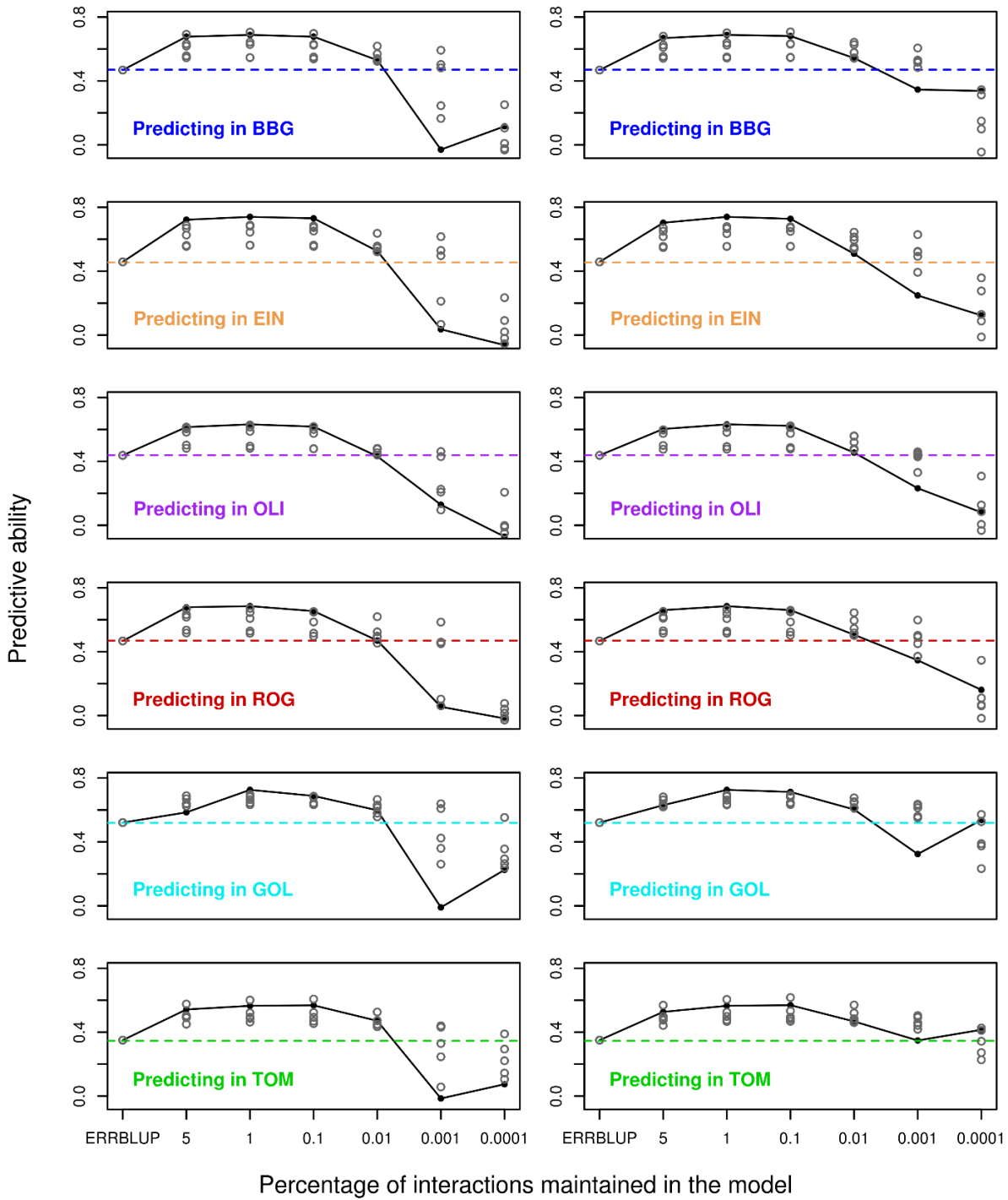
**Fig. S. 3.21:** Comparison of predictive ability of univariate GBLUP within environments (filled squares) and the maximum predictive ability of univariate sERRBLUP across EIN when the SNP interaction selections are based on estimated effects variances (orange circles) for trait PH-V4 in KE (left side plot) and in PE (right side plot). The colors dark blue, purple, red, light blue and green represent the environments BBG, OLI, ROG, GOL and TOM, respectively.



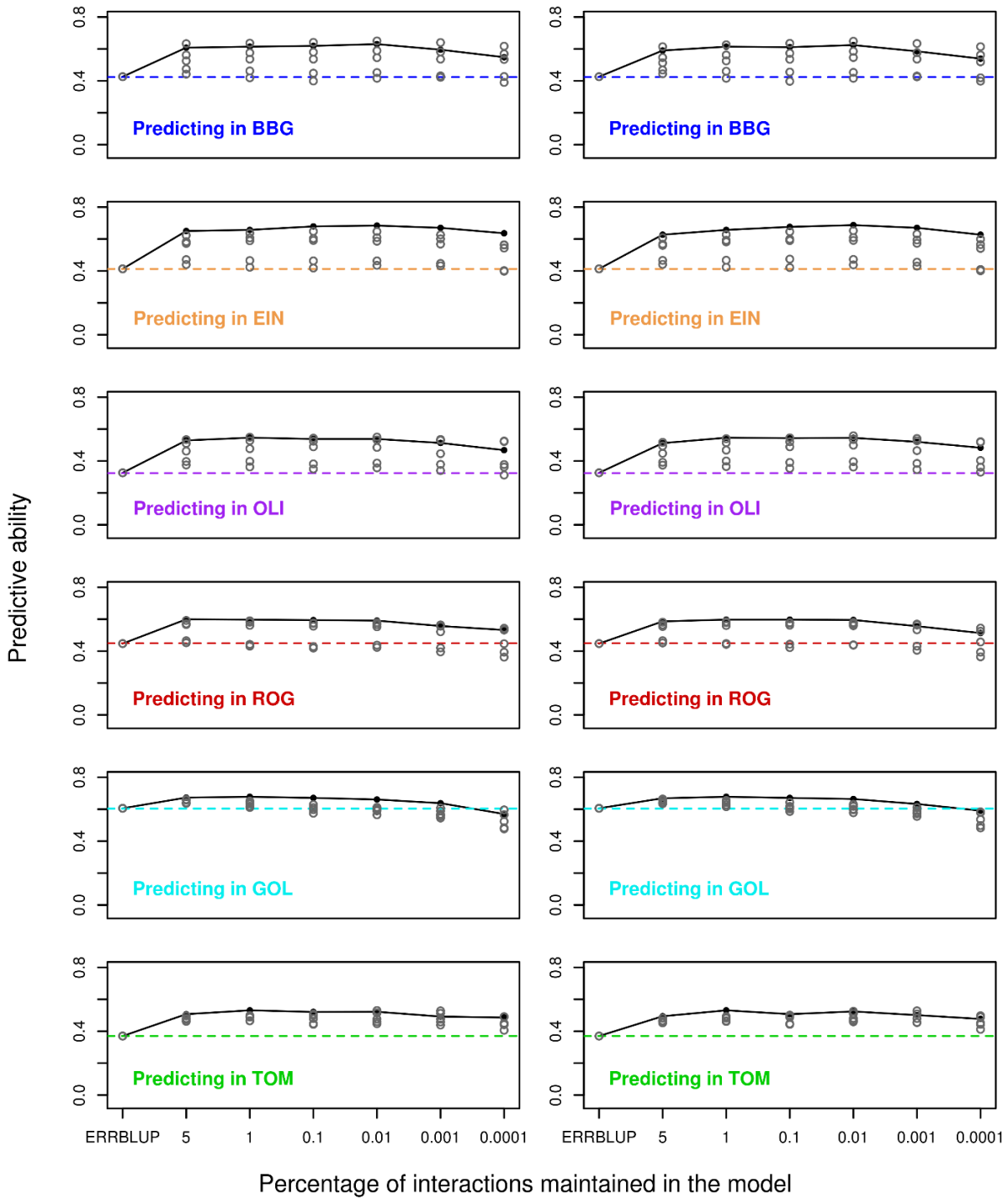
**Fig. S. 3.22:** Absolute increase in prediction accuracy from univariate GBLUP within environments to the maximum prediction accuracy of univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances in KE (left side plot) and in PE (right side plot). The average of absolute increase in prediction accuracy for each trait and environments are display in rows and columns, respectively.



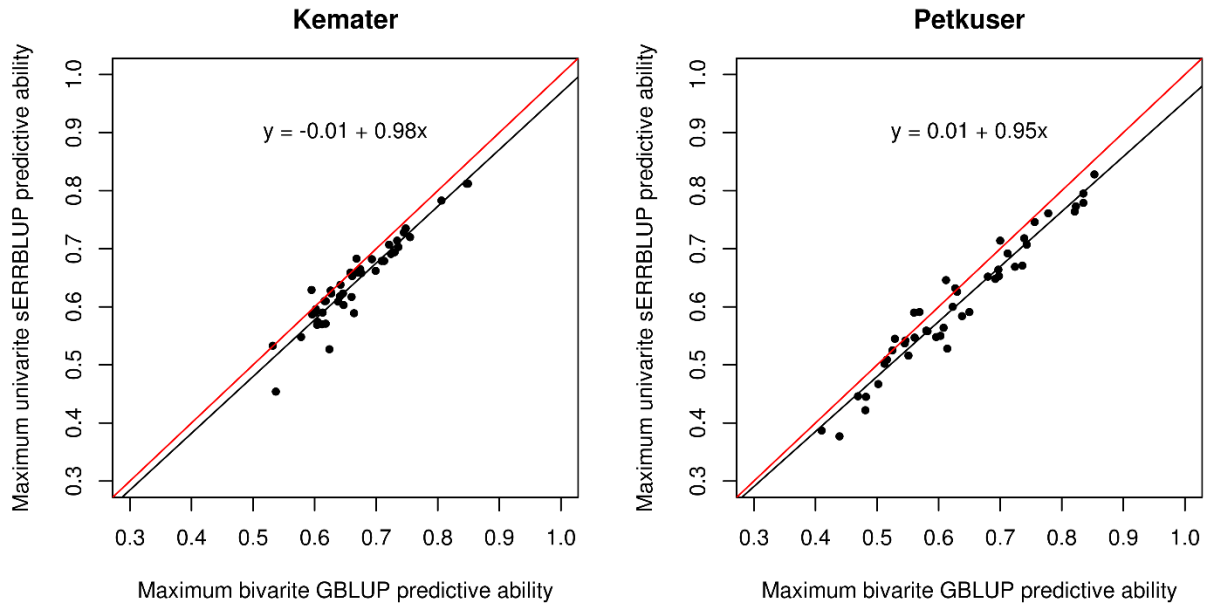
**Fig. S. 3.23:** Absolute increase in prediction accuracy from maximum bivariate GBLUP the maximum prediction accuracy of bivariate sERRBLUP when the SNP interaction selections are based on estimated effects variances in KE (left side plot) and in PE (right side plot). The average of absolute increase in prediction accuracy for each trait and environments are display in rows and columns, respectively.



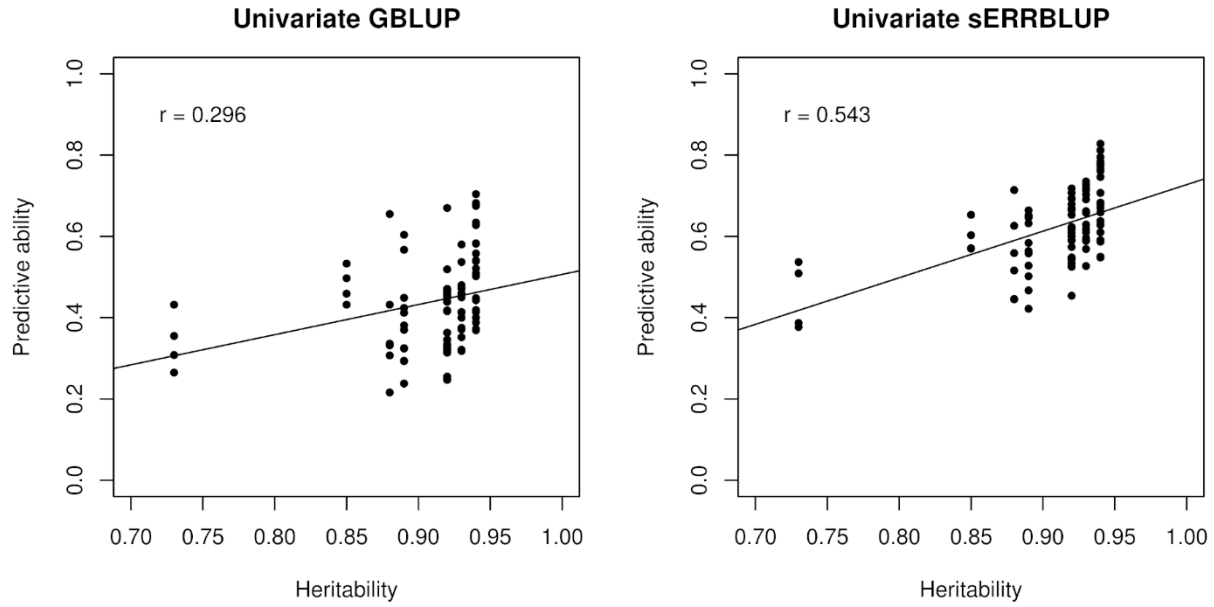
**Fig. S. 3.24:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (gray open circle), univariate sERRBLUP using a single environment for selecting the SNP interactions (gray open circles) and univariate sERRBLUP using all 5 environments jointly (filled black circles and solid line) for the SNP interaction selection based on estimated effects sizes (left side) and estimated effects variances (right side) for trait PH-V4 in KE.



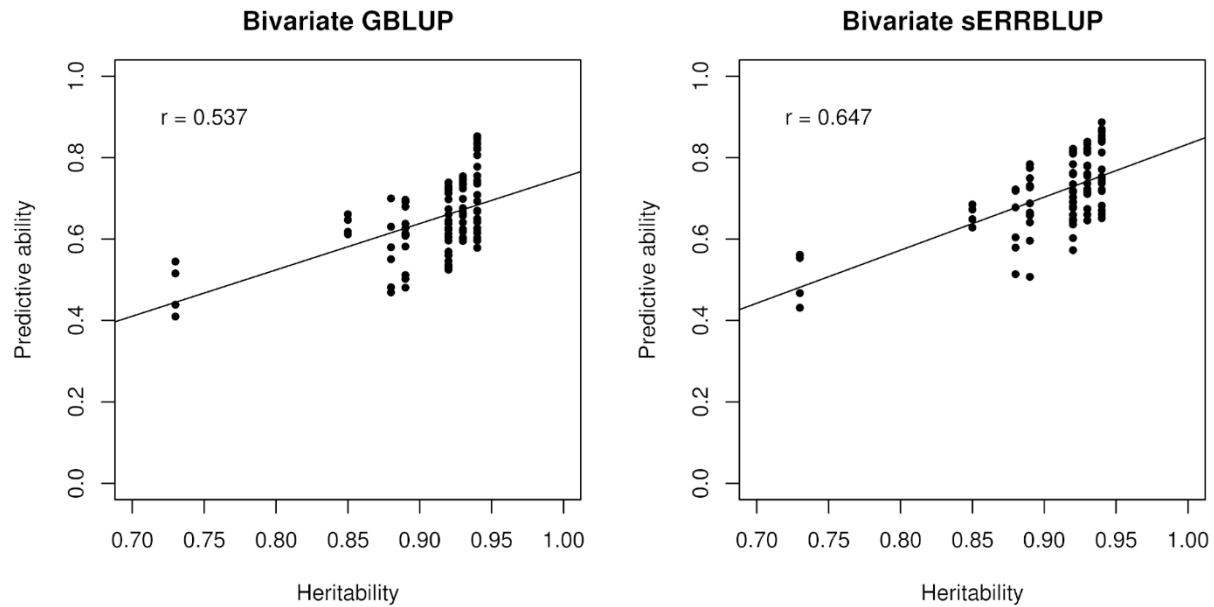
**Fig. S. 3.25:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (gray open circle), univariate sERRBLUP using a single environment for selecting the SNP interactions (gray open circles) and univariate sERRBLUP using all 5 environments jointly (filled black circles and solid line) for the SNP interaction selection based on estimated effects sizes (left side) and estimated effects variances (right side) for trait PH-V4 in PE.



**Fig. S. 3.26:** The comparison between the maximum predictive abilities of bivariate GBLUP within all six environments and the maximum predictive abilities of univariate sERRBLUP across environments for all traits in KE (left side plot) and PE (right side plot). In each plot, the diagonal line (red line) and the overall linear regression line (black line) with the regression formula are shown.



**Fig. S. 3.27:** The correlation between all eight traits' heritabilities and predictive abilities of univariate GBLUP within environments (left side) and maximum predictive abilities of univariate sERRBLUP across environments (right side) in both landraces. The black lines indicate the overall linear regression lines.



**Fig. S. 3.28:** The correlation between all eight traits' heritabilities and the maximum predictive abilities of bivariate GBLUP (left side) and maximum predictive abilities of bivariate sERRBLUP (right side) for all environments in both landraces. The black lines indicate the overall linear regression lines.



## **4 Bivariate genomic prediction of phenotypes by selecting epistatic interactions across years based on haplotype blocks and pruned sets of SNPs**

This chapter contains the manuscript “Bivariate genomic prediction of phenotypes by selecting epistatic interactions across years” has been submitted to the journal TAG (Theoretical and Applied Genetics). In order to have uniform style in the thesis, the journal style is not used in this chapter.

This manuscript is the joint work of Elaheh Vojgani<sup>1\*</sup>, Torsten Pook<sup>1</sup>, Armin C. Hölker<sup>2</sup>, Manfred Mayer<sup>2</sup>, Chris-Carolin Schön<sup>2</sup>, Henner Simianer<sup>1</sup> which is focused on prediction across years through epistasis models in bivariate statistical framework and investigation on the influential factors on bivariate model’s predictive ability.

1: University of Goettingen, Center for Integrated Breeding Research, Animal Breeding and Genetics Group, Goettingen, Germany

2: Plant Breeding, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

### **Author contributions by EV**

EV analyzed the data, generated the pruned set of data, proposed epistasis models based on haplotype blocks, derived the results, wrote the initial manuscript and led the revision of the manuscript.

## 4.1 Key Message

A bivariate epistasis model increases the prediction accuracy similarly for selected subsets of epistatic interactions when utilizing haplotype blocks and pruned sets of SNPs by incorporating phenotypic data across years.

## 4.2 Abstract

The importance of accurate genomic prediction of phenotypes in plant breeding is undeniable, as higher prediction accuracy can increase selection responses. In this study, we investigated the ability of three models to improve prediction accuracy by including phenotypic information from the last growing season. This was done by considering a single biological trait in two growing seasons (2017 and 2018) as separate traits in a multi-trait model. Thus, bivariate variants of the Genomic Best Linear Unbiased Prediction (GBLUP) as an additive model, Epistatic Random Regression BLUP (ERRBLUP) and selective Epistatic Random Regression BLUP (sERRBLUP) as epistasis models were compared with respect to their prediction accuracies for the second year. The results indicate that bivariate ERRBLUP is almost identical to bivariate GBLUP in prediction accuracy, while bivariate sERRBLUP has the highest prediction accuracy in most cases. The obtained prediction accuracies were similar when utilizing pruned sets of SNPs and haplotype blocks, while utilizing haplotype blocks reduces the computational load significantly compared to utilizing pruned sets of SNPs. The prediction accuracies of bivariate GBLUP, ERRBLUP and sERRBLUP have been assessed across eight phenotypic traits and studied datasets from 471/402 doubled haploid lines in the European maize landrace Kemater Landmais Gelb/Petkuser Ferdinand Rot. We further investigated the genomic correlation, phenotypic correlation and trait heritability as factors affecting the bivariate models' prediction accuracy, with genetic correlation between growing seasons being the most important one. For all three considered model architectures results were far worse when using a univariate version of the model.

### Keywords:

Epistasis, Bivariate GBLUP, Prediction across years, Genomic correlation, Haplotype blocks

## 4.3 Introduction

In plant breeding, genomic prediction has become a regular tool (Bernal-Vasquez *et al.*, 2014; Stich and Ingheland, 2018) which enables the optimization of phenotyping costs of breeding programs (Akdemir and Isidro-Sánchez, 2019). The importance of genomic prediction of phenotypes is not restricted to plants. Livestock (Daetwyler *et al.*, 2013) and human research (de los Campos, Vazquez, *et al.*, 2013) also have been widely developed in this regard. In the context of plant and animal breeding, accurately predicting phenotypic traits is of special importance, since

raising all animals and growing all crops to measure their performances requires a considerable amount of money under limited resources (Martini *et al.*, 2016).

Several statistical models have been compared over the last decades in the term of prediction accuracy. In this context, genomic best linear unbiased prediction (GBLUP) (Meuwissen *et al.*, 2001; VanRaden, 2007) as an additive linear mixed model has been widely used due to its high robustness, computing speed and superiority in predictive ability to alternative prediction models like Bayesian methods, especially in small reference populations (Da *et al.*, 2014; Rönnegård and Shen, 2016; Covarrubias-Pazarán *et al.*, 2018; Wang *et al.*, 2018). Furthermore, the inclusion of genotype  $\times$  environment interaction into additive genomic prediction models can result in an increase in prediction accuracy (Hallauer *et al.*, 2010; Bajgain *et al.*, 2020). Such approaches allow borrowing information across environments which potentially leads to higher accuracy in phenotype prediction in multi environment models (Burgueño *et al.*, 2012). In fact, multivariate mixed models have been originally proposed in the context of animal breeding (Henderson and Quaas, 1976) with the purpose of modeling the genomic correlation among traits, longitudinal data, and modeling genotype by environment interactions across multiple years or environments (Mrode, 2014; Lee and van der Werf, 2016; Covarrubias-Pazarán *et al.*, 2018). A multivariate GBLUP model was reported to have higher prediction accuracy than univariate GBLUP (Jia and Jannink, 2012) when the genetic correlations were medium (0.6) or high (0.9) (Covarrubias-Pazarán *et al.*, 2018). It was also shown that aggregating the phenotypic data over years to train the model and predict the performance of lines in the following years is a possible approach which can improve prediction accuracy (Auinger *et al.*, 2016; Schrag *et al.*, 2019b).

In addition, the inclusion of epistasis, defined as the interaction between loci (Falconer and Mackay, 1996; Lynch and Walsh, 1998), into the genomic prediction model results in more accurate phenotype prediction (Hu *et al.*, 2011; Wang *et al.*, 2012; Mackay, 2014; Martini *et al.*, 2016; Vojgani *et al.*, 2021) due to the considerable contribution of epistasis in genetic variation of quantitative traits (Mackay, 2014). In this context, several statistical models have been proposed. Extended genomic best linear unbiased prediction (EG-BLUP, Jiang and Reif 2015) and categorical epistasis (CE, Martini *et al.* 2017) models are using a marker-based epistatic relationship matrix that is constructed in a highly efficient manner. It has been shown that the CE model is as good as or better than EG-BLUP and does not possess undesirable features of EG-BLUP such as coding-dependency (Martini *et al.*, 2017).

Moreover, it was shown that the accuracy of the epistasis genomic prediction model can be increased in one environment by variable selection in another environment (Martini *et al.*, 2016). In this approach, the full epistasis model was reduced to a model with a subset of the largest epistatic interaction effects, resulting in an increase in predictive ability (Martini *et al.*, 2016), through borrowing information across environments. Vojgani *et al.* (2021) showed that the prediction accuracy can be increased even further by selecting the interactions with the highest absolute effect sizes / variances in the epistasis model. The resulting higher computational needs were offset by the development of a highly efficient software package “EpiGP” (Vojgani *et al.*, 2019) to perform computations in a bit-wise manner (Schlather, 2020). Thus, enabling to conduct

such predictions with data sets of practically relevant size across environments in the same year, both with respect to sample size and number of markers (Vojgani *et al.*, 2021). As the number of interactions to account for increases quadratically with the number of included variables, the computational load of methods like EpiGP can quickly go out of control as a model with 600.000 SNPs, as present in high density arrays (Kranis *et al.*, 2013; Unterseer *et al.*, 2014), would result in more than a hundred billion interactions to account for. The most common methods for variable reduction applied here is LD pruning (Purcell *et al.* 2007), but new linkage-based haplotyping methods (Pook *et al.* 2019) have recently been proposed to even further reduce the dimensionality of genomic data without much information loss.

The aim of this study is to assess the bivariate genomic prediction models which incorporate epistatic interactions with the target of borrowing information across years to maximize the predictive ability based on both, a pruned set of SNPs and haplotype blocks. Since the accuracy of genomic prediction of phenotypes was shown to be increased by both borrowing information across environments and years (Covarrubias-Pazarán *et al.*, 2018; Schrag *et al.*, 2019a) and inclusion of epistasis into the prediction model (Martini *et al.*, 2016; Vojgani *et al.*, 2020), we combine these two approaches to make the best use of the available information. We further aim to assess the optimum proportion of epistatic interactions to be kept in the model in the variable selection step across years and compare the obtained predictive ability when utilizing pruned sets of SNPs and haplotype blocks. The data used for this purpose were generated in multi-location trials of doubled haploid (DH) lines generated from two European maize landraces in 2017 and 2018.

## **4.4 Materials and Methods**

### **4.4.1 Data used for analysis**

A set of 948 doubled haploid lines of the European maize landraces Kemater Landmais Gelb (KE, Austria, 516 lines) and Petkuser Ferdinand Rot (PE, Germany, 432 lines) were genotyped with the 600 k Affymetrix® Axiom® Maize Array (Unterseer *et al.*, 2014).

After quality filtering and imputation, 910 DH lines remained (501 lines in KE and 409 lines in PE) and the panel of markers reduced to 501,124 markers (Hölker *et al.*, 2019). Additionally, loci that were in high level of pairwise linkage disequilibrium (LD) were removed (Calus and Vandenplas, 2018) through linkage disequilibrium based SNP pruning with PLINK v1.07 (Purcell *et al.*, 2007; Chang *et al.*, 2015). LD pruning was done by the parameters of 50, 5 and 2 which considered as the SNPs window size, the number of SNPs at which the SNP window shifts and the variance inflation factor, respectively. This resulted in a data panel containing 25'437 SNPs for KE and 30'212 SNPs for PE (Vojgani *et al.*, 2020). Note that even a panel of 25'000 SNPs results in more than 1 billion SNP interactions to account for. Therefore, in order to have further variable

reduction, haplotype blocks as a combination of closely linked markers, which has shown to be an alternative approach for genomic prediction improving the prediction accuracy (Meuwissen *et al.*, 2014; Jiang *et al.*, 2018), have been generated from the full panel of markers with the software HaploBlocker (Pook *et al.* 2019) using default settings. This resulted in a data panel containing 2'972 haplotype blocks in KE and 3'330 haplotype blocks in PE.

Out of 910 genotyped lines only 873 DH lines were phenotyped (471 lines in KE and 402 lines in PE). Einbeck (EIN, Germany), Roggenstein (ROG, Germany), Golada (GOL, Spain) and Tomeza (TOM, Spain) were the four locations that these lines were phenotyped for a series of traits in both 2017 and 2018.

The means, standard deviations, maximum and minimum values of studied phenotypic traits in 2017 and 2018 in each landrace are compared in Table 4.1 which were derived from the Best Linear Unbiased Estimations (BLUEs) of the genotype mean for each phenotypic trait by Hölker *et al.* (2019). The comparison of the respective detailed values for each trait in each environment and landrace in 2017 and 2018 are illustrated in the supplementary (Table S. 4.1). *V<sub>i</sub>* in phenotypic traits represents the vegetative growth stage when *i* leaf collars are visible based on the leaf collar method of the corn growth (Abendroth *et al.*, 2011). Early vigour at V3 stage (EV\_V3), female flowering (FF) and root lodging (RL) were not phenotyped in all four environments for both years. EV\_V3 was not phenotyped in EIN in 2018, FF was not phenotyped in GOL in 2017 and RL was not phenotyped in TOM and GOL in both 2017 and 2018.

The number of phenotyped lines per year and environment for trait PH\_V4, as the main trait in this study, are summarized in Table 4.2. For EIN and ROG a higher number of phenotyped lines were generated in 2017. On the contrary, more lines were phenotypes in GOL and TOM in 2018.

**Table 4.1:** Phenotypic trait description and the mean, minimum, maximum and standard deviation of the BLUEs for each phenotypic trait in KE and PE landraces in the years 2017 and 2018.

Trait	Definition	Landrace	Year	Mean	Minimum	Maximum	Standard deviation
<b>EV_V3</b>	Early vigour at V3 stage scored on scale from 1 (very poor early vigour) to 9 (very high early vigour)	KE	2017	4.94	0.78	9.00	1.35
			2018	5.06	0.32	8.67	1.33
		PE	2017	5.57	1.00	9.03	1.20
			2018	5.47	1.38	8.93	1.13
<b>EV_V4</b>	Early vigour at V4 stage scored on scale from 1 (very poor early vigour) to 9 (very high early vigour)	KE	2017	4.84	0.67	8.29	1.30
			2018	5.08	0.96	8.65	1.30
		PE	2017	5.45	0.93	8.49	1.15
			2018	5.25	1.63	9.07	1.19
<b>EV_V6</b>	Early vigour at V6 stage scored on scale from 1 (very poor early vigour) to 9 (very high early vigour)	KE	2017	5.13	0.54	8.75	1.31
			2018	5.54	1.07	9.60	1.35
		PE	2017	5.64	0.84	8.39	1.12
			2018	5.38	1.07	9.68	1.29
<b>PH_V4</b>	Mean plant height of three plants of the plot at V4 stage in cm	KE	2017	33.10	6.90	88.24	13.95
			2018	42.01	8.48	89.24	16.47
		PE	2017	38.01	11.89	95.30	14.96
			2018	46.19	16.14	93.20	17.78
<b>PH_V6</b>	Mean plant height of three plants of the plot at V6 stage in cm	KE	2017	62.03	8.34	127.54	19.95
			2018	92.27	21.90	173.66	21.04
		PE	2017	69.84	14.78	130.51	19.26
			2018	97.80	50.37	169.71	19.44
<b>PH_final</b>	Final plant height after flowering in cm	KE	2017	139.10	49.27	245.00	27.14
			2018	146.04	35.41	265.02	35.74
		PE	2017	124.09	30.21	211.14	24.54
			2018	128.08	35.76	248.43	35.99
<b>FF</b>	Days after sowing until female flowering (days until 50% of the plot showed silks)	KE	2017	79.72	62.45	102.02	6.27
			2018	76.99	62.22	100.14	6.09
		PE	2017	78.85	59.10	101.50	6.33
			2018	76.70	60.14	93.96	6.52
<b>RL</b>	Root lodging score from 1 to 9 (1 = no lodging and 9= severe lodging)	KE	2017	3.38	0.59	9.58	2.50
			2018	1.42	0.73	8.52	0.90
		PE	2017	2.14	0.03	9.22	1.74
			2018	1.21	0.32	4.69	0.51

**Table 4.2:** Number of KE and PE lines phenotyped in each location for the years 2017 (blue numbers) and 2018 (red numbers) for trait PH\_V4.

	EIN (2017\2018)	ROG (2017\2018)	GOL (2017\2018)	TOM (2017\2018)
Phenotyped lines in KE	462\365	461\365	211\222	211\222
Phenotyped lines in PE	393\365	390\365	204\240	204\240

#### 4.4.2 Statistical models for phenotype prediction

We used the bivariate statistical framework as the basis of the genomic prediction models which has been proposed in the recent work by Vojgani *et al.* (2020). In this regard, GBLUP, ERRBLUP and sERRBLUP as three different methods described in Vojgani *et al.* (2021) were used for genomic prediction of phenotypes which differ in dispersion matrices representing their covariance structure of the genetic effects. GBLUP as an additive model is based on a genomic relationship matrix calculated according to VanRaden (2008). ERRBLUP (Epistatic Random Regression BLUP) as a full epistasis model is based on all pairwise SNP interactions which generates a new marker matrix considered as a marker combination matrix. The marker combination matrix is a 0, 1 matrix indicating the absence (0) or presence (1) of each marker combination for each individual. sERRBLUP (selective Epistatic Random Regression BLUP) as a selective epistasis model is based on a selected subset of SNP interactions (Vojgani *et al.*, 2021). Vojgani *et al.* (2020) proposed estimated effect variances in the training set as the selection criterion of pairwise SNP interactions due to its robustness in predictive ability specifically when only a small proportion of interactions are maintained in the model.

#### 4.4.3 Assessment of genomic prediction models

GBLUP, ERRBLUP and sERRBLUP models have been assessed via 5-fold cross validation by randomly partitioning the original sample into 5 equal size subsamples in which one subsample was considered as the test set to validate the model, and the remaining 4 subsamples were considered as a joint training set (Erbe *et al.*, 2010). The 5-fold cross validation technique was utilized with 5 replicates through which the Pearson correlation between the predicted genetic values and the observed phenotypes in the test set was considered as the predictive ability in each fold of each replicate, which then was averaged across 25 replicates. In this study, predictive ability was separately assessed for KE and PE for a series of phenotypic traits in four different environments. Besides, we calculated the traits' prediction accuracies by dividing their predictive abilities by the square-root of the respective traits' heritabilities (Dekkers, 2007) derived from all environments in both 2017 and 2018 jointly (Table S. 4.2).

Univariate GBLUP within 2018 was assessed by training the model in the same year (2018) as the test set was sampled from. However, bivariate GBLUP, ERRBLUP and sERRBLUP were assessed by training the model with both the training set of the target environment in 2018 and the full

dataset of the respective environment in 2017. The interaction selection step in bivariate sERRBLUP is done by first using the complete dataset of target environment in 2017 to estimate all pairwise SNP interaction effect variances. Then, an epistatic relationship matrix for all lines is constructed based on the subset of top ranked interaction effect variances, which is finally used to predict phenotypes of the target environment test set in 2018 (Vojgani *et al.*, 2020).

#### **4.4.4 Variance component estimation**

Variance component estimation in univariate GBLUP was done by EMMREML (Akdemir and Godfrey, 2015) based on the training set in each run of 5-fold cross validation with 5 replicates. In bivariate models this was done by ASReML-R (Butler *et al.*, 2018) with the approach specified by Vojgani *et al.* (2020) for pre estimating the variance components from the full dataset to derive the initial values for the variance components in ASReML models in 100 iterations for each combination. If the variance estimation based on the full set did not converge after 100 iterations, then the estimated variance components at the 100<sup>th</sup> iteration were extracted as initial values of the bivariate model in the cross validation step. Afterwards, the model used these values to re-estimate the variance components based on the training set in each run of 5-fold cross validation in 50 iterations. The estimated variance components in the converged models based on the full set deviated only slightly from the estimated variance components based on the training set (Fig. S. 4.1). However, the variance component estimations did not converge in all folds of 5-fold cross validation with 5 replicates. In such cases, the initial values were set as the fixed values for the model to predict the breeding values. This approach appears justifiable in the case of non-convergence of the bivariate model, since we have shown in Fig. S. 4.2 that the difference in mean predictive ability of all folds and only the converged folds is not critical. This difference can get higher as the number of non-converged folds increases. The percentage of converged folds in all studied material is shown in the supplementary (Table S. 4.3).

#### **4.4.5 Genomic correlation estimation**

Genomic correlations were estimated from the genetic variances and covariance derived from the ASReML bivariate model based on the full dataset of each location in both 2017 and 2018.

### **4.5 Results**

#### **Comparison of univariate GBLUP, bivariate GBLUP, bivariate ERRBLUP and bivariate sERRBLUP based on pruned set of SNPs and haplotype blocks in PH\_V4**

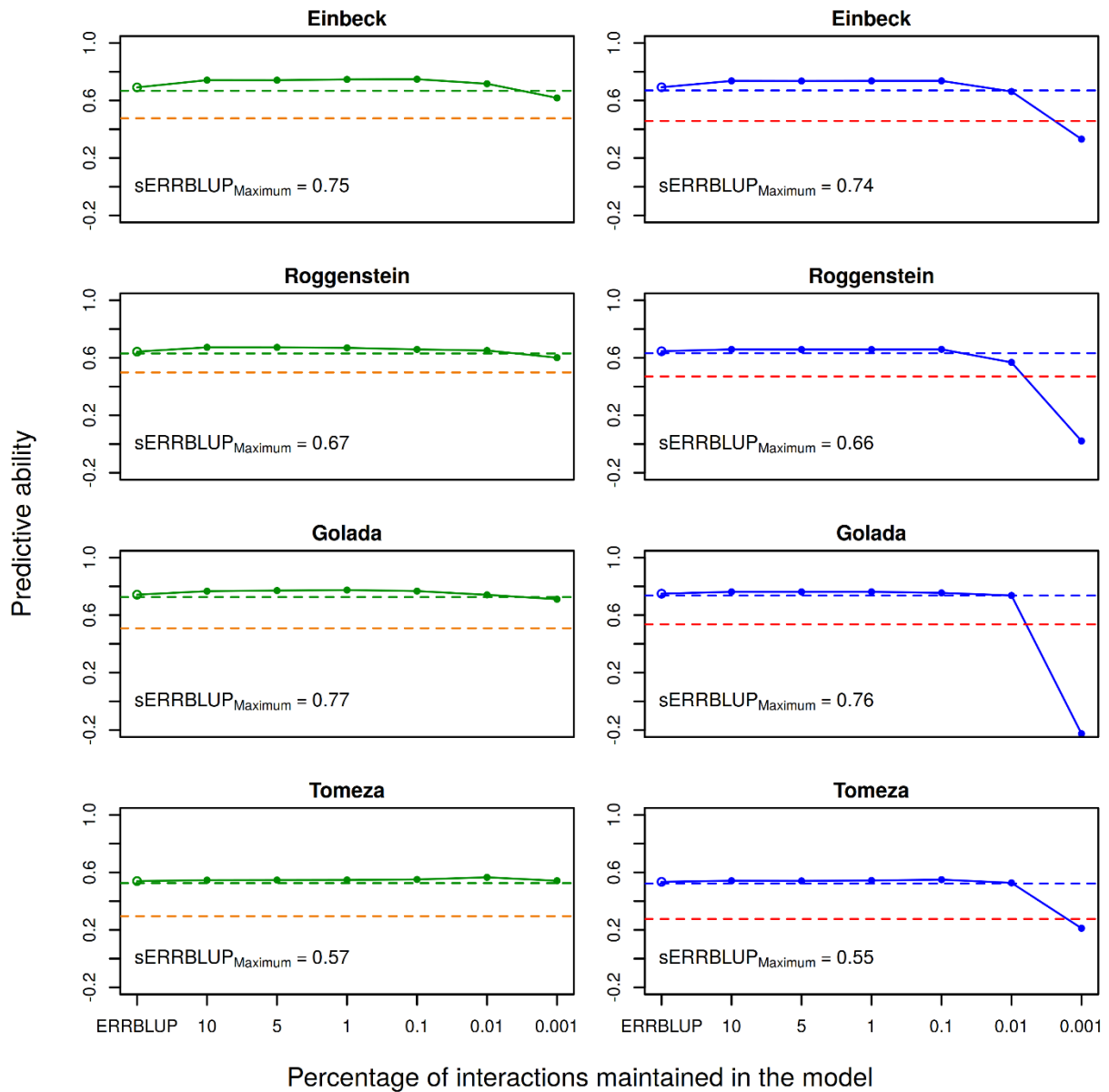
Our results confirm that bivariate models outperform the univariate models (Vojgani *et al.*, 2020) as illustrated by the comparison in predictive ability of bivariate GBLUP and univariate GBLUP for the trait PH-V4 in both landraces indicating the superiority of bivariate GBLUP to univariate GBLUP in most cases (see Fig. 4.1 and Fig. 4.2). Among the bivariate genomic prediction models,



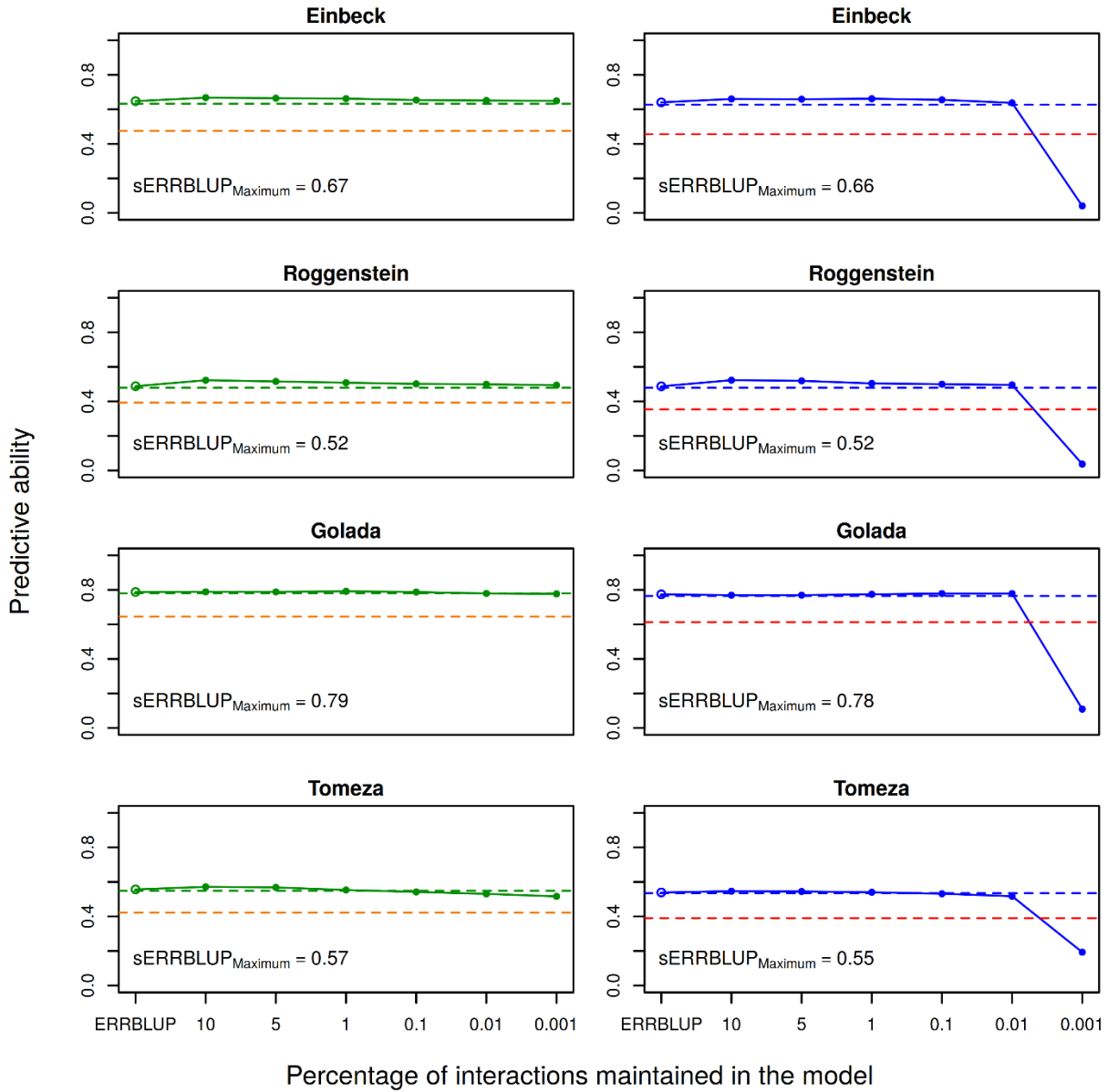
the predictive ability obtained from bivariate ERRBLUP is almost identical to bivariate GBLUP. This predictive ability increases in bivariate sERRBLUP and the highest gain in accuracy is generally obtained when the top 10 or 5 percent of pairwise SNP interactions are kept in the model. A too strict selection like using only the top 0.001 percent interactions, results in a decrease in predictive ability (see Fig. 4.1 and Fig. 4.2). The number of interactions maintained in the model for each proportion of interactions are tabulated in the supplementary (Table S. 4.4). Robustness of the predictive ability depending on the share of selected markers was higher in PE than KE. Moreover, Fig. 4.1 and Fig. 4.2 illustrate the comparison between the predictive abilities obtained from the respective genomic prediction models in KE and PE when utilizing pruned set of SNPs and haplotype blocks. It is shown that the GBLUP, ERRBLUP and sERRBLUP (for the optimum proportions of interactions) predictive abilities are almost identical in both pruned set of SNPs and haplotype blocks. It should be noted that the robustness of sERRBLUP when a very small proportions of interactions maintained in the model is higher when utilizing the pruned set of SNPs compared to haplotype blocks. This should not be surprising as the total number of interactions in the HaploBlocker panel is much smaller, thus, leading to a dataset with an extremely low number of explanatory variables (Table S. 4.4). Similar patterns are observed across a series of other traits for bivariate models which are shown in the supplementary (Fig. S. 4.3 - Fig. S. 4.16). Additionally, the predictive ability of univariate GBLUP by training the model on the average phenotypic values of both 2017 and 2018, when utilizing pruned set of SNPs, was evaluated for a series of phenotypic traits, which yielded quite similar predictive ability as obtained with univariate GBLUP within year 2018 or worse in some cases (Table S. 4.5 (KE) and Table S. 4.6 (PE)).

### **Correlation between prediction accuracy and the genomic correlation in bivariate models**

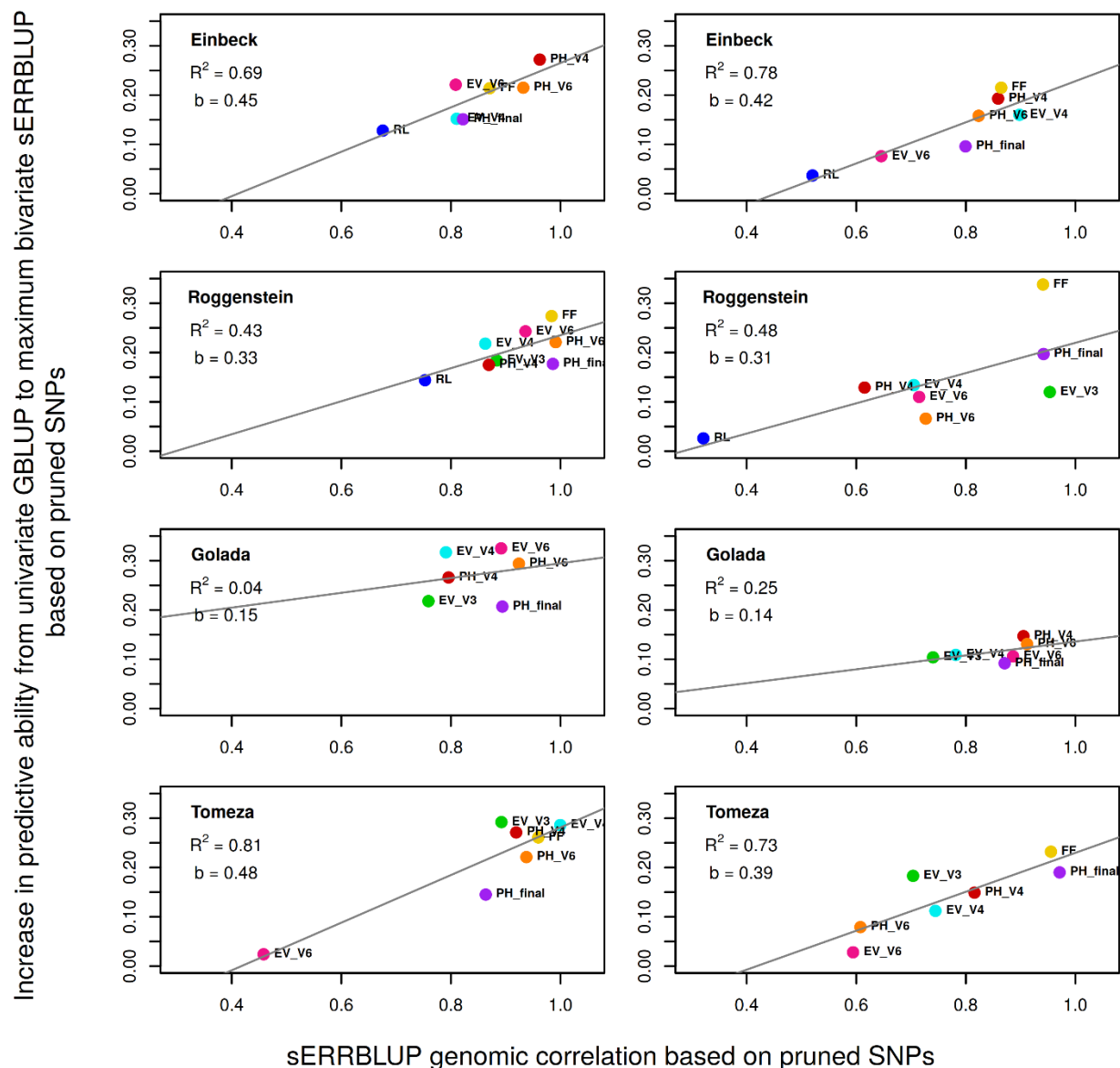
The absolute gain in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP, when utilizing pruned set of SNPs, was regressed on the respective sERRBLUP genomic correlation between the two respective environments and across the series of studied traits (Fig. 4.3). Regression coefficients range between 0.14 and 0.48 and thus show a clear association between the absolute gain in prediction accuracy and the genomic correlation between environments. When combining all traits and environments, this correlation is 0.65 (p-value = 0.00018) in KE and 0.69 (p-value = 4.393e-05) in PE. This correlation is also significant for most of the environments when utilizing haplotype blocks (Fig. S. 4.17).



**Fig. 4.1:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH-V4 in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP<sub>Maximum</sub> indicates the maximum predictive ability obtained from bivariate sERRBLUP.



**Fig. 4.2:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH-V4 in PE based on pruned sets of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.



**Fig. 4.3:** Regression of the absolute increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP on the respective sERRBLUP genomic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits. In each panel, the overall linear regression line (gray solid line) with the regression coefficient (*b*) and R-squared (*R*<sup>2</sup>) are shown.

### Interplay of GBLUP and sERRBLUP prediction accuracy and genomic correlations

The genomic correlations across years estimated with GBLUP and sERRBLUP based on pruned set of SNPs for the trait PH\_V4 are illustrated in Table 4.3, indicating that the proportion of interactions in bivariate sERRBLUP which maximized the predictive ability are not necessarily linked to the highest genomic correlation. In contrast, the best sERRBLUP for trait PH\_V4 is

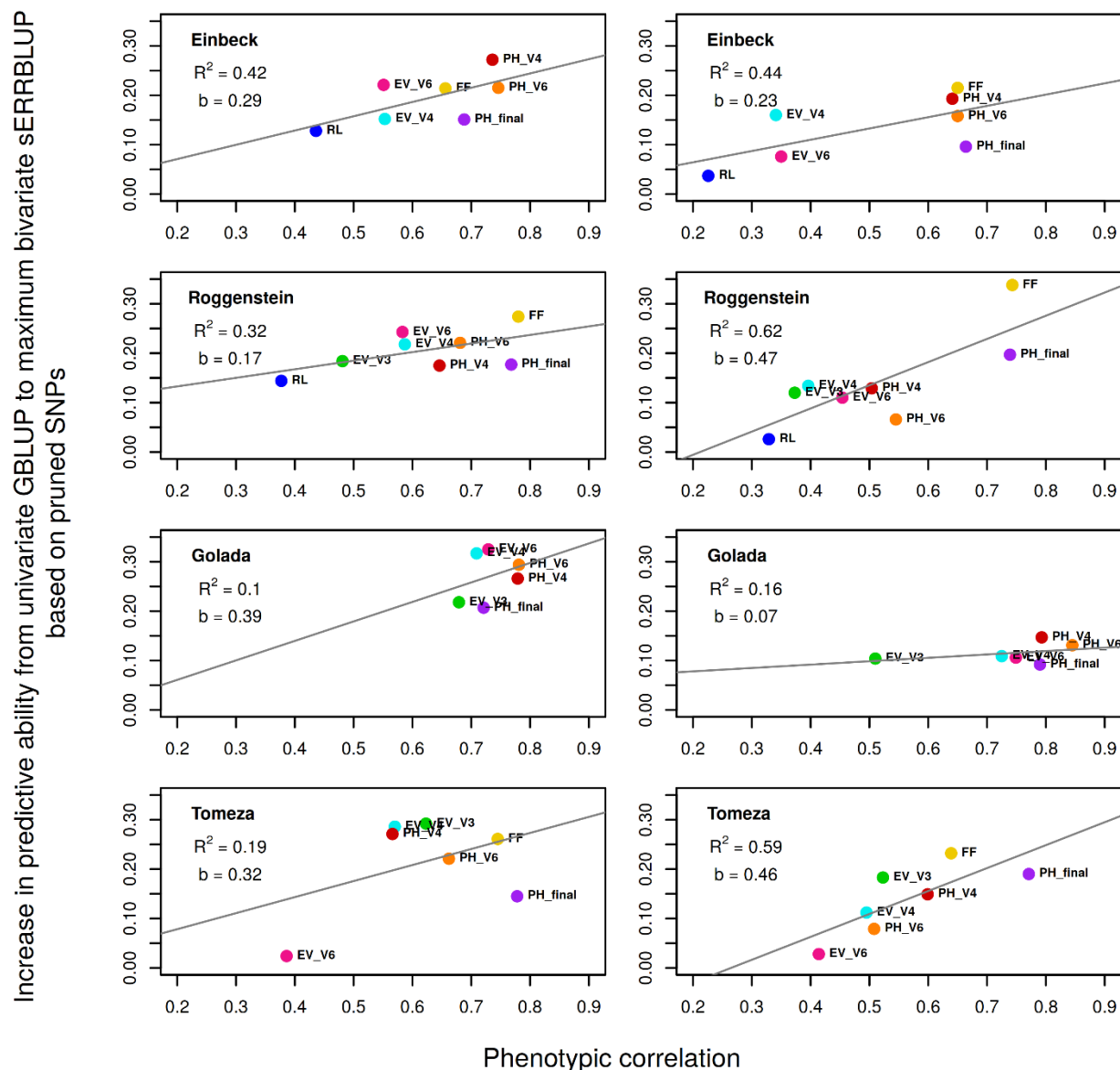
linked to the lowest genomic correlation in most cases. However, this is not the general pattern observed for a series of other traits and the best sERRBLUP for some traits and environments combinations are linked to the highest genomic correlation (Table S. 4.7 - Table S. 4.13). In fact, there is a significant correlation between the absolute increase in predictive ability from bivariate GBLUP to maximum bivariate sERRBLUP and the difference between genetic correlations estimated with GBLUP and maximum sERRBLUP in both KE and PE when utilizing pruned set of SNPs (Fig. S. 4.18).

**Table 4.3:** Genomic correlation between 2017 and 2018 in each environment for trait PH\_V4 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.

Bivariate Models	EIN	ROG	GOL	TOM
GBLUP	0.945 / 0.898	0.940 / 0.658	0.942 / 0.969	0.954 / 0.923
sERRBLUP top 10%	0.955 / <b>0.859*</b>	<b>0.869*</b> / <b>0.615*</b>	0.835 / 0.895	0.929 / <b>0.816*</b>
sERRBLUP top 5%	0.958 / 0.868	0.850 / 0.631	0.797 / 0.888	0.912 / 0.826
sERRBLUP top 1%	<b>0.949*</b> / 0.895	0.848 / 0.820	<b>0.796*</b> / <b>0.905*</b>	0.918 / 0.863
sERRBLUP top 0.1%	0.962 / 0.966	0.917 / 0.922	0.884 / 0.948	0.929 / 0.959
sERRBLUP top 0.01%	0.963 / 0.980	0.951 / 0.985	0.911 / 0.983	<b>0.919*</b> / 0.987
sERRBLUP top 0.001%	0.997 / 0.976	0.963 / 0.970	0.908 / 0.973	0.933 / 0.968

### Correlation between prediction accuracy and the phenotypic correlation in bivariate models

There might be some tendency that including phenotypes of the previous year into prediction becomes more efficient when the phenotypic correlation between years is high. In this context, the correlation between the absolute gain in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP and the phenotypic correlation among the years (see Table S. 4.14) over all studied traits in all four environments and in both landraces was studied. Fig. 4.4 demonstrates that the maximum correlation between the absolute gain in the respective predictive ability based on the pruned set of SNPs and the phenotypic correlation is obtained in EIN for KE (0.42) and in ROG for PE (0.62). Across all studied traits and environments, there is a significant correlation of 0.55 in KE (p-value= 0.003) and 0.50 in PE (p-value= 0.007). This correlation is also significant in most of the environments when utilizing haplotype blocks (Fig. S. 4.19).

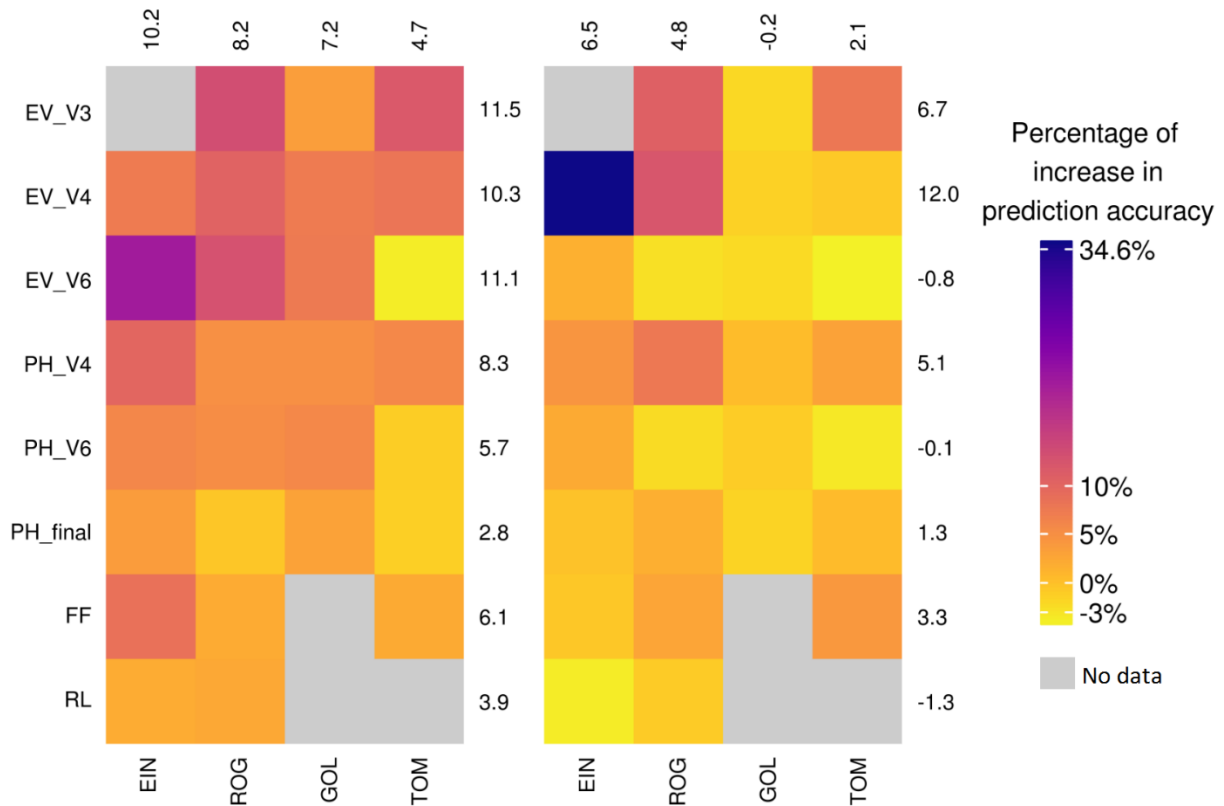


**Fig. 4.4:** Regression of the absolute increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP on the phenotypic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits. In each panel, the overall linear regression line (gray solid line) with the regression coefficient ( $b$ ) and R-squared ( $R^2$ ) are shown.

### Relative increase in prediction accuracy across all traits in all environments and landraces with bivariate models

Overall, the percentage of relative increase in prediction accuracy from the bivariate GBLUP to the maximum bivariate sERRBLUP based on pruned set of SNPs in both landraces are illustrated in Fig. 4.5 with the average increase of 7.61 percent in KE and 3.47 percent in PE over all studied

traits. Among all traits, the maximum relative increase in prediction accuracy for KE is 22.63 percent which was obtained in EV\_V6 in EIN, and for PE is 34.59 percent which was obtained in EV\_V4 in EIN. However, Fig. 4.5 shows some slight decreases in prediction accuracy from bivariate GBLUP to maximum bivariate sERRBLUP for some combinations of traits and environment in both landraces. This is more often observed in PE than KE, where the maximum decrease was found in EV\_V6 in TOM for both PE (-3.198 percent) and KE (-2.795 percent). Overall, the average relative increase from bivariate GBLUP to maximum bivariate sERRBLUP was over 3 percent in most cases. The absolute increase in prediction accuracy is also illustrated in the supplementary (Fig. S. 4.20) indicating the average increase of 0.046 in KE and 0.015 in PE over all combinations of traits and environments. In addition, the absolute increase in prediction accuracy from the bivariate GBLUP to the maximum bivariate sERRBLUP based on haplotype blocks is shown in the supplementary indicating the average absolute increase of 0.034 in KE and 0.013 in PE (Fig. S. 4.21). Overall, the increase in prediction accuracy from bivariate GBLUP to maximum bivariate sERRBLUP is significantly higher in KE than PE in both cases of utilizing pruned set of SNPs and haplotype blocks (Fig. S. 4.22).



**Fig. 4.5:** Percentage of change in prediction accuracy from bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP based on pruned set of SNPs in KE (left side plot) and in PE (right side plot). The average percentage of change in prediction accuracy for each trait and environment is displayed in all rows and columns, respectively.

## 4.6 Discussion

In this study, bivariate ERRBLUP as a full epistasis model incorporating all pairwise SNP interactions is almost identical to bivariate GBLUP. This was expected, since ERRBLUP incorporates a high number of interactions by which a large number of unimportant variables are introduced into the model (Martini *et al.*, 2016), thus introducing potential ‘noise’ which can prevent gains in predictive ability. In contrast, bivariate sERRBLUP substantially increases the predictive ability compared to bivariate GBLUP which is only caused by inclusion of relevant pairwise SNP interactions. Note that all bivariate models substantially outperformed univariate GBLUP, as phenotypic data of the respective environment in the previous year was used. ERRBLUP and sERRBLUP models have shown to display similar behaviors in the context of univariate statistical setting in the maize dataset for prediction across environments (Vojgani *et al.*, 2020). ERRBLUP, which had been introduced as categorical epistasis (CE) model by Martini *et al.* (2017), performs as good as the best EG-BLUP which is EG-BLUP with symmetric marker coding. Similarly, selection of optimum proportions of interactions in EG-BLUP has shown to increase the predictive ability compared to the EG-BLUP which includes all pairwise SNP interactions (Martini *et al.*, 2016) for a wheat dataset (Pérez and de los Campos, 2014). Consequently, ERRBLUP and sERRBLUP leads to higher predictive ability than EG-BLUP and reduced EG-BLUP with non-symmetric coded markers, respectively. This was shown by Vojgani *et al.* (2021) in the wheat dataset (Pérez and de los Campos, 2014).

Furthermore, in this study we have found that GBLUP, ERRBLUP and maximum sERRBLUP predictive abilities when utilizing haplotype blocks are very similar to the respective models’ predictive abilities when utilizing pruned sets of SNPs. This finding is of high relevance in practice, since it helps to overcome the high computational load of epistatic models. The required computational time for sERRBLUP based on 3’330 haplotype blocks indicating 5’546’115 interactions was 9 minutes out of which 4 minutes were needed to estimate the pairwise SNP interaction effect variances and 5 minutes were needed to generate the sERRBLUP relationship matrix for a selected proportion of interactions by utilizing the R-package miraculix with 15 cores on a server cluster with Intel E5-2650 (2X12 core 2.2GHz) processors in the released EpiGP R-package (Vojgani *et al.*, 2019). As the computing time is increasing approximately quadratically in the number of included markers, the computing time for the respective SNP-based model with 30’212 SNPs took 90 times as long (Vojgani *et al.*, 2020), while it resulted in similar predictive abilities with the absolute difference being less than 0.01 in most cases across all traits in all environments and both landraces (Fig. S. 4.23). Although this difference in predictive abilities is statistically significant based on paired t-test, it is not of practical relevance.

Although, sEERBLUP is a method that is using multiple environments, it is not a GxE model (de Leon *et al.*, 2016) in the traditional sense. While GxE models typically assign effect to combination of specific genotypes depending on the environment, the second environment in sEERBLUP is “only” used to detect which markers affect a given trait and use the information to put more focus on these marker in the actual prediction step. The estimation of marker effects itself is then executed only based on the environment itself (or in the case of the bivariate model with some contributions from the second environment but still not in the sense of a traditional GxE model).



As a different model is used for each environment, sEERBLUP will of course still assign different marker effects in different environments. As the set of selected marker interactions will however be different between different models, a direct comparison of the effects assigned to specific marker interactions is not statistically sound. Similar to most GxE models (Shin and Lee, 2020), the computational load of sEERBLUP using markers is extremely high. However, our suggested use of haplotype blocks massively reduced this problem, while the predictive ability is almost as good as sERRBLUP based on pruned set of SNPs.

It was shown that multivariate GBLUP is superior in predictive ability compared to univariate GBLUP with of medium ( $\sim 0.6$ ) to high ( $\sim 0.9$ ) genomic correlations, and that low genomic correlations results in no increase in multivariate GBLUP compared to univariate GBLUP (Covarrubias-Pazaran *et al.*, 2018). Calus *et al.* (2011) also found an increase of 3 to 14 percent in predictive ability of multi-trait SNP-based models in a simulation study when genetic correlations ranged from 0.25 to 0.75. In our study, we also found a significant correlation between the absolute gain in prediction accuracy from univariate GBLUP to maximum bivariate sERRBLUP and the respective genomic correlation based on both pruned sets of SNPs ( $r_{KE} = 0.65$ ,  $r_{PE} = 0.69$ ) and haplotype blocks ( $r_{KE} = 0.47$ ,  $r_{PE} = 0.49$ ) across all traits and environments combinations.

Moreover, Martini *et al.* (2016) showed that the predictive ability in one environment can be increased by variable selection in the other environment under the assumption of positive phenotypic correlation between environments. It was shown in a wheat dataset (Pérez and de los Campos, 2014), where environments 2 and 3 had the highest phenotypic correlation (0.661), that the predictive ability for phenotype prediction in environment 2 was maximized by variable selection in environment 3 and vice versa (Martini *et al.*, 2016). Therefore, the increase in prediction accuracy is expected to be influenced by the phenotypic correlations between the environments or between the years in the same environment in bivariate models. In our study, although 2017 and 2018 were climatically quite different, since 2018 suffered from a major heat stress compared to 2017 (Table 4.1), we see a significant correlation between the absolute gain in predictive ability from univariate GBLUP to maximum predictive ability of bivariate sERRBLUP and the phenotypic correlation between years in each environment based on both pruned sets of SNPs ( $r_{KE} = 0.55$ ,  $r_{PE} = 0.50$ ) and haplotype blocks ( $r_{KE} = 0.55$ ,  $r_{PE} = 0.56$ ).

In addition to the genomic and phenotypic correlations between the years, the trait heritability is another factor which is expected to be influential for such an increase in bivariate sERRBLUP predictive ability as well. Therefore, the traits with lower heritability are expected to obtain less gain in sERRBLUP predictive ability than the traits with higher heritability. This was confirmed in our study, as traits with low heritability (e.g. 0.59 for RL in PE) showed only a small increase in prediction accuracy from univariate GBLUP to maximum bivariate sERRBLUP. However, not all traits with higher heritabilities did necessarily show a higher gain in predictive ability for all traits. It should be noted that the trait heritabilities were calculated on an entry-mean basis (Hallauer *et al.*, 2010) within each KE and PE landraces by Hölker *et al.* (2019) over all four environments in both years 2017 and 2018 jointly. The trait heritabilities obtained only from 2017 are significantly higher than the trait heritabilities obtained only from 2018 in both KE and PE based on a paired t-test (Table S. 4.2). This also results in an increase in predictive ability from

univariate GBLUP to maximum bivariate sERRBLUP in KE and PE, since multi-trait models have the potential of increasing the predictive ability when traits with low heritability are joined with traits with higher heritability, given they are genomically correlated (Thompson and Meyer, 1986).

It should be noted that the increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP is caused by both borrowing information across years and capitalizing on epistasis, while the increase in predictive ability from bivariate GBLUP to maximum bivariate sERRBLUP is caused by accounting for epistasis alone. Overall, the traits behave differently among different environments and landraces due to their genomic correlations, phenotypic correlations and heritabilities. To shed light on this, the maximum increase in prediction accuracy from bivariate GBLUP to bivariate sERRBLUP based on pruned set of SNPs in KE was observed for the trait EV\_V6 (0.112) in EIN where the corresponding sERRBLUP genomic correlation (0.809) is higher than the GBLUP genomic correlation (0.768). This trait has a high heritability (0.90) and high phenotypic correlation (0.551) as well. In contrast, the respective prediction accuracy decreases (-0.018) for EV\_V6 in TOM for KE indicating the lower sERRBLUP genomic correlation (0.458) than GBLUP genomic correlation (0.703) and the particularly low phenotypic correlation (0.383). It should be noted that the phenotypic correlation does not play a major role for the increase in prediction accuracy from bivariate GBLUP to bivariate sERRBLUP, since both models are bivariate and benefit from the same phenotypic correlations. Therefore, EV\_V6 obtaining the maximum and minimum increase in the respective prediction accuracy for KE indicates the significant role of genomic correlation among the possible causes. In general, bivariate sERRBLUP improves the prediction accuracy compared to bivariate GBLUP more in KE than PE which is potentially due to significantly higher sERRBLUP genomic correlation and heritability in KE compared to PE, based on paired t-test.

Overall, our results indicate that incorporating a suitable subset of epistatic interactions besides utilizing information across years can substantially increase the predictive ability. The amount of this increase is affected by the genomic and phenotypic correlations between the years and the heritability of the phenotypic trait. Moreover, utilizing haplotype blocks instead of pruned sets of SNPs in epistasis model is proposed, since the obtained predictive abilities of the best epistasis model are quite similar, while the required computational time when utilizing haplotype blocks is significantly lower than the required computational time when utilizing pruned set of SNPs. Therefore, this computationally efficient approach is potentially beneficial for genomic prediction of phenotypes under the assumption of sufficient genomic and phenotypic correlation between years for highly heritable traits. This may allow to reduce the number of lines which have to be phenotyped over several years and thus reduce phenotyping costs which and thus be of high interest in practical plant breeding.

## 4.7 Supplementary material

### 4.7.1 Supplementary Tables

**Table S. 4.1:** The mean, minimum, maximum and standard deviation of BLUEs of phenotypic traits in each location for KE (blue numbers) and PE (red numbers) in 2017 and 2018.

Trait	Location/Year	Mean	Minimum	Maximum	Standard deviation	
EV_V3	ROG/2017	5.35\5.84	1.71\2.90	7.90\7.92	0.95\0.75	
	ROG/2018	5.68\5.91	1.21\3.46	7.91\8.08	1.01\0.89	
	GOL/2017	6.31\6.67	4.07\5.49	8.49\7.98	0.69\0.51	
	GOL/2018	4.26\4.62	0.32\1.69	7.63\7.99	1.20\0.97	
	TOM/2017	5.51\6.15	1.93\3.84	7.34\8.45	0.99\0.67	
	TOM/2018	4.84\5.67	0.59\1.38	8.67\8.93	1.43\1.13	
	EIN/2017	4.24\4.82	0.94\1.52	7.07\7.46	1.11\0.98	
	EIN/2018	4.56\4.61	0.98\1.99	7.01\6.99	1.05\0.78	
	EV_V4	ROG/2017	5.44\5.85	2.65\2.88	7.86\7.94	0.92\0.78
		ROG/2018	5.61\5.84	1.39\2.89	8.41\9.07	1.22\1.15
GOL/2017		5.71\5.98	3.37\3.91	7.89\7.89	0.81\0.83	
GOL/2018		5.16\5.29	1.50\1.63	8.36\8.44	1.26\1.36	
TOM/2017		5.26\5.75	2.59\3.92	6.89\7.35	0.83\0.61	
TOM/2018		5.01\5.27	0.96\2.08	8.56\8.07	1.49\1.11	
EIN/2017		5.03\5.54	0.97\1.51	8.05\8.39	1.24\1.06	
EIN/2018		4.73\4.73	1.07\2.58	6.95\6.08	0.78\0.56	
EV_V6		ROG/2017	5.55\5.91	1.02\2.52	8.07\7.76	0.95\0.77
		ROG/2018	6.14\6.42	2.21\3.36	8.81\9.68	1.29\1.20
	GOL/2017	6.24\6.24	3.90\3.81	8.45\7.94	0.85\0.85	
	GOL/2018	5.12\4.77	1.21\1.17	8.23\7.51	1.29\1.26	
	TOM/2017	5.58\5.86	2.96\3.90	7.66\7.91	0.92\0.68	
	TOM/2018	6.30\5.43	2.44\1.07	9.60\9.08	1.36\1.26	
	EIN/2017	34.49\38.73	6.90\20.43	53.14\57.94	7.24\6.17	
	EIN/2018	32.54\35.23	8.48\19.60	49.63\50.29	5.65\5.24	
	PH_V4	ROG/2017	25.50\28.10	9.23\13.63	42.29\41.54	4.60\4.53
		ROG/2018	29.08\31.75	11.11\17.45	43.70\45.04	4.79\4.92

Trait	Location/Year	Mean	Minimum	Maximum	Standard deviation
	<b>GOL/2017</b>	62.88\68.98	34.30\38.39	88.24\95.30	9.79\10.96
	<b>GOL/2018</b>	60.37\64.49	23.27\16.14	89.24\93.20	12.54\14.32
	<b>TOM/2017</b>	41.60\47.45	11.98\25.37	63.89\72.12	8.71\8.27
	<b>TOM/2018</b>	60.32\66.49	28.43\47.45	84.24\88.27	9.59\8.03
	<b>EIN/2017</b>	62.40\69.36	21.41\36.53	95.54\98.80	11.89\9.62
	<b>EIN/2018</b>	78.81\85.14	21.90\52.02	105.05\115.61	10.53\10.07
<b>PH_V6</b>	<b>ROG/2017</b>	61.46\68.91	32.17\30.35	89.74\94.77	9.34\9.52
	<b>ROG/2018</b>	82.64\90.89	41.48\57.90	118.69\123.27	11.00\11.17
	<b>GOL/2017</b>	94.21\98.30	37.28\54.75	127.54\130.51	15.05\15.29
	<b>GOL/2018</b>	101.90\104.82	53.69\50.37	137.67\146.02	15.42\18.24
	<b>TOM/2017</b>	83.86\92.35	48.46\57.81	119.07\124.98	14.41\12.79
	<b>TOM/2018</b>	120.46\120.57	68.48\58.96	173.66\169.71	19.56\18.63
	<b>EIN/2017</b>	159.18\141.35	100.84\69.01	228.96\211.14	21.57\21.10
	<b>EIN/2018</b>	114.90\93.35	82.28\49.97	172.12\136.25	16.46\16.27
<b>PH_final</b>	<b>ROG/2017</b>	137.04\122.25	74.25\63.56	211.14\201.92	22.32\20.56
	<b>ROG/2018</b>	163.71\142.70	103.82\70.16	249.35\208.81	25.52\23.66
	<b>GOL/2017</b>	115.68\102.69	49.27\30.21	167.58\149.14	21.73\23.59
	<b>GOL/2018</b>	129.94\117.16	35.41\35.76	186.09\173.10	26.35\27.58
	<b>TOM/2017</b>	157.99\144.61	81.92\79.28	245.00\195.36	24.82\18.95
	<b>TOM/2018</b>	184.54\169.57	115.10\118.15	265.02\248.43	26.34\22.77
<b>FF</b>	<b>EIN/2017</b>	82.55\81.78	70.36\68.86	102.02\101.50	5.23\5.17
	<b>EIN/2018</b>	78.80\79.59	63.37\68.12	94.35\93.96	5.44\5.40
	<b>ROG/2017</b>	73.06\71.91	62.45\59.10	91.22\88.03	4.82\4.47
	<b>ROG/2018</b>	79.16\79.05	66.74\67.74	100.14\92.87	4.72\4.41
	<b>TOM/2017</b>	76.88\74.16	63.93\62.13	93.28\92.17	5.58\4.64
	<b>TOM/2018</b>	70.31\68.76	62.22\60.14	83.64\90.06	4.11\3.73
<b>RL</b>	<b>EIN/2017</b>	3.48\2.23	0.63\0.76	9.21\8.08	2.29\1.54
	<b>EIN/2018</b>	1.58\1.25	0.73\0.32	8.52\4.69	1.10\0.59
	<b>ROG/2017</b>	2.39\1.50	0.96\0.95	9.01\8.50	2.21\1.13
	<b>ROG/2018</b>	1.27\1.17	0.95\0.95	7.01\3.52	0.61\0.41

**Table S. 4.2:** The traits heritabilities in 2017, 2018 and both years jointly in KE (blue numbers) and PE (red numbers).

Traits	2017	2018	Both 2017 and 2018
EV_V3	0.91 / 0.85	0.79 / 0.67	0.92 / 0.86
EV_V4	0.90 / 0.82	0.83 / 0.71	0.91 / 0.84
EV_V6	0.88 / 0.84	0.75 / 0.68	0.90 / 0.85
PH_V4	0.89 / 0.82	0.81 / 0.72	0.92 / 0.87
PH_V6	0.90 / 0.88	0.85 / 0.81	0.93 / 0.91
PH_final	0.91 / 0.93	0.83 / 0.85	0.94 / 0.94
FF	0.90 / 0.92	0.86 / 0.83	0.94 / 0.93
RL	0.78 / 0.55	0.49 / 0.29	0.80 / 0.59

**Table S. 4.3:** The percentage of bivariate models convergence in 5-fold cross validation with 5 replicates based on pruned set of SNPs for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The stars represent the non-convergence of pre estimated variance components based on the full set.

Traits	Predicted Environments	GBLUP	ERRBLUP	sERRBLUP Top 10%	sERRBLUP Top 5%	sERRBLUP Top 1%	sERRBLUP Top 0.1%	sERRBLUP Top 0.01%	sERRBLUP Top 0.001%
EV_V3	ROG	92%	88%	96%	88%	52%*	68%*	60%*	16%*
	GOL	100%	100%	0%*	0%*	100%	72%	84% / 44%	1 / 60%
	TOM	100%	100%	92%	96%	100%	96% / 96%	92% / 64%	40%* / 28%*
EV_V4	EIN	100%	100%	100%	100%	100%	80% / 100%*	76%	100%
	ROG	96%	96%	100%	100%	96%	100%	96% / 96%	76% / 96%
	GOL	100%	100%	100%	100%	100%	100%*	100%	68%
	TOM	40%* / 96%	52% / 96%	88%	100%	96%	100%*	100%	96%
EV_V6	EIN	100%	100%	100%	100%	100%	96%	96%	100%* / 96%
	ROG	100%	100%	100%	100%	100%	100%	56%* / 92%	0%*
	GOL	100%	100%	0%* / 0%*	0% / 0%*	0%*	16%* / 0%*	0%*	100%
	TOM	100%	100%	100%	100%	100%	100%	100%	100%

Traits	Predicted Environments	GBLUP	ERRBLUP	sERRBLUP Top 10%	sERRBLUP Top 5%	sERRBLUP Top 1%	sERRBLUP Top 0.1%	sERRBLUP Top 0.01%	sERRBLUP Top 0.001%
PH_V4	EIN	100%	100%	100%	96%	100%	100%	96%	100%
	ROG	100%	100%	100%	100%	100%	84%	84%	92%
	GOL	100%	100%	0%*	0%*	0%*	12%*	100%	100%
	TOM	96% / 92%	96% / 96%	100%	100%	100%	92%	92%	96%
PH_V6	EIN	100%	100%	100%	100%	100%	100%	100%	100%
	ROG	100%	100%	100%	100%	100%	84%	92%	84%*
	GOL	96%	96%	96%	96%	100%	100%	96%	84%*
	TOM	100%	100%	100%	100%	100%	100%	96%	100%
PH_final	EIN	100%	100%	100%	100%	96%	84%	76%*	72%*
	ROG	100%	100%	0%* / 88%	0%* / 0%*	0%* / 0%*	0%* / 0%*	92% / 0%*	48%*
	GOL	100%	100%	100%	100%	96%	100%	100%	68%
	TOM	100%	100%	4%* / 96%	92%	20%* / 96%	4%* / 68%*	96%* / 52%*	76%
FF	EIN	100%	100%	96%	100%	64%	0%*	64%	92%
	ROG	100%	100%	96% / 0%*	0%*	96% / 0%*	4%* / 0%*	56%* / 0%*	24%*
	TOM	88% / 80%	64%* / 84%	96% / 80%	96% / 80%	92% / 76%	60% / 52%*	36%* / 32%*	76% / 72%
RL	EIN	100%	100%	100%	100%	100%	100%	100%	100%
	ROG	100%	100%	100%	100%	96%	92%	72%	84%

**Table S. 4.4:** The number of epistasis interactions maintained in the model based on haplotype blocks and pruned set of SNPs for the each proportions of interactions in KE and PE.

Proportions of interactions	Landrace	Pruned set of SNPs	Haplotype Blocks
<b>100%</b>	<b>KE</b>	323'533'203	4'416'392
	<b>PE</b>	456'397'578	5'546'115
<b>Top 10%</b>	<b>KE</b>	32'353'320	441'639
	<b>PE</b>	45'639'758	554'612
<b>Top 5%</b>	<b>KE</b>	16'176'660	220'820
	<b>PE</b>	22'819'879	277'306
<b>Top 1%</b>	<b>KE</b>	3'235'332	44'164
	<b>PE</b>	4'563'976	55'461
<b>Top 0.1%</b>	<b>KE</b>	323'533	4'416
	<b>PE</b>	456'398	5'546
<b>Top 0.01%</b>	<b>KE</b>	32'353	443
	<b>PE</b>	45'640	555
<b>Top 0.001%</b>	<b>KE</b>	3'235	44
	<b>PE</b>	4'564	55

**Table S. 4.5:** GBLUP predictive ability based on pruned set of SNPs for prediction in 2018 with training the model either on 2018 data or the average phenotypic values of 2017 and 2018 in each environment for series of phenotypic traits in KE.

Trait	Training set	EIN	ROG	GOL	TOM
<b>EV_V3</b>	2018	NA	0.335	0.435	0.291
	2017 and 2018 average	NA	0.338	0.425	0.311
<b>EV_V4</b>	2018	0.448	0.385	0.410	0.299
	2017 and 2018 average	0.429	0.386	0.387	0.316
<b>EV_V6</b>	2018	0.355	0.397	0.403	0.566
	2017 and 2018 average	0.314	0.406	0.410	0.532
<b>PH_V4</b>	2018	0.470	0.495	0.506	0.299
	2017 and 2018 average	0.483	0.488	0.518	0.294
<b>PH_V6</b>	2018	0.527	0.494	0.464	0.463
	2017 and 2018 average	0.523	0.458	0.470	0.471
<b>PH_final</b>	2018	0.554	0.597	0.513	0.656
	2017 and 2018 average	0.542	0.631	0.526	0.635
<b>FF</b>	2018	0.474	0.502	0.134	0.496
	2017 and 2018 average	0.458	0.551	-	0.506
<b>RL</b>	2018	0.391	0.252	-	-
	2017 and 2018 average	0.358	0.212	-	-

**Table S. 4.6:** GBLUP predictive ability based on pruned set of SNPs for prediction in 2018 with training the model either on 2018 data or the average phenotypic values of 2017 and 2018 in each environment for series of phenotypic traits in PE.

Trait	Training set	EIN	ROG	GOL	TOM
<b>EV_V3</b>	2018	NA	0.209	0.337	0.326
	2017 and 2018 average	NA	0.210	0.332	0.323
<b>EV_V4</b>	2018	0.185	0.261	0.625	0.363
	2017 and 2018 average	0.205	0.269	0.646	0.371
<b>EV_V6</b>	2018	0.323	0.362	0.627	0.583
	2017 and 2018 average	0.288	0.382	0.645	0.572
<b>PH_V4</b>	2018	0.464	0.510	0.641	0.437
	2017 and 2018 average	0.480	0.396	0.661	0.439
<b>PH_V6</b>	2018	0.547	0.529	0.690	0.515
	2017 and 2018 average	0.541	0.440	0.706	0.512
<b>PH_final</b>	2018	0.620	0.537	0.682	0.580
	2017 and 2018 average	0.547	0.370	0.699	0.572
<b>FF</b>	2018	0.466	0.456	0.306	0.401
	2017 and 2018 average	0.495	0.079	-	0.363
<b>RL</b>	2018	0.268	0.419	-	-
	2017 and 2018 average	0.235	0.114	-	-

**Table S. 4.7:** Genomic correlation between 2017 and 2018 in each environment for trait EV\_V3 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.

Bivariate Models	ROG	GOL	TOM
<b>GBLUP</b>	0.864 / 0.442	0.984 / 0.768	0.913 / 0.642
<b>sERRBLUP top 10%</b>	0.869 / 0.643	0.804 / <b>0.740*</b>	<b>0.892*</b> / <b>0.704*</b>
<b>sERRBLUP top 5%</b>	0.872 / 0.976	<b>0.759*</b> / 0.725	0.877 / 0.686
<b>sERRBLUP top 1%</b>	0.844 / 0.947	0.760 / 0.824	0.896 / 0.765
<b>sERRBLUP top 0.1%</b>	<b>0.884*</b> / <b>0.953*</b>	0.812 / 0.931	0.920 / 0.997
<b>sERRBLUP top 0.01%</b>	0.829 / 0.968	0.781 / 0.973	0.993 / 0.938
<b>sERRBLUP top 0.001%</b>	0.857 / 0.940	0.892 / 0.962	0.962 / 0.908



**Table S. 4.8:** Genomic correlation between 2017 and 2018 in each environment for trait EV\_V4 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.

Bivariate Models	EIN	ROG	GOL	TOM
GBLUP	0.843 / 0.525	0.897 / 0.592	0.940 / 0.998	0.960 / 0.835
sERRBLUP top 10%	<b>0.810*</b> / 0.713	<b>0.863*</b> / 0.618	0.872 / 0.815	0.972 / <b>0.744*</b>
sERRBLUP top 5%	0.807 / 0.732	0.861 / 0.632	0.830 / <b>0.782*</b>	0.997 / 0.751
sERRBLUP top 1%	0.865 / 0.769	0.889 / <b>0.705*</b>	<b>0.791*</b> / 0.824	0.941 / 0.733
sERRBLUP top 0.1%	0.966 / 0.888	0.936 / 0.766	0.874 / 0.946	0.934 / 0.784
sERRBLUP top 0.01%	0.979 / 0.908	0.947 / 0.780	0.883 / 0.999	0.979 / 0.806
sERRBLUP top 0.001%	0.974 / <b>0.898*</b>	0.966 / 0.758	0.861 / 0.932	<b>0.999*</b> / 0.834

**Table S. 4.9:** Genomic correlation between 2017 and 2018 in each environment for trait EV\_V6 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.

Bivariate Models	EIN	ROG	GOL	TOM
GBLUP	0.768 / 0.712	0.989 / 0.965	0.900 / 0.940	0.703 / 0.764
sERRBLUP top 10%	0.817 / <b>0.646*</b>	0.922 / <b>0.715*</b>	0.895 / <b>0.886*</b>	<b>0.458*</b> / <b>0.594*</b>
sERRBLUP top 5%	<b>0.809*</b> / 0.635	0.900 / 0.736	0.861 / 0.883	0.412 / 0.566
sERRBLUP top 1%	0.809 / 0.690	0.888 / 0.818	0.823 / 0.898	0.409 / 0.542
sERRBLUP top 0.1%	0.890 / 0.842	<b>0.936*</b> / 0.882	<b>0.892*</b> / 0.942	0.489 / 0.544
sERRBLUP top 0.01%	0.991 / 0.840	0.954 / 0.908	0.914 / 0.916	0.496 / 0.530
sERRBLUP top 0.001%	0.909 / 0.899	0.969 / 0.968	0.856 / 0.932	0.567 / 0.573

**Table S. 4.10:** Genomic correlation between 2017 and 2018 in each environment for trait PH\_V6 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.

Bivariate Models	EIN	ROG	GOL	TOM
GBLUP	0.942 / 0.880	0.952 / 1.000	0.937 / 0.952	0.994 / 0.758
sERRBLUP top 10%	<b>0.932*</b> / <b>0.823*</b>	0.909 / <b>0.727*</b>	0.943 / <b>0.912*</b>	0.949 / <b>0.608*</b>
sERRBLUP top 5%	0.928 / 0.801	0.909 / 0.749	0.935 / 0.908	<b>0.938*</b> / 0.568
sERRBLUP top 1%	0.910 / 0.823	0.951 / 0.826	0.887 / 0.934	0.936 / 0.501
sERRBLUP top 0.1%	0.934 / 0.915	<b>0.991*</b> / 0.934	<b>0.924*</b> / 0.957	0.955 / 0.964
sERRBLUP top 0.01%	0.902 / 0.874	0.959 / 0.963	0.969 / 0.948	0.968 / 0.526
sERRBLUP top 0.001%	0.918 / 0.848	0.984 / 0.955	0.919 / 0.941	0.993 / 0.713

**Table S. 4.11:** Genomic correlation between 2017 and 2018 in each environment for trait PH\_final for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.

Bivariate Models	EIN	ROG	GOL	TOM
GBLUP	0.831 / 0.851	0.984 / 0.935	0.958 / 0.943	0.985 / 0.981
sERRBLUP top 10%	<b>0.822*</b> / <b>0.799*</b>	<b>0.986*</b> / 0.910	<b>0.894*</b> / <b>0.871*</b>	<b>0.864*</b> / <b>0.971*</b>
sERRBLUP top 5%	0.820 / 0.738	0.987 / 0.900	0.861 / 0.863	0.816 / 0.965
sERRBLUP top 1%	0.881 / 0.835	0.994 / 0.924	0.822 / 0.867	0.808 / 0.981
sERRBLUP top 0.1%	0.986 / 0.934	0.999 / 0.959	0.989 / 0.956	0.841 / 0.981
sERRBLUP top 0.01%	0.962 / 0.922	0.999 / 0.967	0.982 / 0.994	0.873 / 0.984
sERRBLUP top 0.001%	0.965 / 0.879	0.992 / <b>0.942*</b>	0.988 / 0.998	0.935 / 0.987

**Table S. 4.12:** Genomic correlation between 2017 and 2018 in each environment for trait FF for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.

Bivariate Models	EIN	ROG	TOM
GBLUP	0.888 / 0.951	1.000 / 0.979	0.995 / 0.970
sERRBLUP top 10%	<b>0.871*</b> / 0.882	0.987 / 0.969	<b>0.959*</b> / <b>0.955*</b>
sERRBLUP top 5%	0.859 / <b>0.865*</b>	<b>0.984*</b> / 0.955	0.955 / 0.944
sERRBLUP top 1%	0.852 / 0.855	0.983 / <b>0.941*</b>	0.995 / 0.993
sERRBLUP top 0.1%	0.854 / 0.861	1.000 / 0.955	0.987 / 0.981
sERRBLUP top 0.01%	0.847 / 0.909	0.997 / 0.970	0.957 / 0.977
sERRBLUP top 0.001%	0.814 / 0.898	0.992 / 0.961	0.992 / 0.984

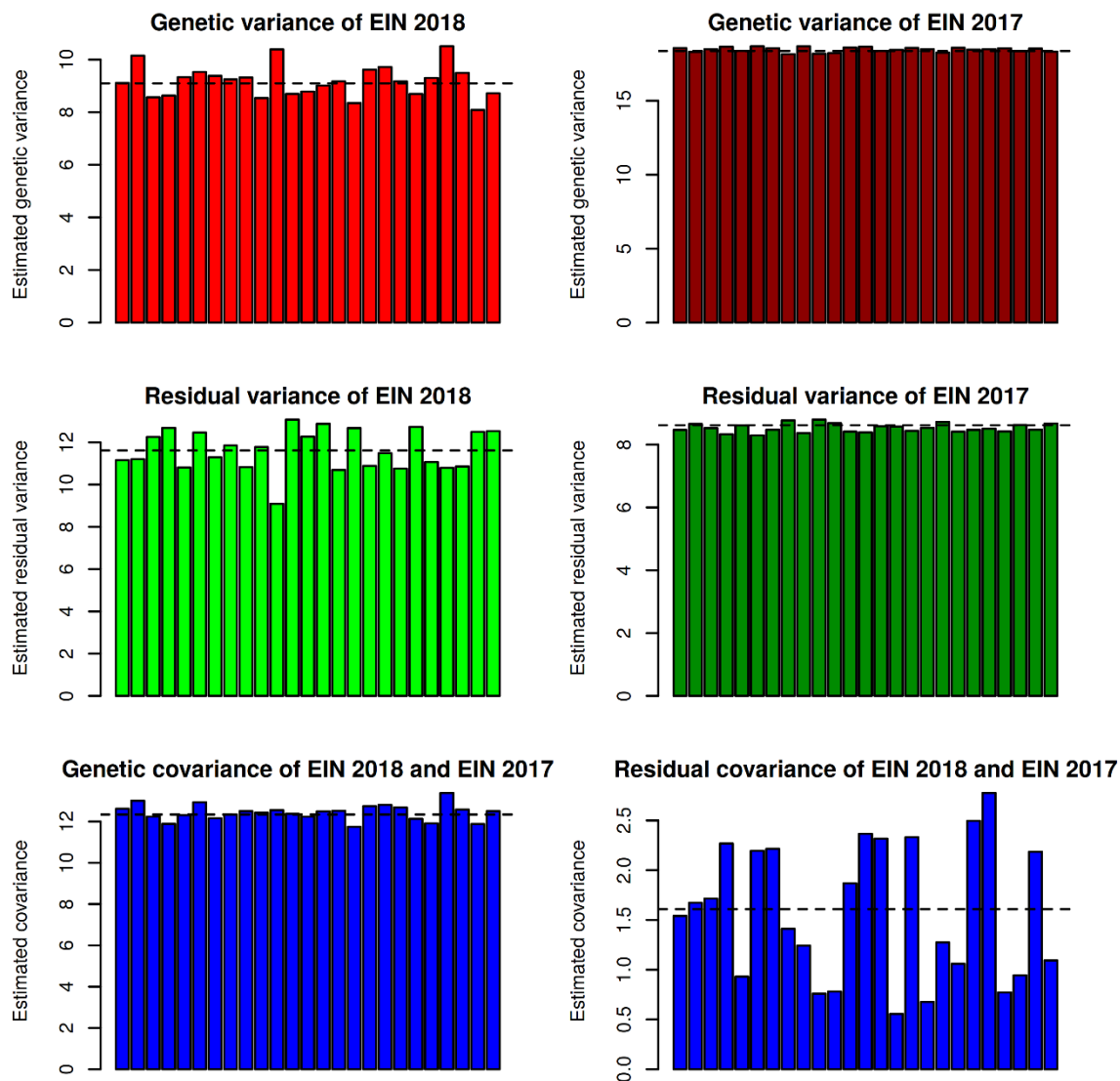
**Table S. 4.13:** Genomic correlation between 2017 and 2018 in each environment for trait RL for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.

Bivariate Models	EIN	ROG
GBLUP	0.662 / 0.818	0.751 / 0.404
sERRBLUP top 10%	0.672 / 0.604	<b>0.753*</b> / <b>0.321*</b>
sERRBLUP top 5%	0.644 / <b>0.520*</b>	0.741 / 0.325
sERRBLUP top 1%	0.603 / 0.546	0.748 / 0.338
sERRBLUP top 0.1%	<b>0.676*</b> / 0.719	0.887 / 0.412
sERRBLUP top 0.01%	0.727 / 0.572	0.973 / 0.380
sERRBLUP top 0.001%	0.712 / 0.531	0.983 / 0.352

**Table S. 4.14:** Phenotypic correlation between 2017 and 2018 in each environment for KE and PE.

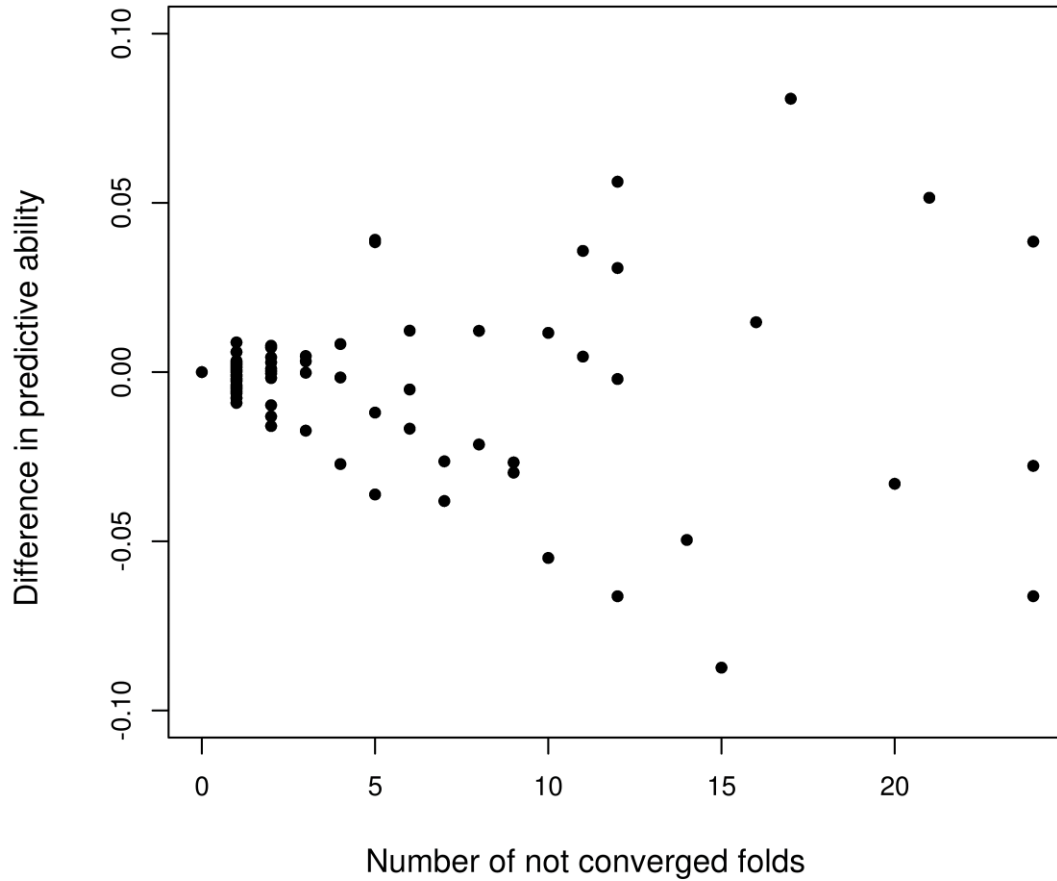
<b>Trait</b>	<b>Landrace</b>	<b>EIN</b>	<b>ROG</b>	<b>GOL</b>	<b>TOM</b>
<b>EV_V3</b>	<b>KE</b>	-	0.481	0.679	0.623
	<b>PE</b>	-	0.373	0.510	0.523
<b>EV_V4</b>	<b>KE</b>	0.553	0.587	0.709	0.570
	<b>PE</b>	0.341	0.396	0.725	0.495
<b>EV_V6</b>	<b>KE</b>	0.551	0.583	0.729	0.383
	<b>PE</b>	0.350	0.454	0.749	0.414
<b>PH_V4</b>	<b>KE</b>	0.736	0.646	0.779	0.566
	<b>PE</b>	0.641	0.504	0.793	0.599
<b>PH_V6</b>	<b>KE</b>	0.746	0.681	0.781	0.662
	<b>PE</b>	0.650	0.545	0.845	0.508
<b>PH_final</b>	<b>KE</b>	0.688	0.768	0.721	0.778
	<b>PE</b>	0.668	0.739	0.790	0.771
<b>FF</b>	<b>KE</b>	0.656	0.780	-	0.745
	<b>PE</b>	0.650	0.743	-	0.639
<b>RL</b>	<b>KE</b>	-	0.436	-	0.377
	<b>PE</b>	-	0.226	-	0.329

## 4.7.2 Supplementary Figures

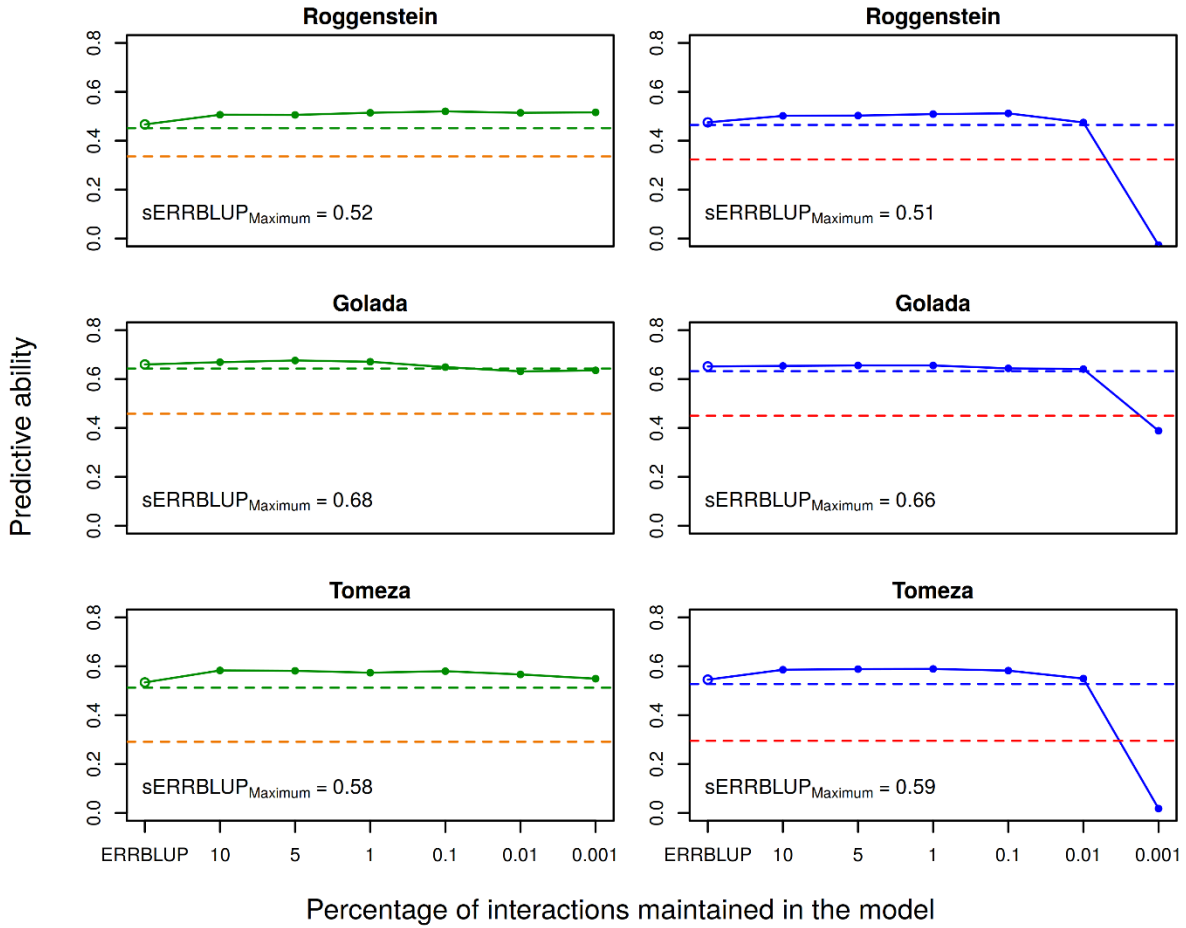


5-fold cross validation with 5 replicates

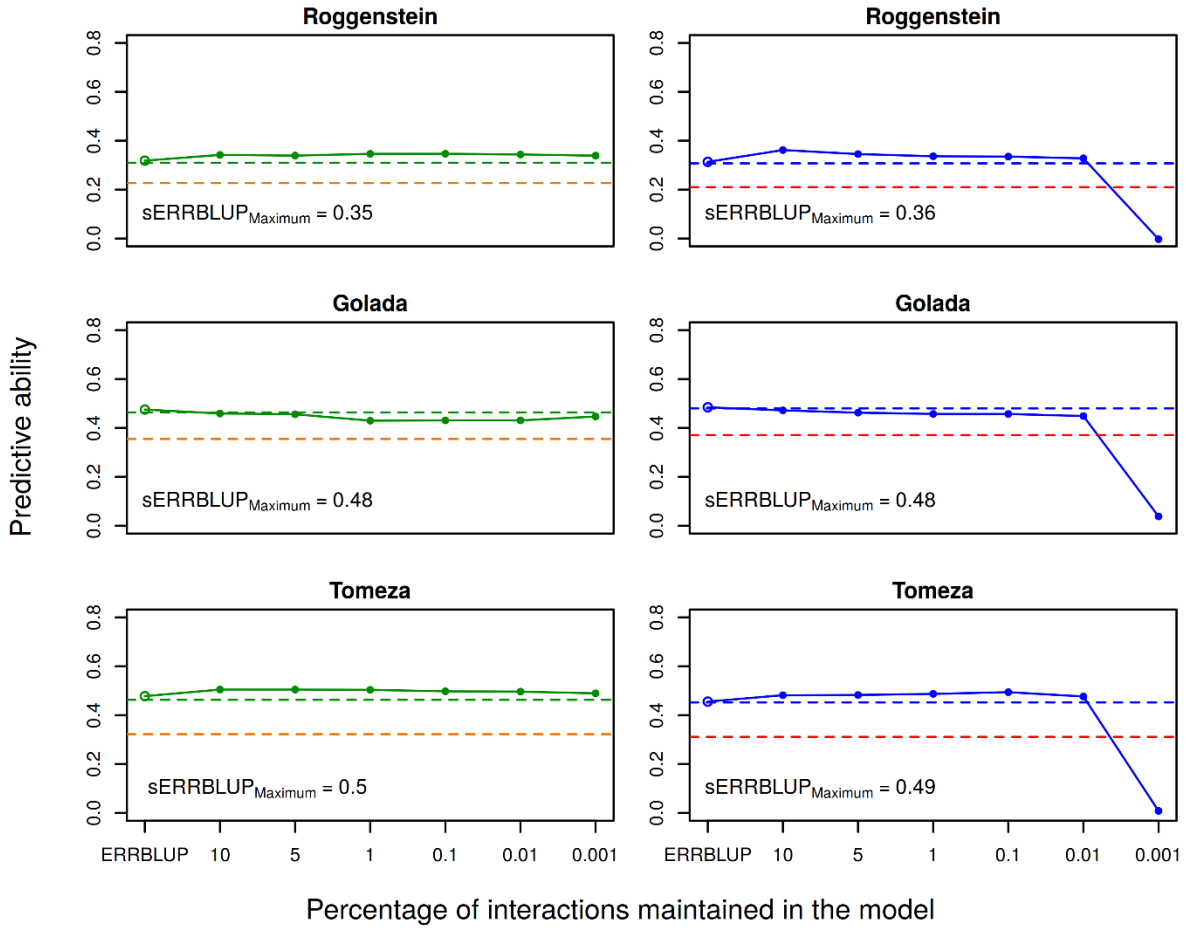
**Fig. S. 4.1:** Comparison of pre estimated genetic and residual variances and covariances of converged bivariate sERRBLUP (top 10%) based on the full dataset (dashed horizontal lines) and estimated genetic and residual variances and covariances of converged bivariate sERRBLUP (top 10%) based on training set in each run of 5-fold cross validation with 5 replicates (colored bars) for predicting EIN in 2018 when the additional environment is EIN in 2017 in KE for trait PH-V4.



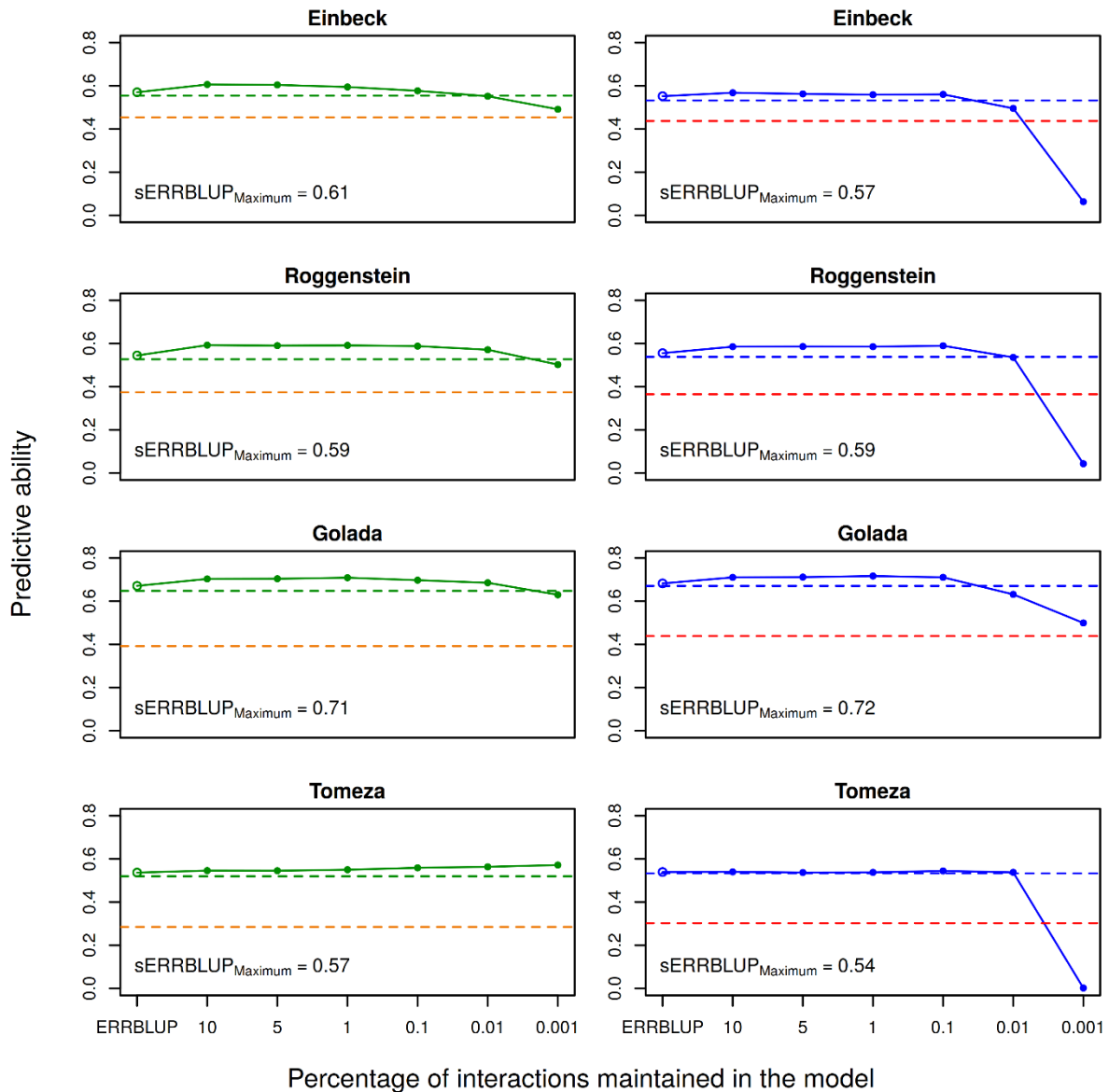
**Fig. S. 4.2:** The difference between the mean predictive ability of only the converged folds and the mean predictive ability of all folds in 5-fold cross validation with 5 replicates versus the number of the folds which did not converged across all traits in all combinations for both KE and PE in bivariate GBLUP, ERRBLUP, sERRBLUP.



**Fig. S. 4.3:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V3 in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

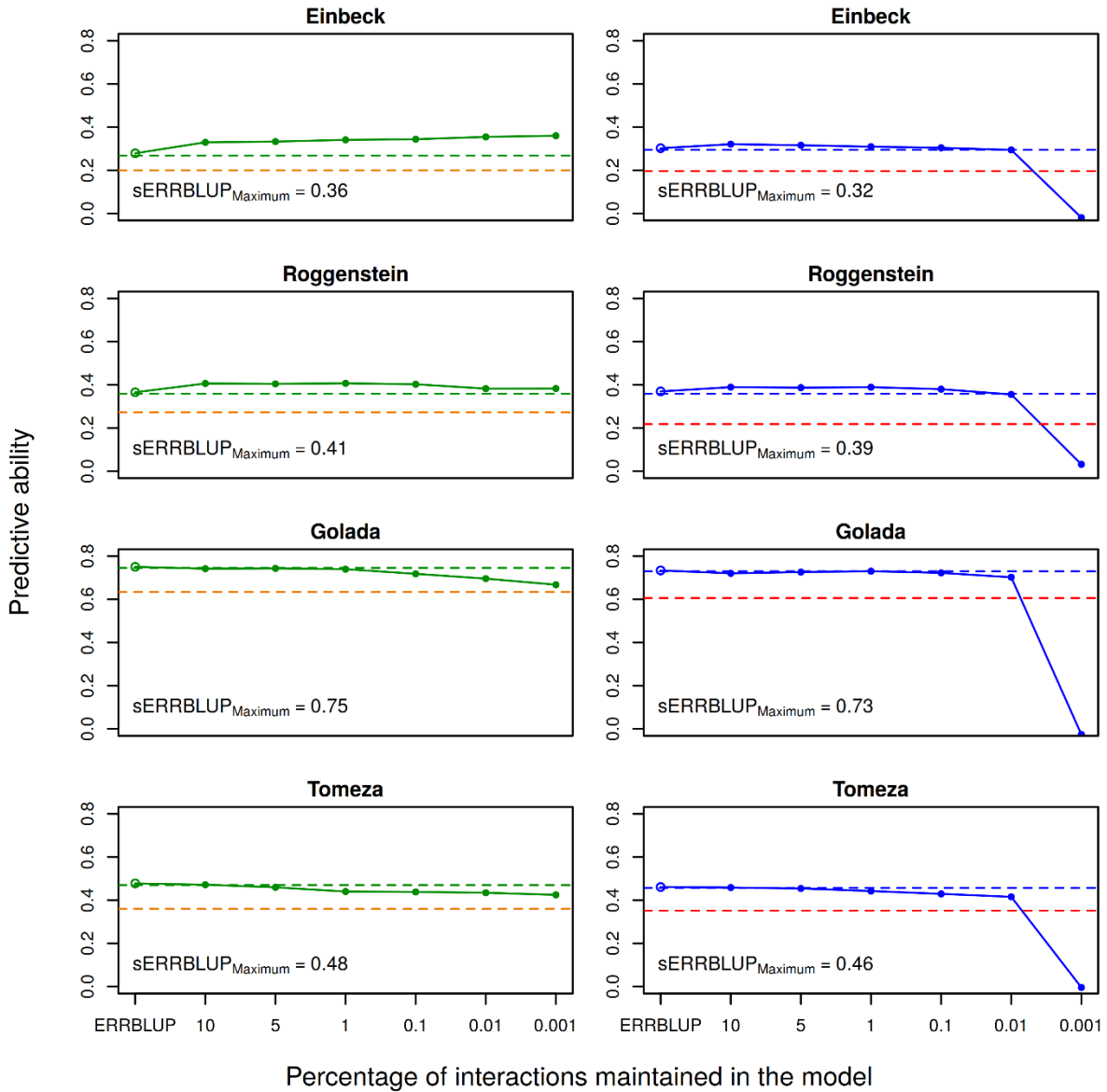


**Fig. S. 4.4:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V3 in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

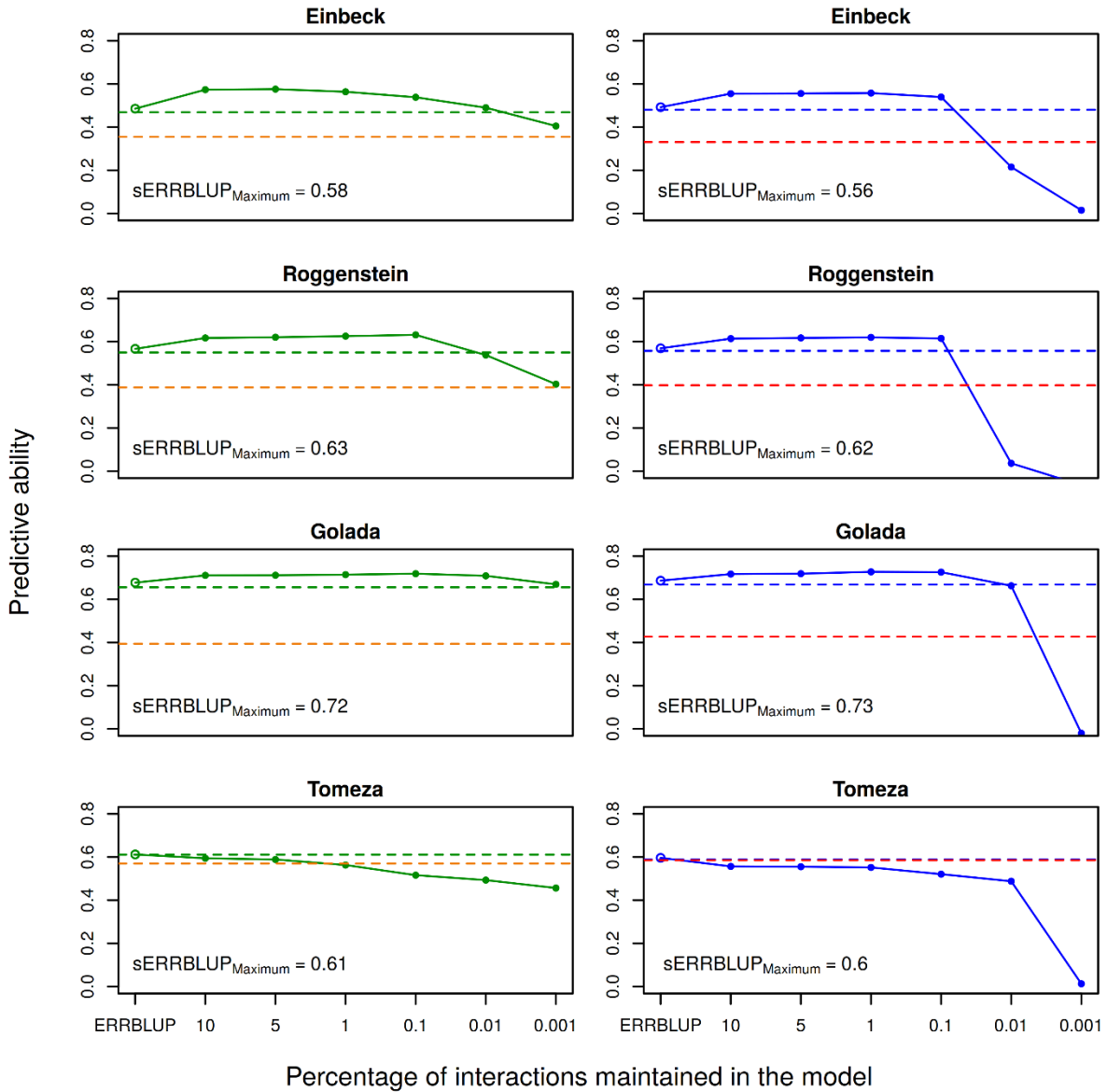


**Fig. S. 4.5:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V4 in KE based on pPruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

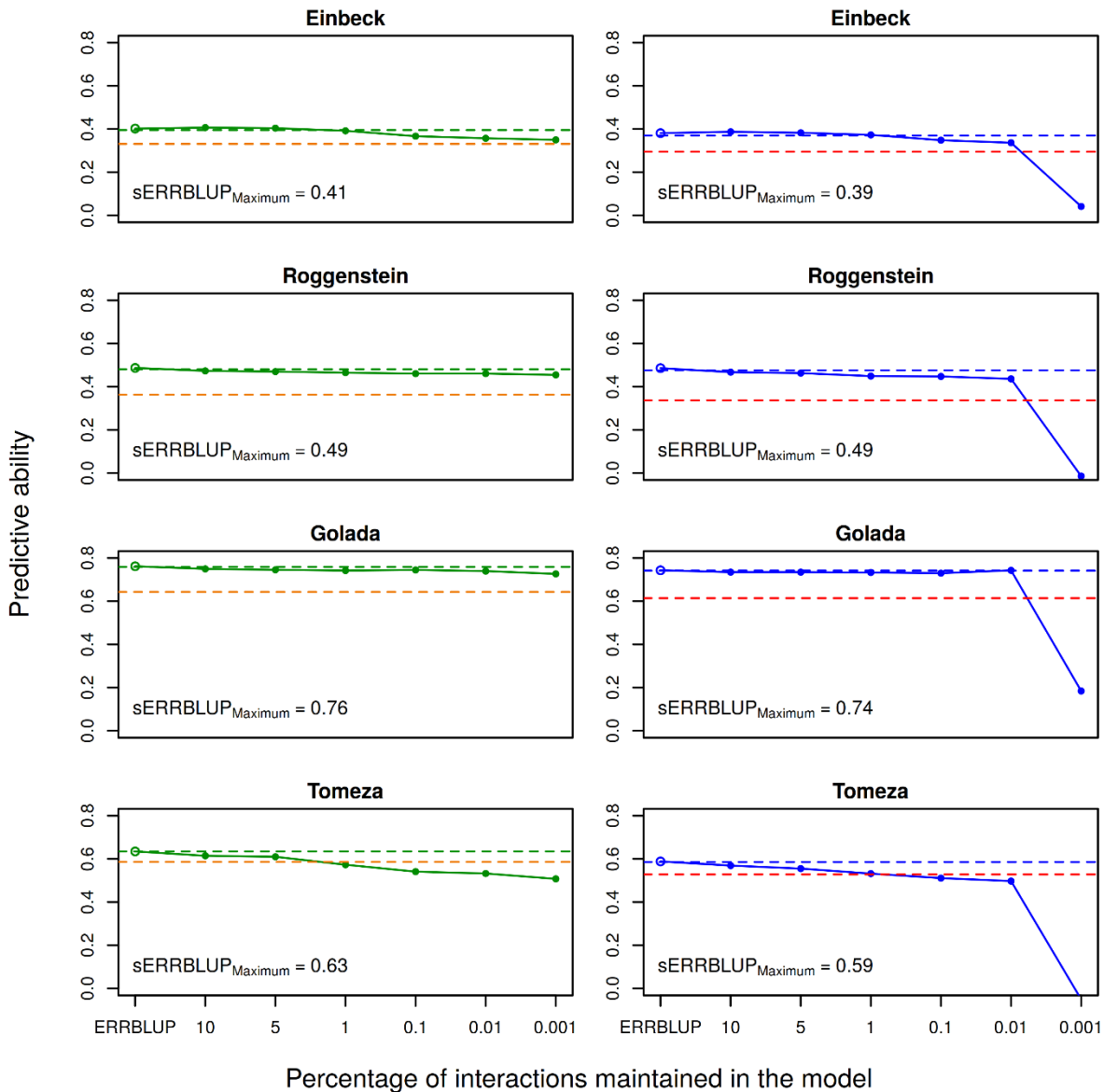




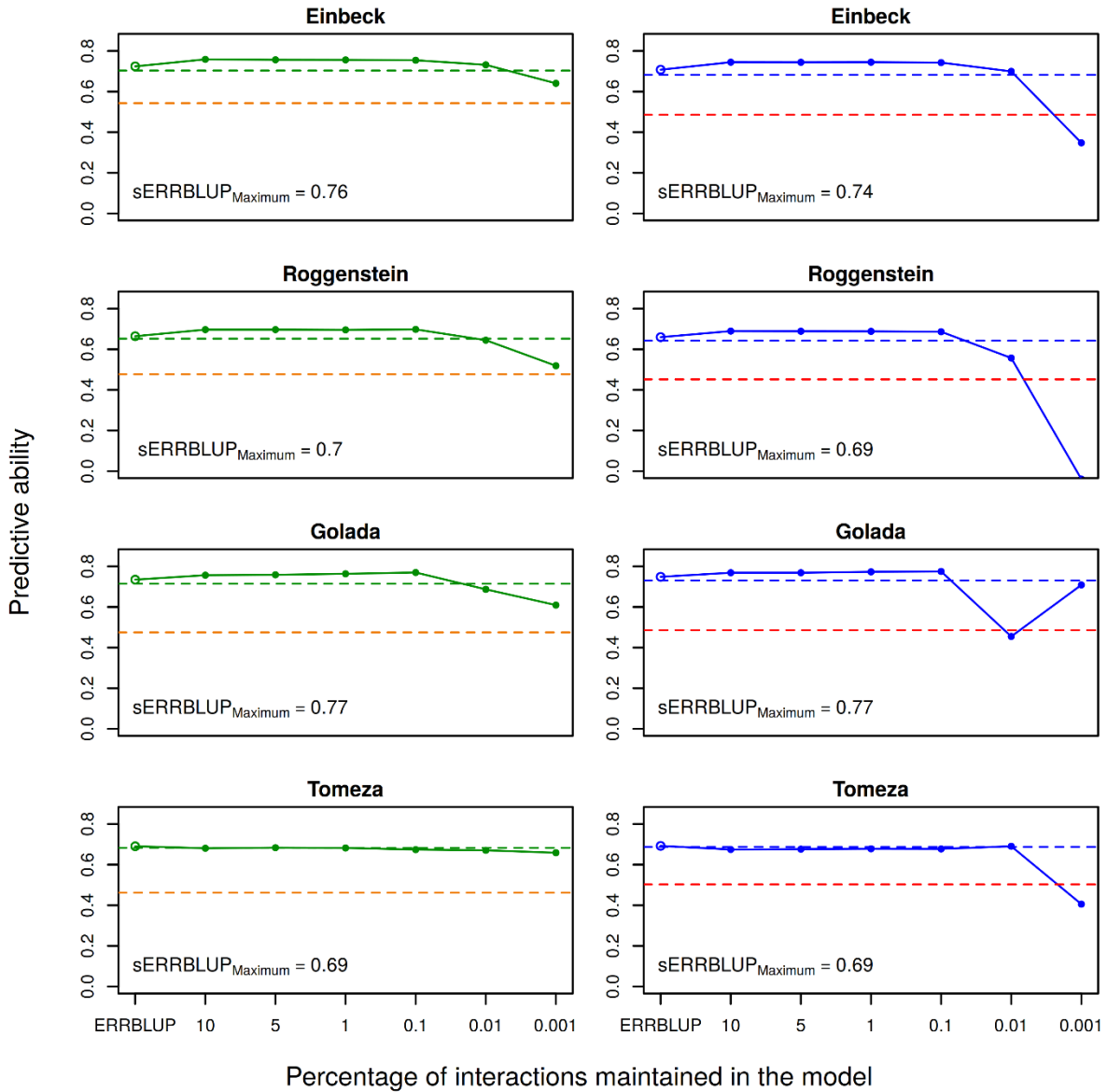
**Fig. S. 4.6:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V4 in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.



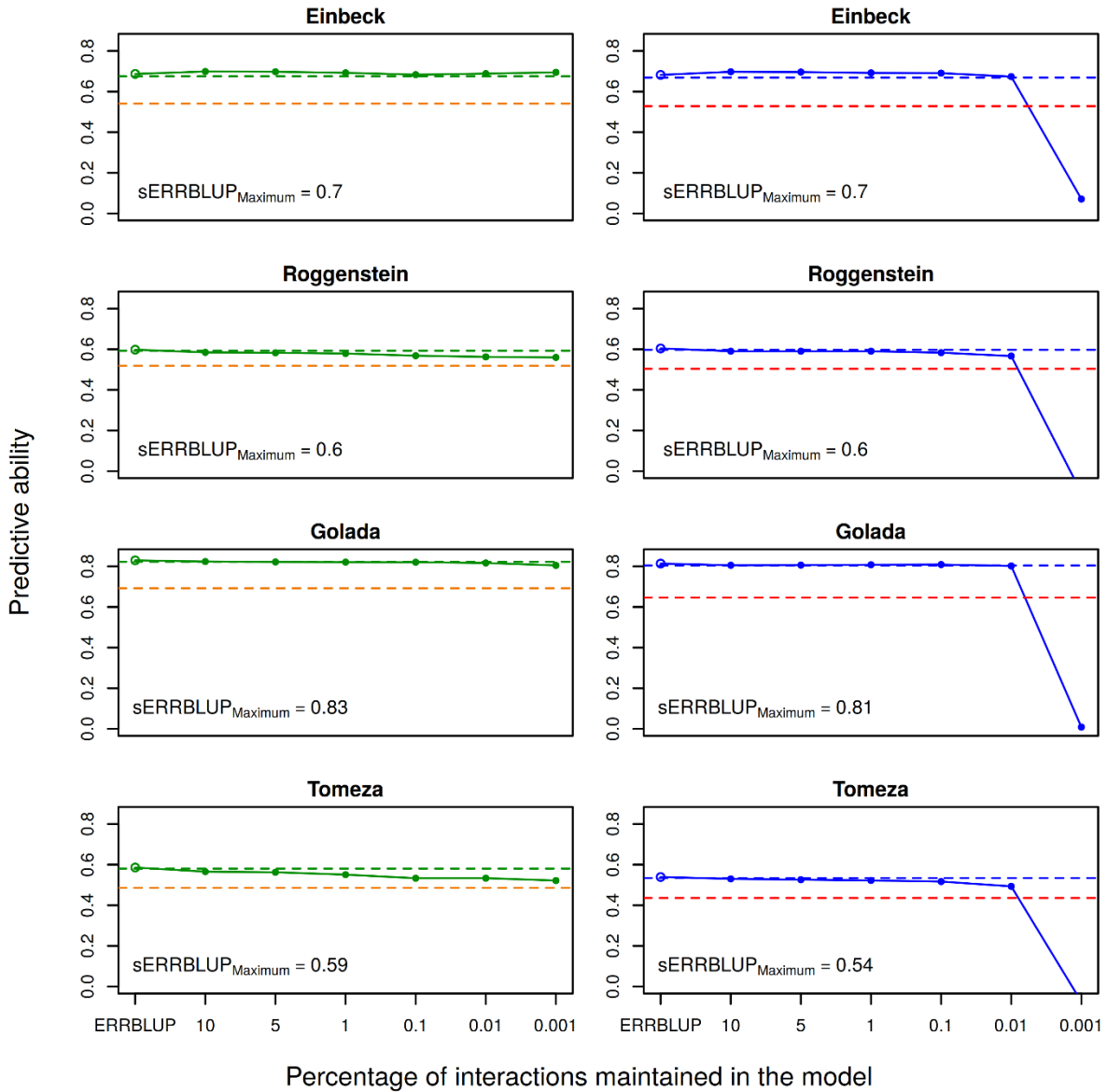
**Fig. S. 4.7:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V6 in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.



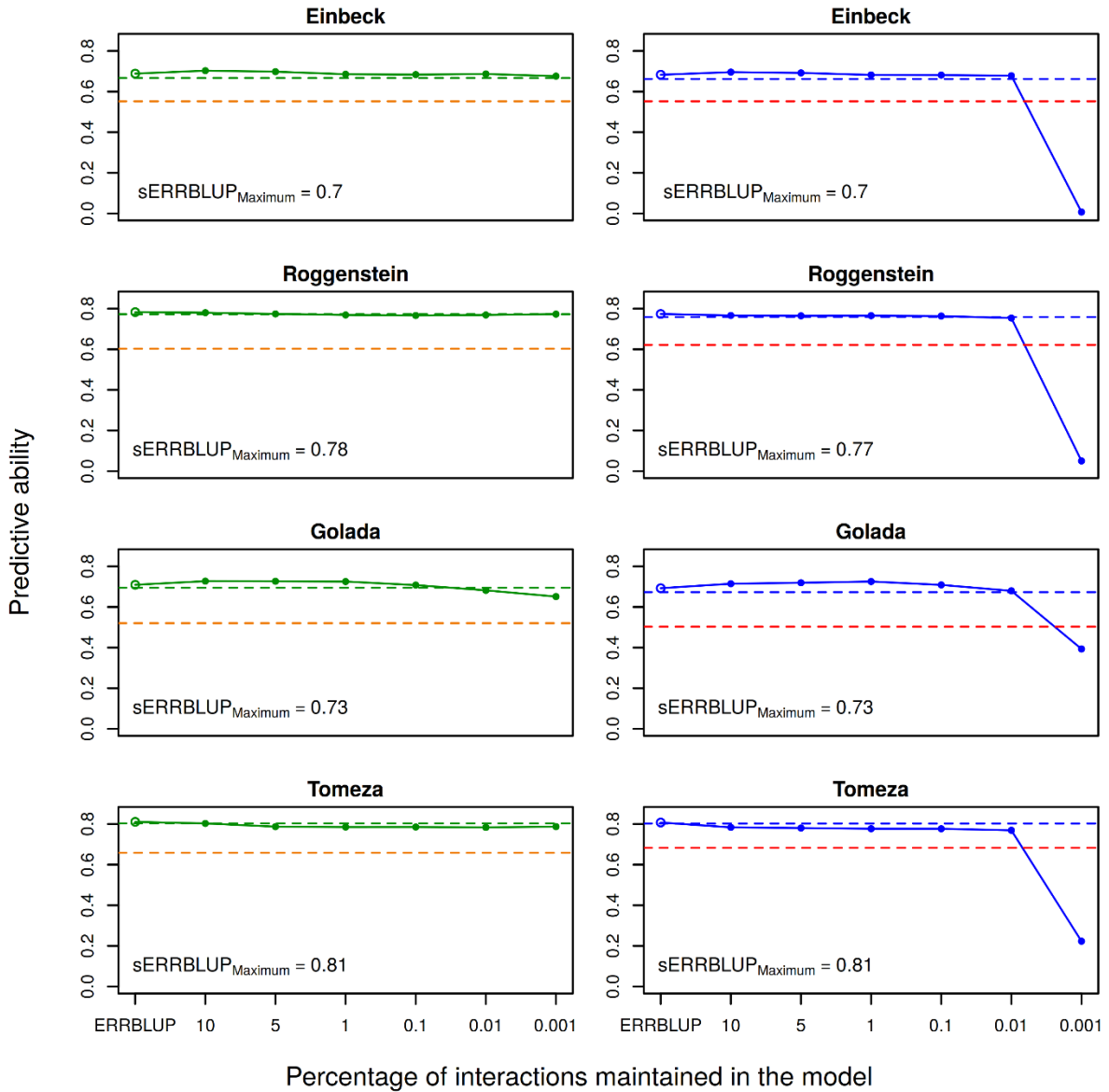
**Fig. S. 4.8:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V6 in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP<sub>Maximum</sub> indicates the maximum predictive ability obtained from bivariate sERRBLUP.



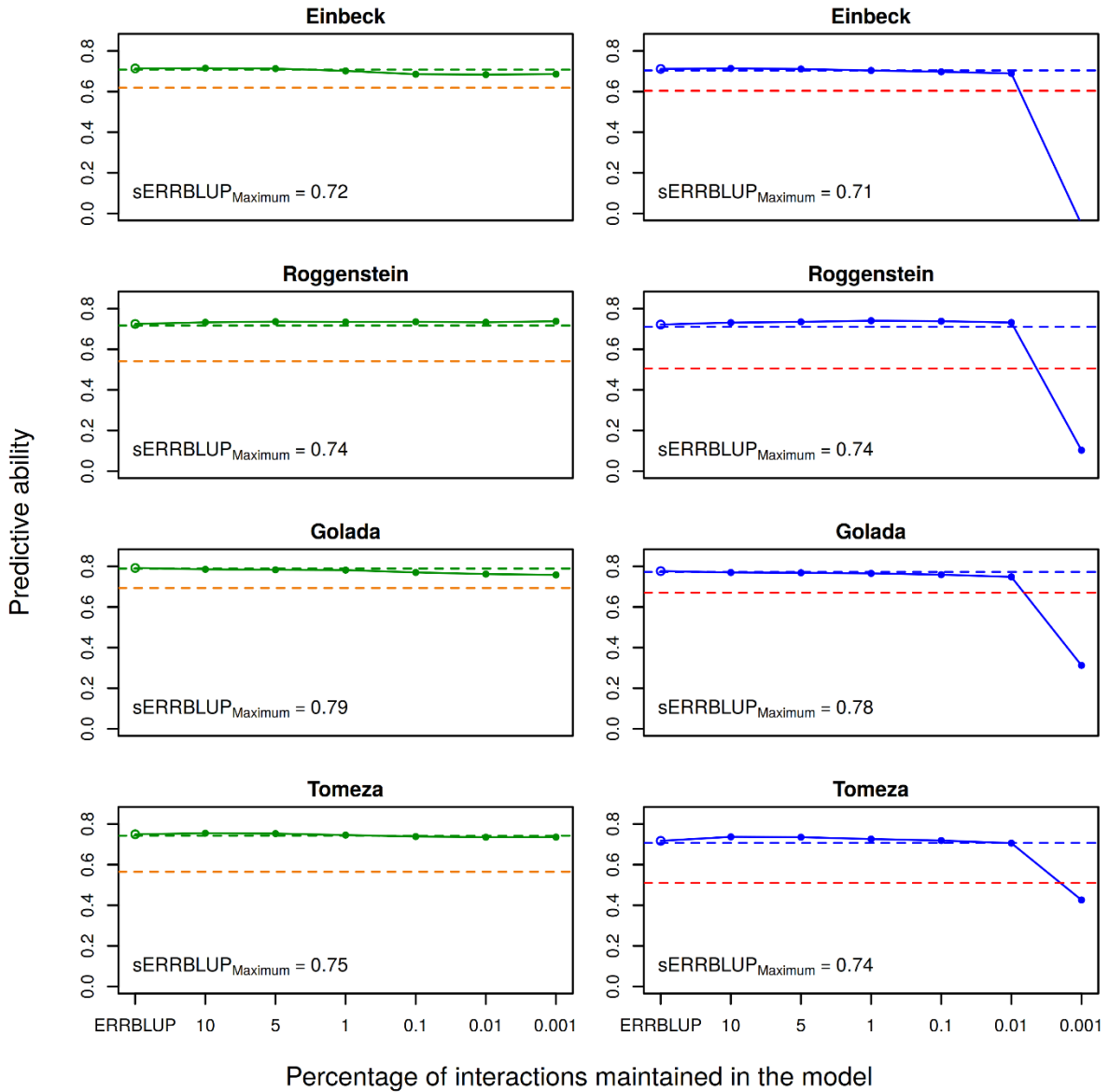
**Fig. S. 4.9:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH\_V6 in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.



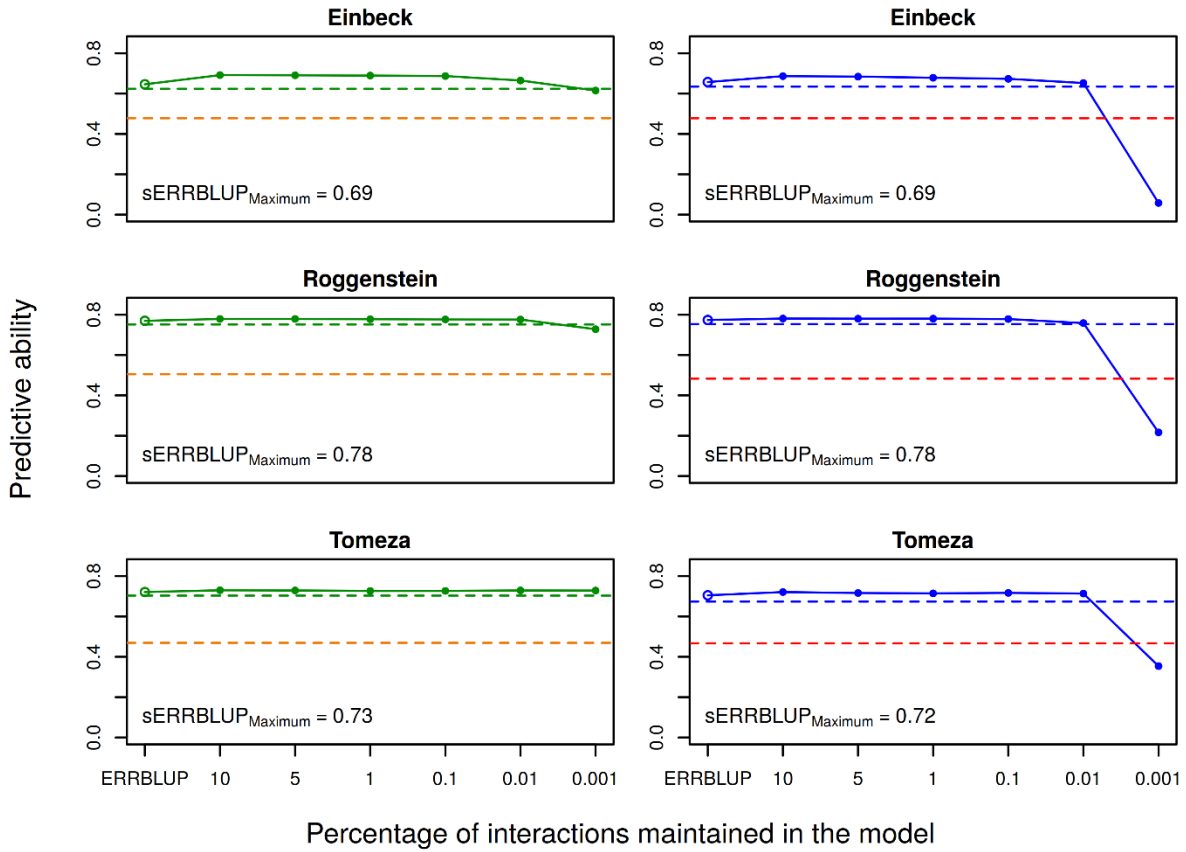
**Fig. S. 4.10:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH\_V6 in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.



**Fig. S. 4.11:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH\_final in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

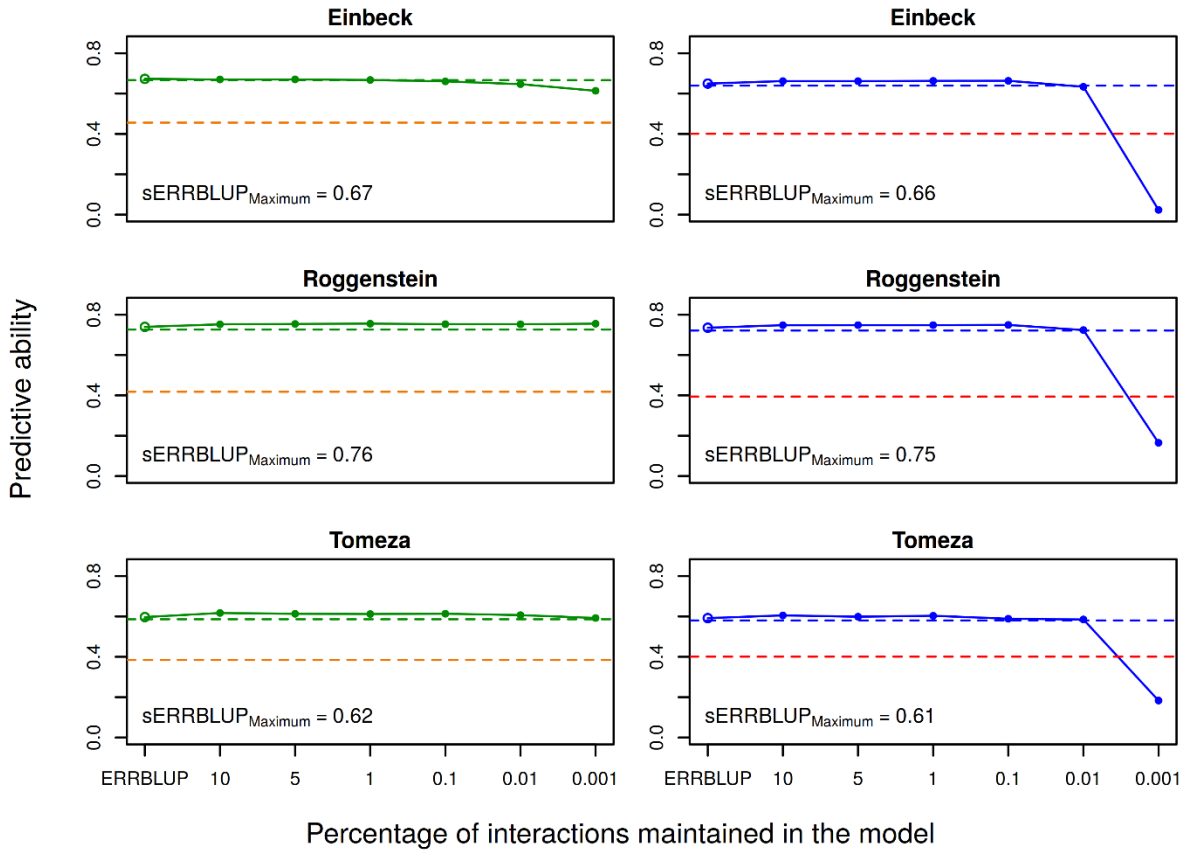


**Fig. S. 4.12:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH\_final in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

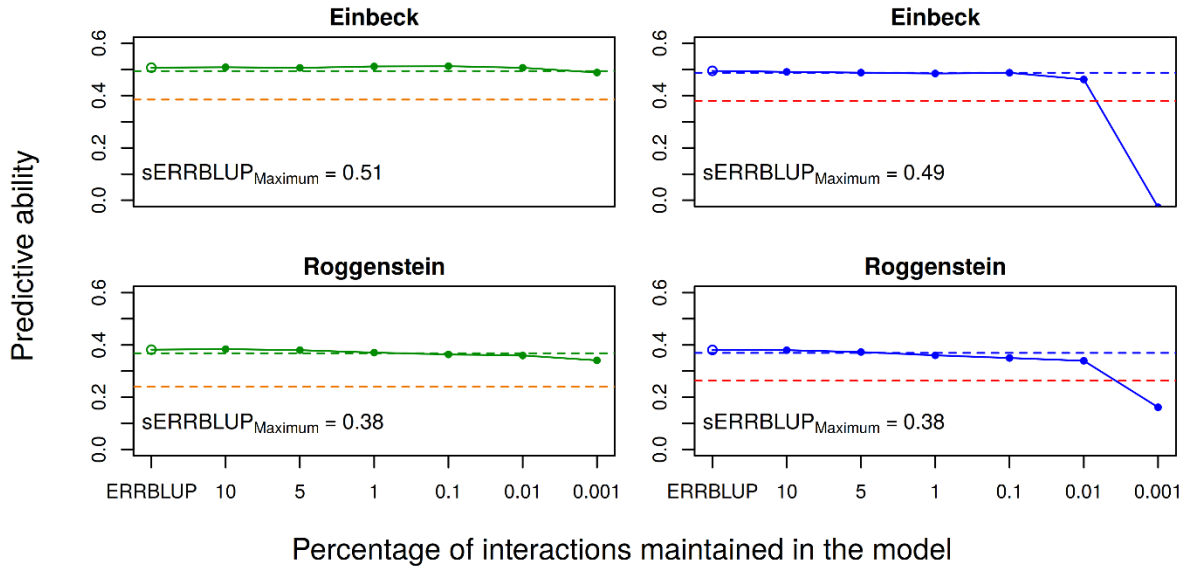


**Fig. S. 4.13:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait FF in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

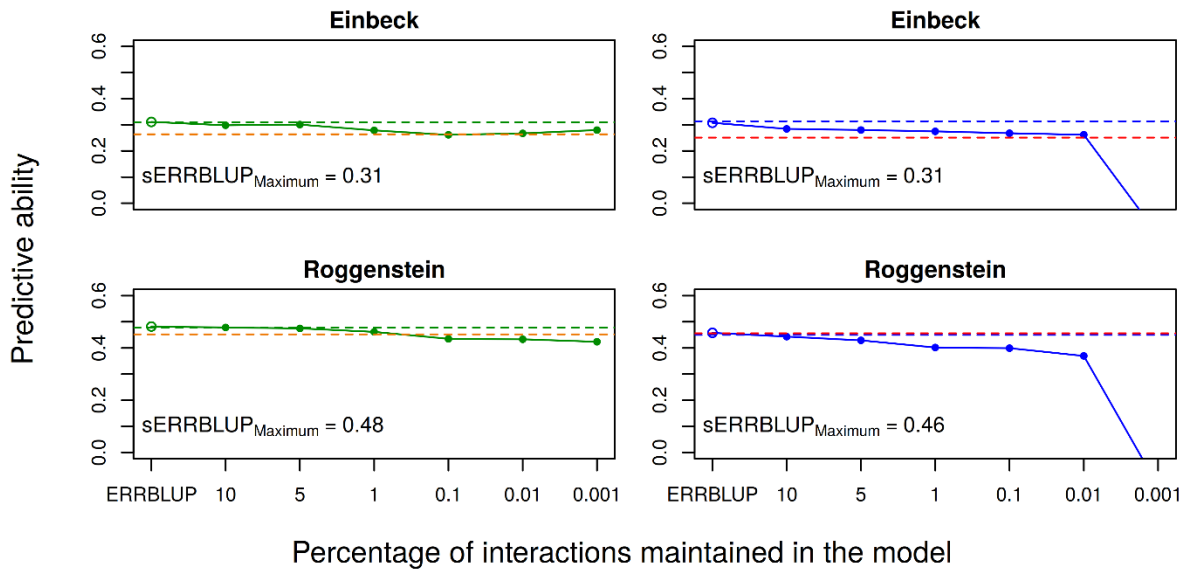




**Fig. S. 4.14:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait FF in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

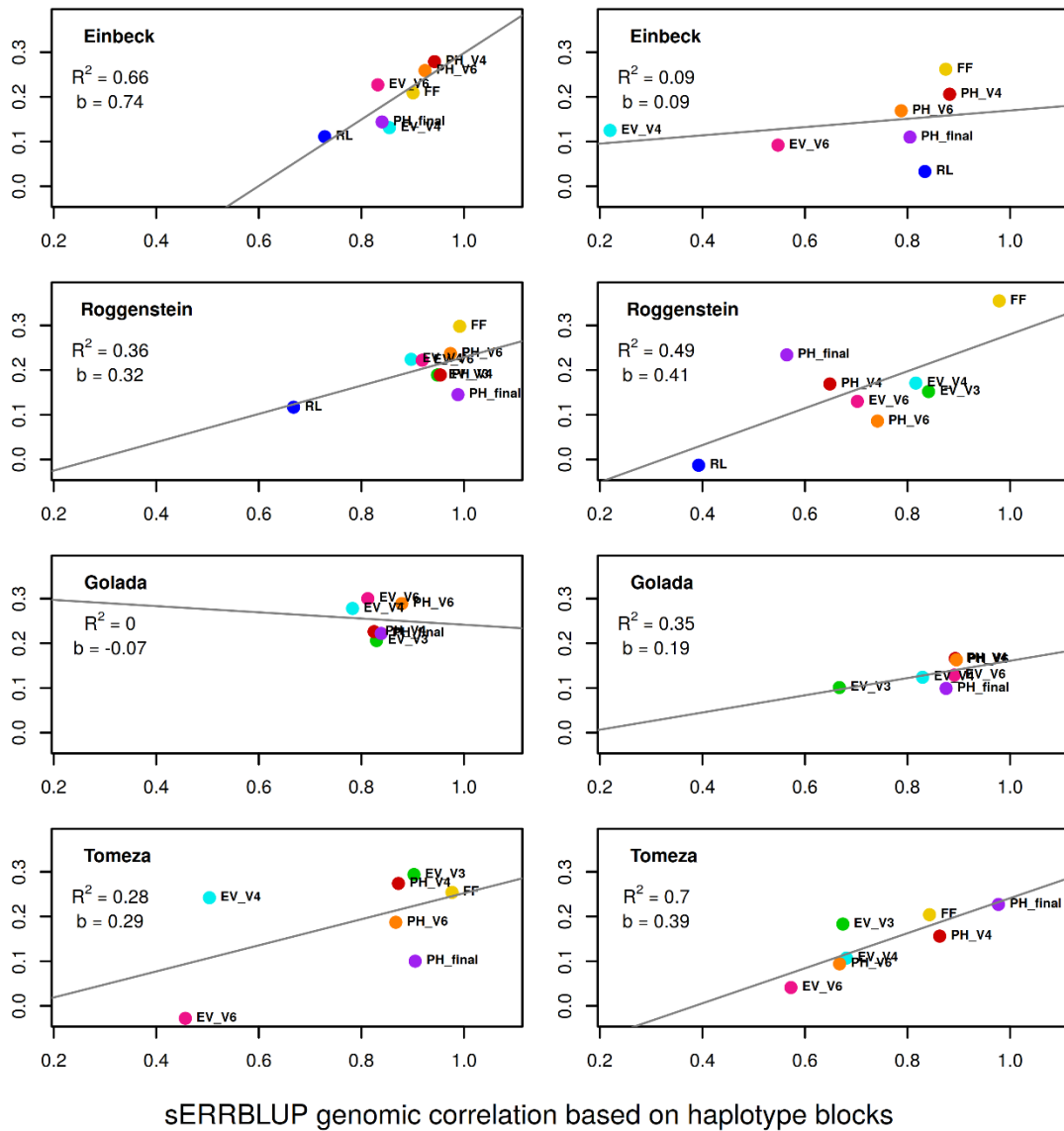


**Fig. S. 4.15:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait RL in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

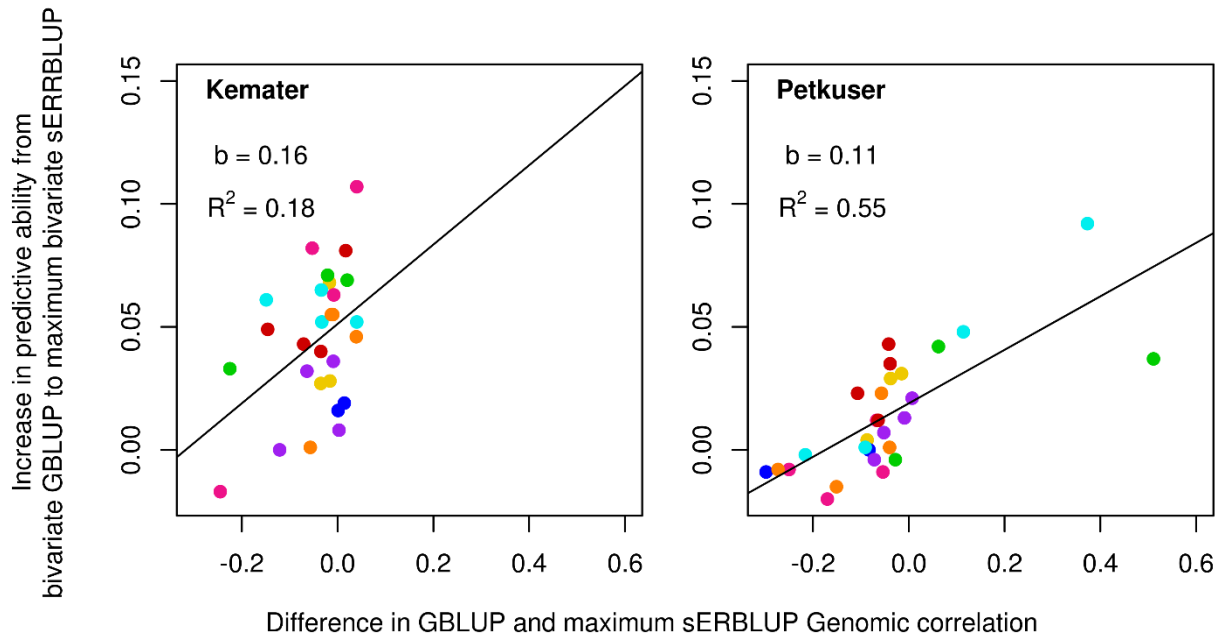


**Fig. S. 4.16:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait RL in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP.

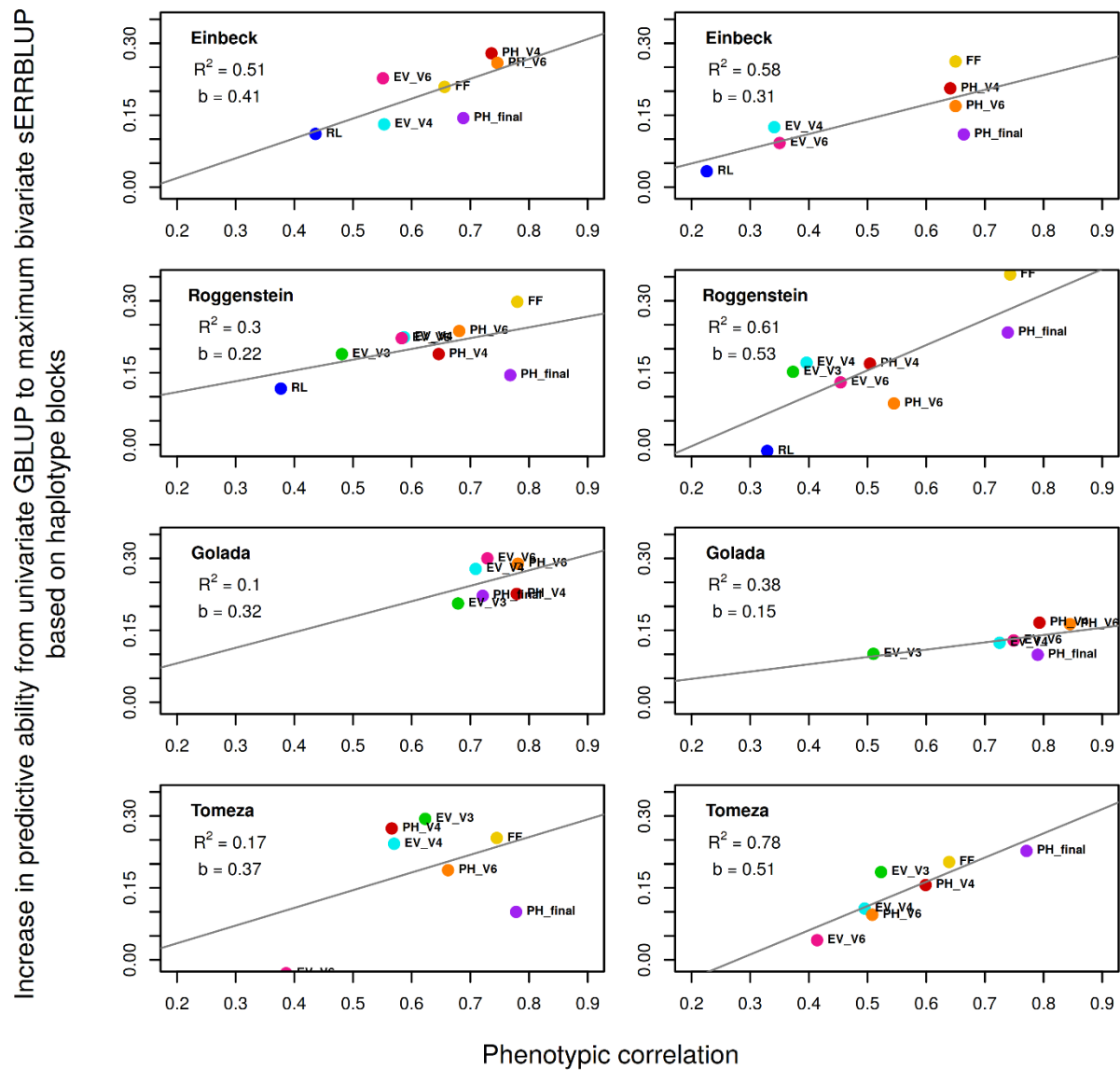
Increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP based on haplotype blocks



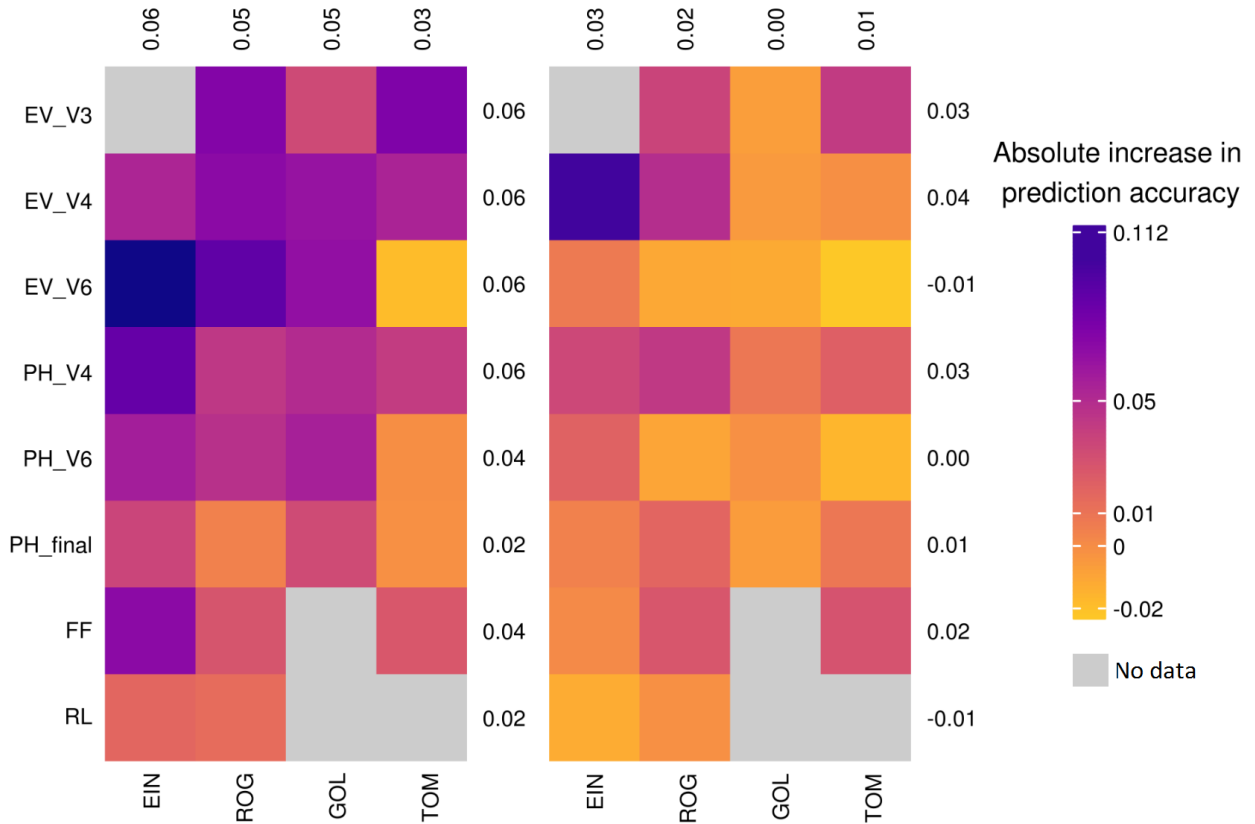
**Fig. S. 4.17:** Regression of the absolute increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP on the respective sERRBLUP genomic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits. In each panel, the overall linear regression line (gray solid line) with the regression coefficient ( $b$ ) and R-squared ( $R^2$ ) are shown.



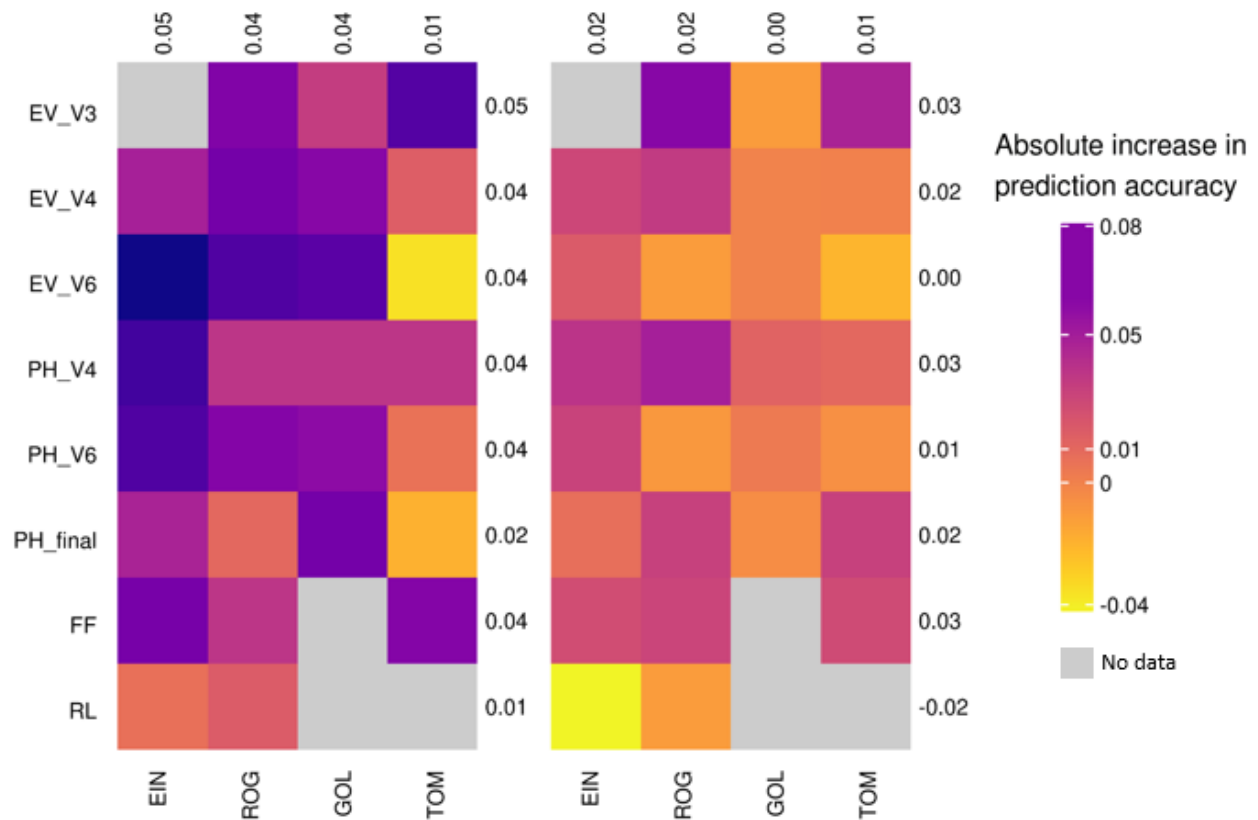
**Fig. S. 4.18:** Regression of the absolute increase in predictive ability from bivariate GBLUP to maximum bivariate sERRBLUP on the difference between the GBLUP genomic correlation and maximum sERRBLUP genomic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits, when utilizing pruned set of SNPs. In each panel, the overall linear regression line with the regression coefficient ( $b$ ) and R-squared ( $R^2$ ) are shown. The colors green, light blue, pink, red, orange, purple, yellow and dark blue represent the phenotypic traits EV\_V3, EV\_V4, EV\_V6, PH\_V4, PH\_V6, PH\_final, FF and RL, respectively.



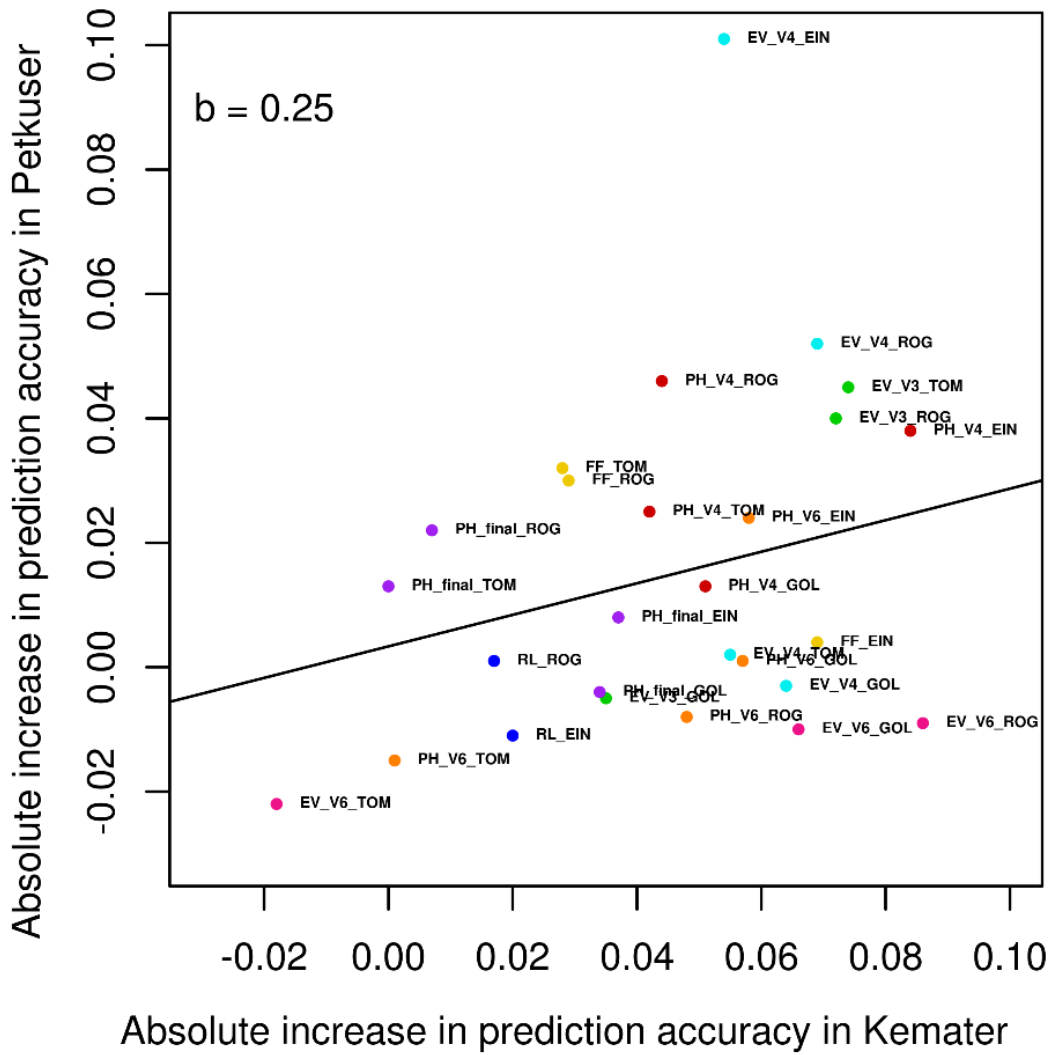
**Fig. S. 4.19:** Regression of the absolute increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP on the phenotypic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits. In each panel, the overall linear regression line (gray solid line) with the regression coefficient ( $b$ ) and R-squared ( $R^2$ ) are shown.



**Fig. S. 4.20:** Absolute change in prediction accuracy from bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP based on pruned set of SNPs in KE (left side plot) and in PE (right side plot). The average absolute change in prediction accuracy for each trait and environment is displayed in all rows and columns, respectively.

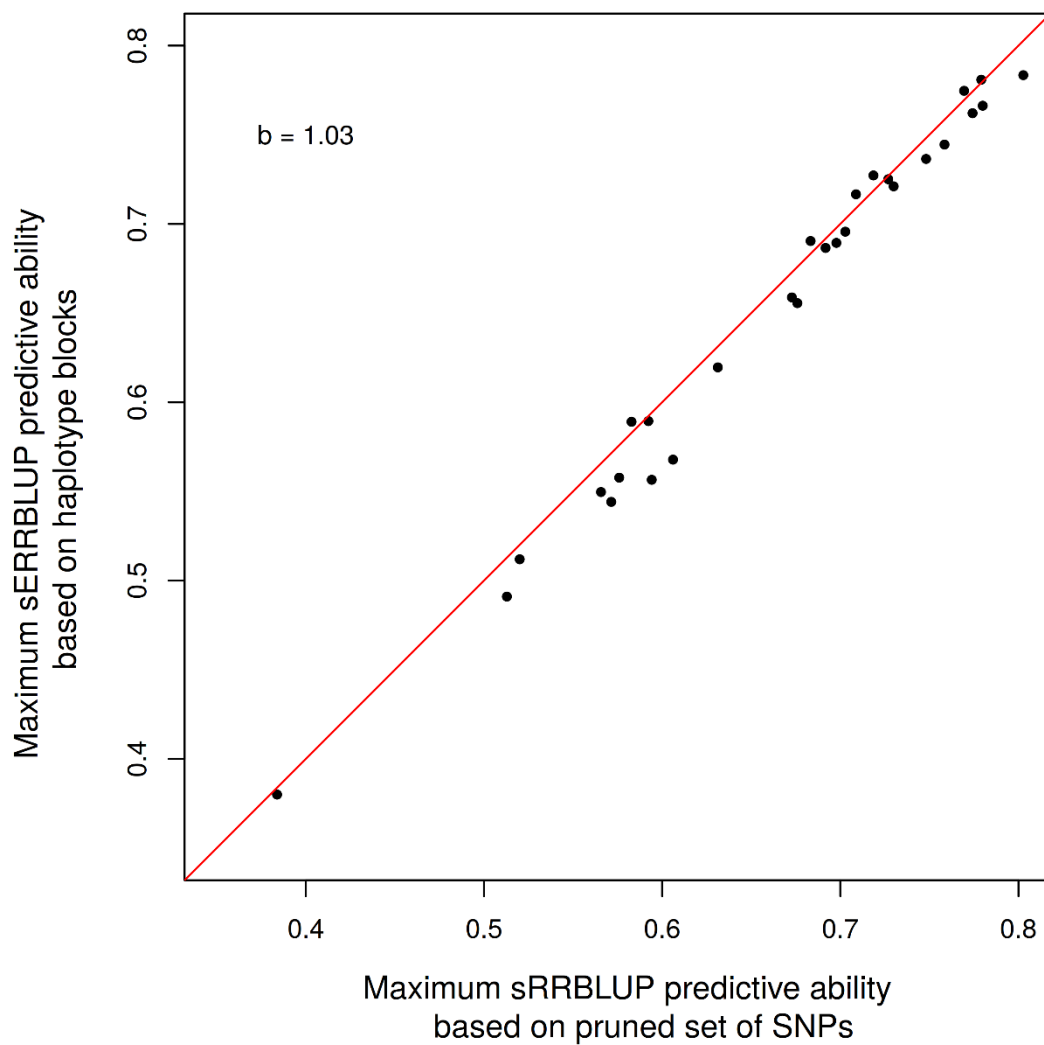


**Fig. S. 4.21:** Absolute change in prediction accuracy from bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP based on haplotype blocks in KE (left side plot) and in PE (right side plot). The average absolute change in prediction accuracy for each trait and environment is displayed in all rows and columns, respectively.



**Fig. S. 4.22:** Absolute change in prediction accuracy from bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP in PE vs. KE. The black line represents the overall linear regression line.





**Fig. S. 4.23:** Maximum sERRBLUP predictive ability based on pruned set of SNPs versus Maximum sERRBLUP predictive ability based on haplotype blocks across all traits in all environments and both KE and PE. ***b*** represents the regression coefficient.



## 5 Discussion

“To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science”

Albert Einstein

In this chapter, other methods of variable selection in epistasis model are investigated and compared to our proposed selection method in sERRBLUP model. This is followed by a review and elaboration of the underlying features affecting the predictive abilities of genomic prediction models, beside presentation of potential future research topics.

### 5.1 Different methods for pairwise SNP interaction selection

In this thesis, we have discussed that ERRBLUP is considered as the categorical epistasis (CE) model (Martini *et al.*, 2017) which does not possess the undesirable coding dependent feature of EG-BLUP and performs as well as the best EG-BLUP obtained through symmetrically coded markers (Martini *et al.*, 2017) (chapter 3). We then positively evaluated the sERRBLUP model as a selective CE model by variable selection across environments in maize dataset for both KE and PE. sERRBLUP was also positively evaluated for the prediction of simulated dataset in a single environment based on the wheat data set genotypes of CIMMYT inbred lines (Pérez and de los Campos, 2014) (chapter 2).

In order to assess different options of variable selection methodologically, our proposed variable selection method in the sERRBLUP model is compared with the best of two variable selection method in reduced EG-BLUP (Martini *et al.*, 2016). Therefore, the obtained predictive abilities in reduced EG-BLUP based on variable selection using the effect sizes obtained by ridge regression are first compared (Martini *et al.*, 2016) with an approach based on an epistatic GWAS (eGWAS) test (Aulchenko *et al.*, 2007) in the wheat dataset (Pérez and de los Campos, 2014). Genome-wide association studies (GWAS) enable comprehensive scanning of genetic components which control heritable traits (Nordborg and Weigel, 2008; Myles *et al.*, 2009; Brachi *et al.*, 2011). In GWAS analysis, statistical association between SNPs and complex traits is tested through individually testing each SNP for an association with the target trait (Yang *et al.*, 2010). Then, the precise p-values are used to account for the numerous significance tests. However, this may result in the occurrence of false positives, whereas many real associations will be missed especially for individual SNPs with small effects on a trait (Yang *et al.*, 2010). However, fitting all SNPs simultaneously is an alternative approach to overcome this problem (Hayes, Visscher, *et al.*, 2009).

Epistasis GWAS (eGWAS) was therefore developed for variable selection as a comparison for variable selection based on ridge regression in reduced EG-BLUP. In both approaches, subnetworks were inferred from one environment to structure the genomic relationship matrix which is then used for genomic prediction in another environment like sERRBLUP. In this context, the EG-BLUP model (Jiang and Reif, 2015) was first considered as a full epistasis model, then EG-BLUP was reduced to the subsets of pairwise SNP interactions with the highest absolute effect sizes at 5 percent steps until 10 percent of interactions with highest absolute effects sizes maintained in the model which is considered as rrBLUP-based selection of interactions (Martini *et al.*, 2016). However, in eGWAS-based selection of interaction, subnetworks were determined in each environment by considering the most significant pairwise SNP interactions through scanning the data for pairwise SNP interaction by GenABEL R-package (Aulchenko *et al.*, 2007). In this regard, the model with only additive effects was tested versus the model with both additive effects and epistasis effects. Afterwards, the relationship matrices were constructed for the desired proportions of interactions selected based on the obtained p-values as (Martini *et al.*, 2016)

$$G = \sum_{k,j \in I} (M_{.,k}M'_{.,k}) \circ (M_{.,j}M'_{.,j})$$

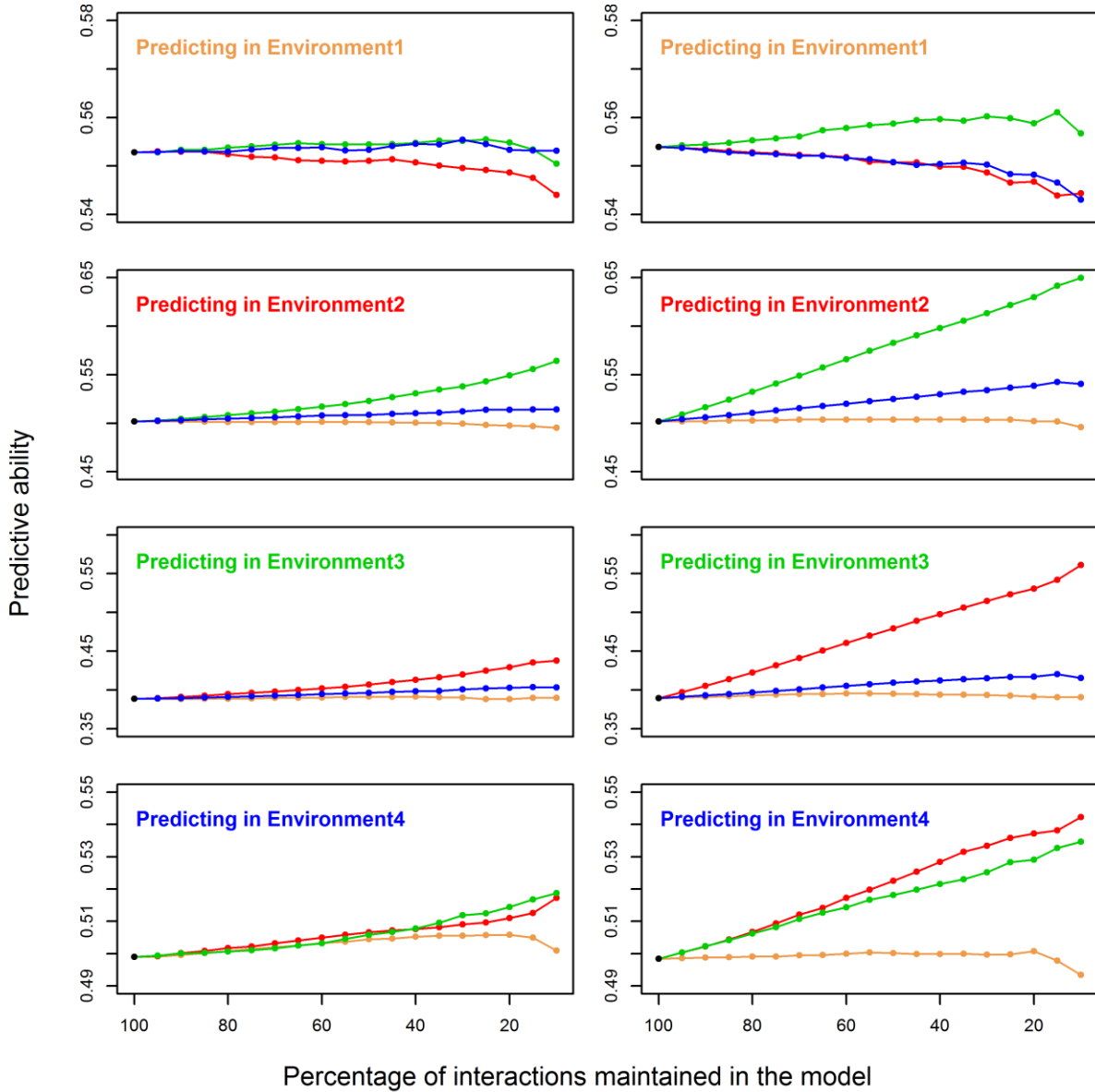
where  $M_{.,j}$  and  $M_{.,k}$  indicate the  $j^{th}$  and  $k^{th}$  marker codes of all individual in which marker  $k$  and  $j$  are the two markers which have interaction and  $I$  indicates the matrix representing the interacting markers for the desired proportion of interactions. This relationship matrix was then used for genomic prediction of phenotypes in the other environment.

The results illustrate an improvement in predictive ability in one environment by variable selection in another environment under both methods of variable selection in the wheat dataset (Pérez and de los Campos, 2014) (Fig. 5.1). However, Fig. 5.1 demonstrates that while the increasing trend with eGWAS-based selection is the same as with rrBLUP-based selection of interactions, the overall predictive ability is lower using the eGWAS approach, especially if a high proportion of epistatic interactions is accounted for. Therefore, reduced EG-BLUP with interaction selection based on effects sizes obtained from rrBLUP was shown to be the preferable method for variable selection.

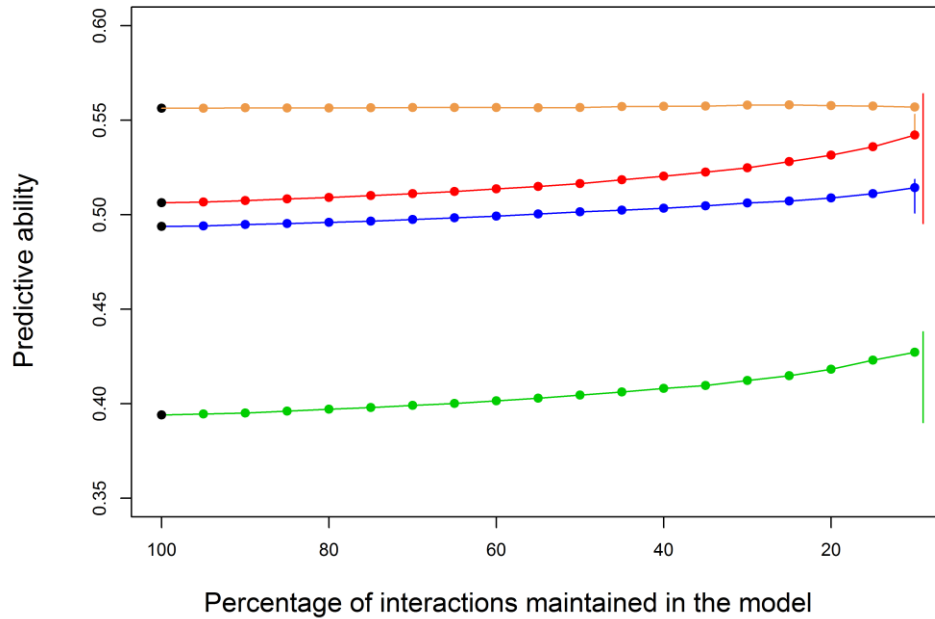
In Chapter 2, we have shown that interaction selection based on absolute effects sizes in sERRBLUP leads to higher predictive ability compared to reduced EG-BLUP. Therefore, sERRBLUP outperforms reduced EG-BLUP based on both rrBLUP-based and eGWAS-based selection of interactions.

Furthermore, eGWAS-based interaction selection by considering three environments jointly to predict the fourth environment was investigated (Fig. 5.2) by a different method than the one we have proposed in Chapter 3. Based on this approach, the corresponding relationship matrices for predicting the fourth environment were determined by averaging the relationship matrices of the other three environments instead of averaging the phenotypic values of three environments and constructing the new relationship matrix as it was done in Chapter 3. This approach in general proved to be robust in predictive abilities which were close to the best predictive abilities obtained by variable selection across a single environment. However, the proposed method for predicting

across multiple environments jointly in Chapter 3 performed as well as or better than prediction based on the single environment. Therefore, the proposed method for predicting across multiple environments jointly in Chapter 3 is a convenient method in this regard.



**Fig. 5.1:** Predictive ability of EG-BLUP within environment (black circle) as a full epistasis model and reduced EG-BLUP across environments (colored circles and solid lines) when interaction selections are based on eGWAS (left) and rrGBLUP (right) for wheat grain yield. In each panel, the solid lines' color indicates the environment in which the relationship matrices were determined by variable selection.



**Fig. 5.2:** Predictive ability of EG-BLUP within environment (black circle) as a full epistasis model and reduced EG-BLUP across all three environments jointly when interaction selections are based on eGWAS. Orange, red, green and blue lines colors indicate the predictive abilities in environment 1, 2, 3, and 4, when the other three environments are used jointly for variable selection. Vertical lines represent the range of accuracies obtained when selection of sub-networks was in a single environment rather than combining three environments and predict the fourth based on eGWAS.

## 5.2 Linkage disequilibrium based SNP pruning and its influence on genomic prediction model's performance

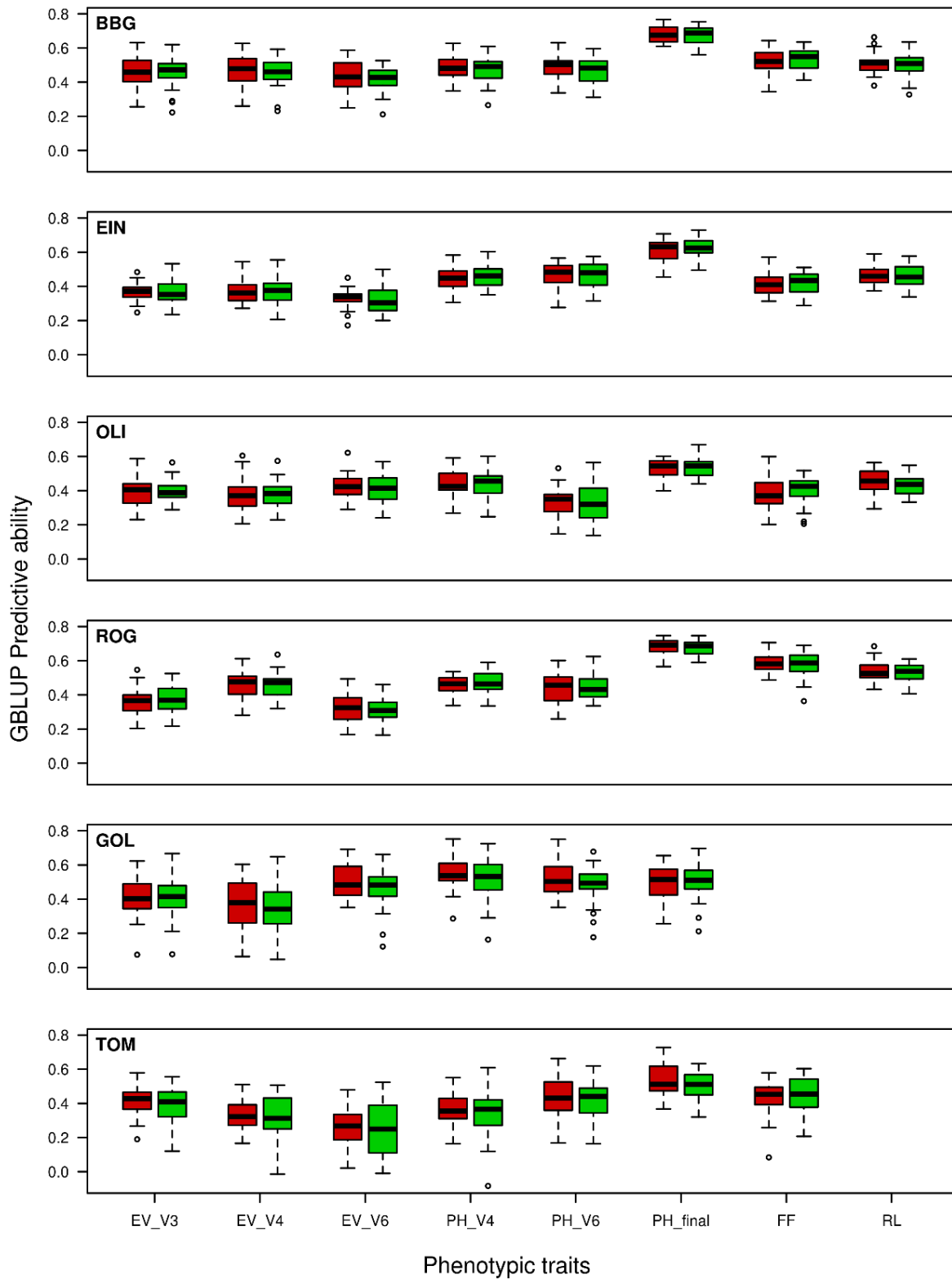
Linkage disequilibrium (LD) as the nonrandom association of alleles at two or more loci – which is measured as the correlation coefficient ( $r^2$ ) between phased alleles (Slatkin, 2008) – is a potential factor affecting the prediction model's accuracy. High level of pairwise LD in SNP array or whole genome sequence data might influence the genomic prediction model's performance and efficiency. Therefore, the loci which show a high level of pairwise LD are removed by LD pruning (Calus and Vandenplas, 2018). So far, several tools have been developed to compute pairwise LD between SNPs (Barrett *et al.*, 2004; Purcell *et al.*, 2007; Calus and Vandenplas, 2018). One of these tools is PLINK (Purcell *et al.*, 2007) which is designed to handle large dataset including hundreds of thousands of genotyped markers for thousands of individuals and prune for LD. Linkage disequilibrium based SNP pruning generates a pruned subset of SNPs which are in approximate linkage equilibrium. This is done based on variance inflation factor (VIF) given by

$$VIF = \frac{1}{(1-r^2)},$$

where  $r^2$  is the squared correlation coefficient between SNPs and the linear combinations of SNPs.

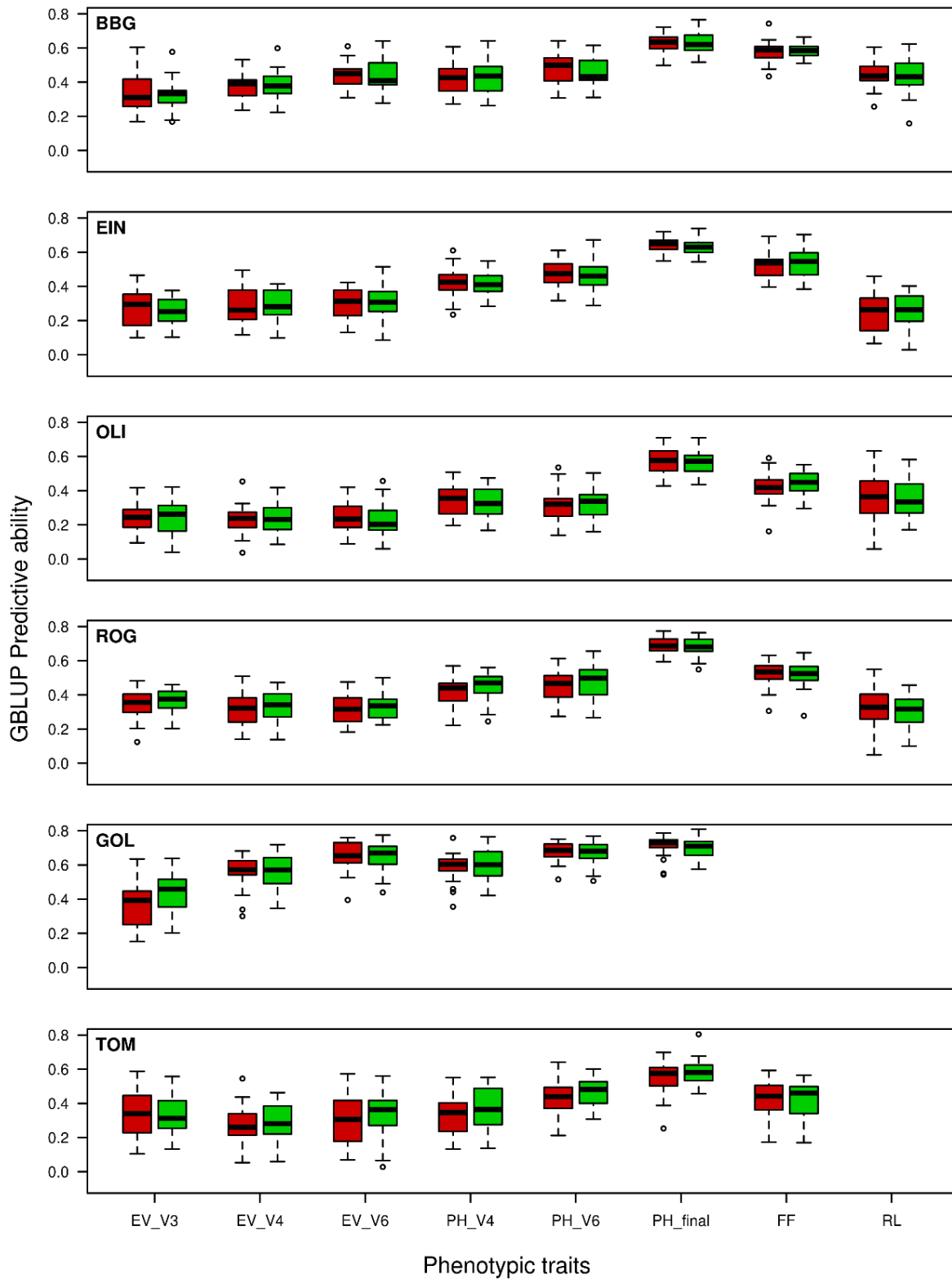
In this thesis, the *VIF* pruning routine was performed by giving three parameters of 50, 5 and 2 explained in Chapter 3; 50 representing the window size in SNPs, 5 representing the number of SNPs to shift the window at each step and 2 representing *VIF*. With this, all variants in 50 SNP window with *VIF* > 2 were removed. This indicates that all the variants with  $r^2 > 0.5$  were removed from the respective window. If *VIF* = 1 was considered, then the SNP was completely independent of all other SNPs. Therefore, *VIF* = 1.5 or 2 should be used practically, since defining a very low *VIF* and/or too large a window size may result in removing too many SNPs (Purcell *et al.*, 2007). In our study, the panel of 501'124 SNPs was pruned and generated a subset of 25'437 SNPs for KE and 30'212 SNPs for PE by PLINK (v1.07).

The comparison in univariate GBLUP predictive ability obtained based on the 501'124 SNPs and the pruned set of SNPs across series of studied traits for KE and PE landraces are shown in Fig. 5.3 and Fig. 5.4, respectively. The average of GBLUP predictive ability is shown to be higher for some of the traits in some environments when the pruned set of markers are utilized instead of the panel of 501'124 SNPs. This improvement in GBLUP predictive ability when utilizing pruned sets is higher in PE than KE. Therefore, considering pruned sets of markers can potentially help to increase the performance and efficiency of our genomic prediction models. Moreover, generation of smaller subset of SNPs based on LD pruning enable us to incorporate all pairwise SNP interactions in ERRBLUP or an optimal subset of pairwise SNP interactions in sERRBUP at a computationally feasible manner.



**Fig. 5.3:** Comparison of predictive abilities obtained from univariate GBLUP when the panel of 501'124 SNPs is used (red) and when the pruned set of 25'437 SNPs is used (green) in KE for prediction within environments across series of all traits.



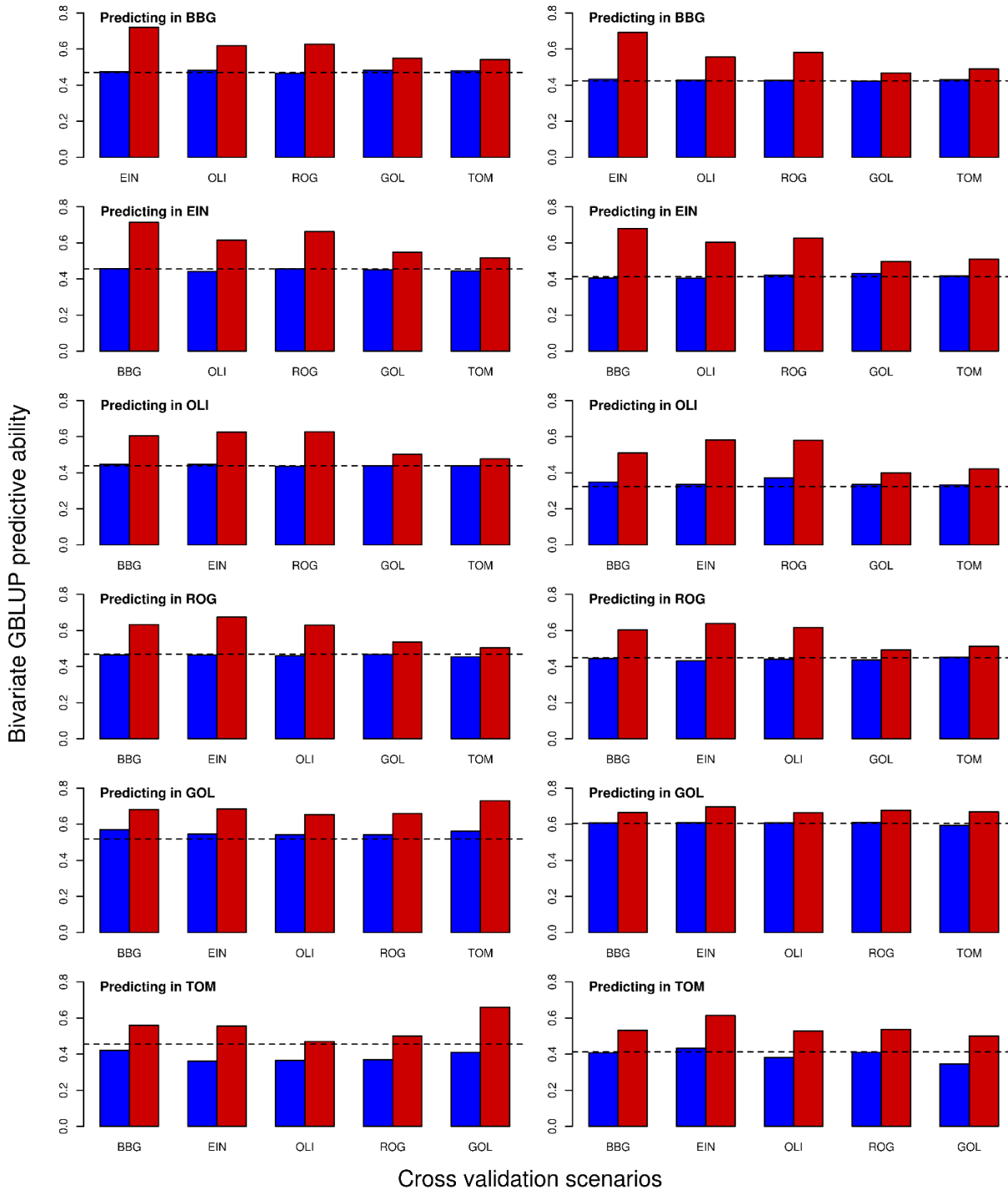


**Fig. 5.4:** Comparison of predictive abilities obtained from univariate GBLUP when the panel of 501'124 SNPs is used (red) and when the pruned set of 30'212 SNPs is used (green) in PE for prediction within environments across series of all traits.

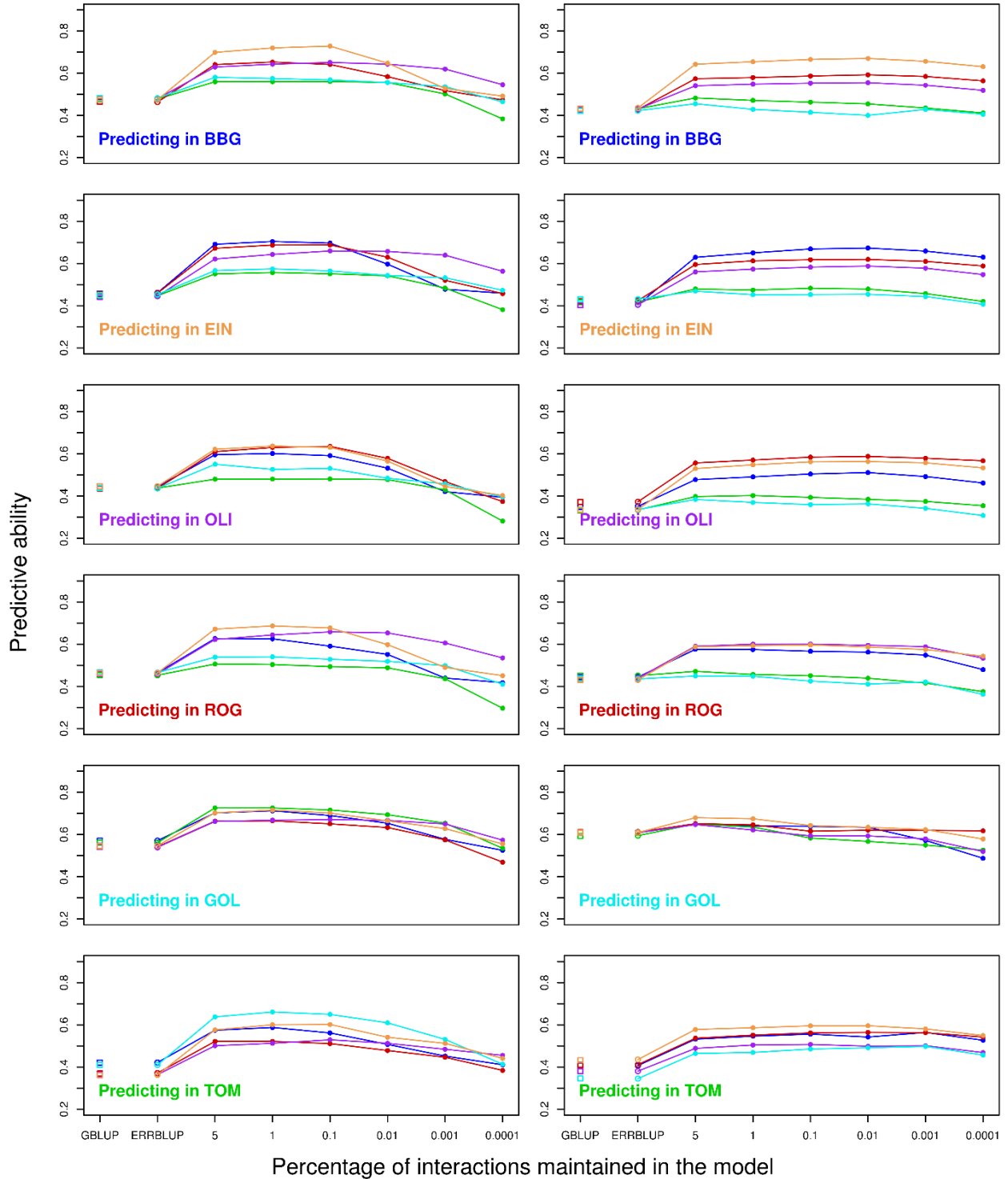
### 5.3 Comparison of cross validation scenarios in bivariate model framework

In this thesis, the cross validation scenario for bivariate model framework was based on training the model using both training sets of the target environment and the full dataset of the additional environment to predict the test set of the target environment. In this scenario, the selection step of bivariate sERRBLUP was done based on the full dataset of the additional environment. In fact, the inclusion of full dataset of the additional environment in this scenario of cross validation has a considerable impact on the predictive ability of bivariate models, although no information from the target environment test set is utilized in the interaction selection step (bivariate sERRBLUP) nor in the prediction step (Bivariate GBLUP, ERRBLUP, sERRBLUP). To shed light on this, Fig. 5.5 illustrates the comparison in bivariate GBLUP predictive ability for the trait PH\_V4 in the year 2017 in both KE and PE under two scenarios of cross validation. The first scenario is based on training the model using both training sets of the target environment and the full dataset of the additional environment which was our scenario in this thesis, whereas the second scenario is based on training the model only using the training set of both target and additional environments. This comparison explicitly demonstrates the superior predictive ability of the first scenario. Therefore, the inclusion of the full dataset of additional environment increases the predictive ability compared to the second scenario even if the additional environment includes half number of lines than the target environment (e.g. GOL and TOM) and it is also significantly higher than univariate GBLUP.

Furthermore, the comparison between bivariate GBLUP, ERRBLUP and sERRBLUP predictive abilities obtained from the second scenario of cross validation by which the models were trained only using the training sets of both environments for the trait PH\_V4 in the year 2017 in both KE and PE are shown in Fig. 5.6 It is shown that bivariate GBLUP and ERRBLUP are almost identical, while bivariate sERRBLUP increases the predictive ability considerably. It should be noted that considering only the training sets of both environments in this cross validation scenario was done at the prediction step, while the interaction selection step of bivariate sERRBLUP was based on the full dataset of the additional environment. Therefore, the observed increase in predictive ability of bivariate sERRBLUP compared to bivariate GBLUP in this scenario is due to both epistasis and borrowing information from the additional environment's test set at the variable selection step. This cross validation scenario results in an average relative increase of 46.5 percent in KE and 42.0 percent in PE from bivariate GBLUP predictive ability to maximum bivariate sERRBLUP predictive ability for the trait PH\_V4. This average relative increase in predictive ability is close to the average relative increase obtained from univariate GBLUP to maximum univariate sERRBLUP (in KE is 49.2 percent and in PE is 44.3 percent) which was shown in Chapter 3. This indicates that the selection of pairwise SNP interactions based on the full dataset of the other environment results in the similar pattern of enhancement in sERRBLUP predictive ability compared to GBLUP predictive ability in both univariate and bivariate models. However, the first cross validation scenario which was utilized in this thesis for bivariate models indicates smaller increases in sERRBLUP predictive ability from GBLUP predictive ability reflecting epistasis only. In general, this scenario of cross validation leads to higher increase in the bivariate model's predictive ability. Therefore, this scenario is recommended when the same lines are phenotyped in different environments or in different growing seasons like multiple years.



**Fig. 5.5:** Comparison of bivariate GBLUP predictive abilities obtained through masking test sets in both target and additional environment (blue bars) and masking only a test set of the target environment (red bars) in cross validation scenario for trait PH\_V4 in 2017 in both KE (left) and PE (right). In each panel, the dashed horizontal line indicates the univariate GBLUP predictive ability for prediction within each environment.



**Fig. 5.6:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH-V4 in 2017 when the test sets in both target and additional environments are masked in cross validation scenario. In each panel, the solid lines' color indicates the additional environment used to predict the target environment.

### **5.3.1 Evaluating cross validation scenario in bivariate models**

Runcie and Cheng (2019) showed that the cross validation in multi-trait models can potentially bias the prediction accuracy. It was illustrated that the cross validation scenario under which the focal trait is predicted by the help of a secondary trait in multi-trait models can be biased when the secondary traits are measured on the individuals which are going to be tested for the focal trait. The arising question here is whether the predictive abilities obtained from our cross validation scenario are biased or not. Although our bivariate models are clustered as the multi-trait models, they have been evaluated for a single trait in two different environments (or years). In fact, the cross validation scenario utilized in this thesis for bivariate models predicted the test set of the focal environment from the training set of the focal environment and the full dataset of the secondary environment for a single phenotypic trait. Therefore, no phenotypic records of the focal environment's test set were used in this scenario of cross validation at the prediction step nor at the interaction selection step. In fact, this scenario benefits from including the phenotypic records of the test set lines in another environment than the focal environment. Therefore, it was shown that considering the environments with fewer recorded lines than the focal environment, as the secondary environment leads to the minimum increase in predictive ability from univariate models' predictive ability. This causes the consideration of all other environments jointly as the secondary environment to be a more convenient approach through which the deficiencies of a single environment are compensated when combined with other environments (Fig. 3.7). Thus, a bias of the bivariate models' prediction accuracies through our cross validation scenario is not expected.

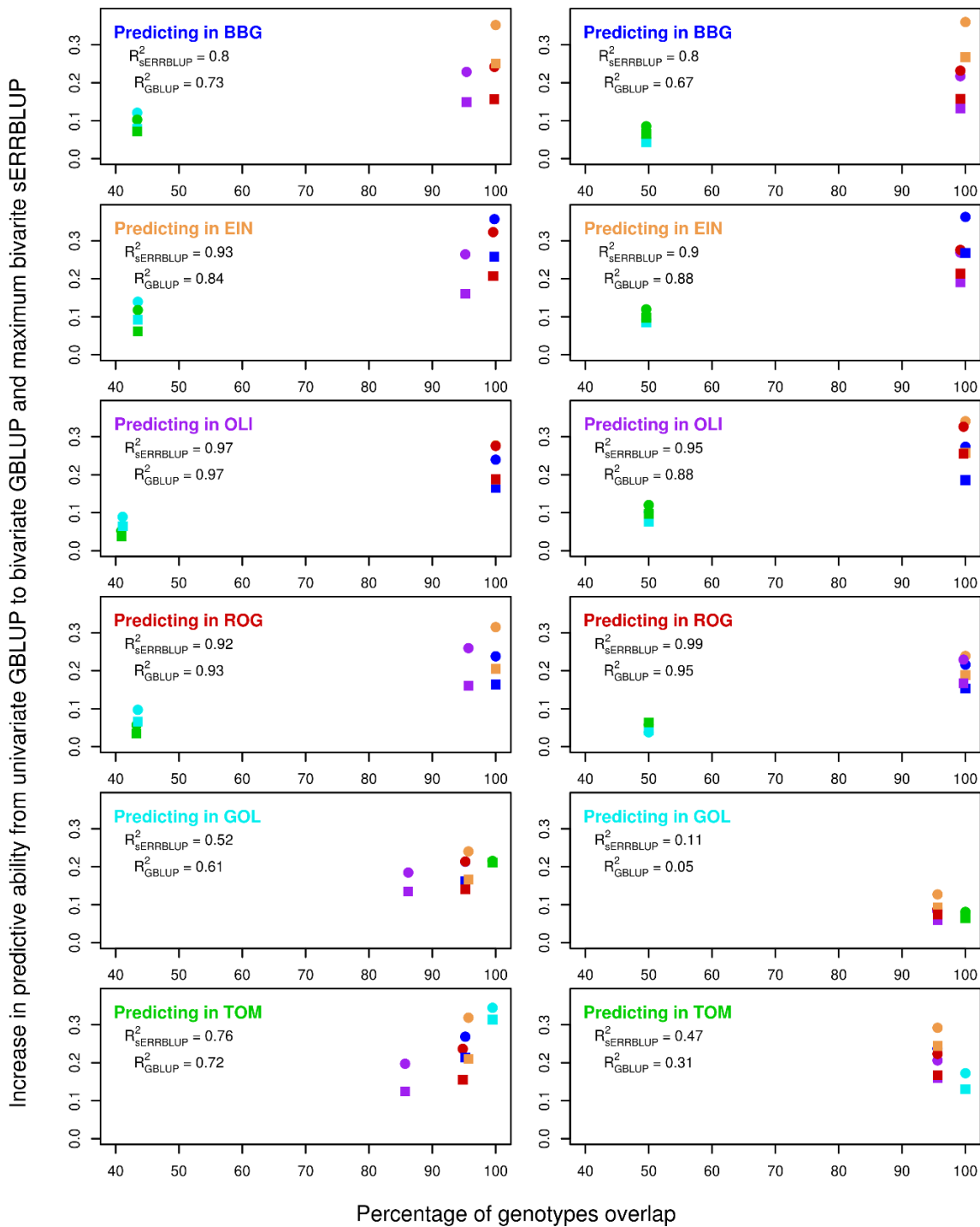
### **5.4 The required genotype overlap in bivariate model framework**

Persa *et al.* (2020) showed that the multi-trait model increases the predictive ability compared to single-trait models in the associated traits for partially overlap of genotypes representing the observation of some traits for the same genotypes in soybeans. This increase was shown to be 6 percent for grain yield and 2 to 9 percent for other traits such as protein content, oil content, plant height, lodging, seed size and fiber. In our study, the number of recorded lines varied between the environments. The exact number of recorded lines for trait PH\_V4 was illustrated in Table 3.2 for the year 2017 and in Table 4.2 for the environments present in both 2017 and 2018. In our bivariate models, the overlap of genotype between the target environment and the additional environment is also an important factor which directly affects the bivariate model's predictive ability. In this regard, the comparison between the overlap of the target environment's and additional environment's genotypes, and the absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP and maximum bivariate sERRBLUP in 2017 for trait PH\_V4 in both KE and PE are illustrated in Fig. 5.7. It is shown that there is a significant correlation between the absolute gain in bivariate GBLUP and maximum bivariate sERRBLUP predictive abilities and the genotype overlap for prediction in all environments of both landraces except for GOL in PE. When combining all environments, there is a significant correlation of 0.84 (p-value = 5.54e-09) in KE and 0.63 (p-value = 0.0001) in PE between the absolute gain in bivariate GBLUP predictive ability

and the percentage of genotype overlap between the target environment and the additional environment. This correlation increases to 0.89 (p-value =  $5.96 \times 10^{-11}$ ) in KE and 0.66 (p-value =  $6.59 \times 10^{-5}$ ) in PE for maximum bivariate sERRBLUP. Therefore, under the assumption of partial genotype overlap such as 50 percent overlap or less, the absolute increase in the respective predictive abilities is less than 0.1 in bivariate GBLUP and less than 0.14 in bivariate sERRBLUP. However, this can increase up to 0.3 in bivariate GBLUP and 0.4 in bivariate sERRBLUP under the assumption of more than 85 percent genotype overlap.

Therefore, the method proposed in Chapter 3 for considering all the other environments jointly as the additional environment in bivariate models leads to 100 percent overlap of genotypes with the target environment. Consequently, this approach leads to the maximum increase in the respective predictive ability which was shown to be as good as or better than the maximum bivariate model predictive ability when considering a single environment as an additional environment (Fig. 3.7).

Moreover, in the context of univariate models, univariate sERRBLUP is also affected by the level of genotype overlap through the variable selection step. As it was illustrated in Chapter 3, the maximum increase in univariate sERRBLUP was obtained when the variable selection step was carried out in the environment with more genotype overlap with the target environment so that TOM and GOL, which provides less than 50 percent genotype overlap with the other four environments (BBG, EIN, OLI, ROG), resulted in the minimum increase for prediction within the four respective environments across all studied traits (Fig. S. 3.10 - Fig. S. 3.11 and Fig. S. 3.3 - Fig. S. 3.9). One may thus conclude that the higher gain in univariate sERRBLUP and bivariate models' predictive abilities are to be expected compared to univariate GBLUP when there is higher level of genotype overlap across the environments. This is also expected for prediction across years in bivariate models for higher level of genotype overlap between the years in each environment.



**Fig. 5.7:** Absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP (squares) and maximum bivariate sERRBLUP (circles) versus the percentage of genotypes overlap between the target environment and the additional environment in 2017 in KE (left) and in PE (right) for trait PH\_V4. In each panel, colored circles and squares represent the additional environment used in bivariate model to predict the target environment with the R-squared representing the correlation coefficient between the respective increase in bivariate GBLUP ( $R^2_{GBLUP}$ ) and bivariate sERRBLUP ( $R^2_{sERRBLUP}$ ) predictive abilities and the percentage of genotype overlap.

### 5.4.1 Genomic and phenotypic correlation impacts on bivariate model's predictive ability under the high level of genotype overlap

The percentage of genotype overlap was shown to be significantly correlated to the absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP and maximum bivariate sERRBLUP in trait PH\_V4 in 2017 based on our cross validation scenario. Fig. 5.7 shows that under the assumption of high genotype overlap such as 100 percent overlap, higher gain in predictive ability is expected. However, one can observe that the respective absolute gain in predictive ability is not the same when the target environment is joined with different environments of the same level of genotype overlap. To shed light on this, the increase in bivariate GBLUP and maximum bivariate sERRBLUP predictive abilities for predicting ROG in KE when it is joined with BBG and EIN are different although both BBG and EIN have 100 percent genotype overlap with ROG. The increase in the respective bivariate GBLUP and maximum bivariate sERRBLUP predictive abilities when ROG is joined with BBG are respectively 0.16 and 0.24, while joining with EIN results in higher increase of 0.20 in bivariate GBLUP and 0.31 in maximum bivariate sERRBLUP predictive abilities. Therefore, an investigation of the underlying factors affecting the respective increase in predictive ability is required.

Genomic correlation was shown to be the most influential factor significantly correlated with the absolute gain in maximum bivariate sERRBLUP predictive ability from univariate GBLUP for prediction across years where we had 100 percent genotype overlap between the years in each environment (Fig. 4.3). This correlation is also significant for prediction across environments for trait PH\_V4 in 2017 in both KE ( $r = 0.93$ ,  $p\text{-value} = 5.78e-14$ ) and PE ( $r = 0.75$ ,  $p\text{-value} = 1.95e-06$ ) in maximum bivariate sERRBLUP and in bivariate GBLUP as well ( $r = 0.72$ ,  $p\text{-value} = 8.22e-06$  in KE and  $r = 0.59$ ,  $p\text{-value} = 0.0006$  in PE). Additionally, phenotypic correlation is another influential factor in bivariate models predictive ability which was shown to be correlated to the increase in bivariate model's predictive ability to some extent (Fig. 4.4). Therefore, the genomic and phenotypic correlation and the absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP and maximum bivariate sERRBLUP for the environments' combinations with more than 85 percent genotype overlap for trait PH\_V4 in both KE and PE are shown in Table 5.1 and Table 5.2, respectively.

Table 5.1 and Table 5.2 illustrate the significant role of genomic correlation in an increase in the respective bivariate models' predictive abilities under the assumption of high level of genotype overlap. Both tables indicate a significant correlation of 0.71 ( $p\text{-value} = 0.0002$ ) and 0.76 ( $p\text{-value} = 3.12e-05$ ) between the absolute gain in bivariate GBLUP predictive ability and the genomic correlation when there is above 85 percent genotype overlap in KE and PE, respectively. This correlation for maximum bivariate sERRBLUP increases to 0.87 ( $p\text{-value} = 1.8e-07$ ) in KE and 0.82 ( $p\text{-value} = 3.75e-06$ ) in PE.

Moreover, it is shown that under the assumption of equal percentage of genotype overlap, the additional environment with higher genomic correlation to the target environment leads to the higher absolute increase in the respective predictive ability. To illustrate this, as it was discussed, predicting ROG in KE when joining with EIN leads to a higher gain in predictive ability (0.20/0.34



in bivariate GBLUP/sERRBLUP) than joining with BBG (0.16/0.20 in bivariate GBLUP/sERRBLUP), since ROG and EIN have higher genomic correlation (0.90/0.93 in bivariate GBLUP/sERRBLUP) than ROG and BBG (0.86/0.85 in bivariate GBLUP/sERRBLUP), although both EIN and BBG have 100 percent genotype overlap with ROG. Therefore, under the assumption of having high genotype overlap, genomic correlation can determine the amount of increase in bivariate model's predictive ability.

Furthermore, Table 5.1 and Table 5.2 illustrate that under the assumption of having equal levels of genotype overlap and equal genomic correlation, the environment with higher phenotypic correlation leads to the higher gain in bivariate model's predictive ability. As an illustration, when predicting TOM in PE by joining with BBG and EIN, there is an equal genotype overlap of 95.6 percent and equal GBLUP genomic correlation of 0.99 between TOM and both BBG and EIN. However, the phenotypic correlation between TOM and BBG is 0.57 where an increase of 0.16 in the bivariate GBLUP predictive ability is observed, whereas the phenotypic correlation between TOM and EIN is 0.68 which leads to the higher increase of 0.24 in bivariate GBLUP predictive ability. Consequently, Table 5.1 and Table 5.2 illustrate the significant correlation of 0.41 in KE (p-value = 0.058) and 0.82 in PE (p-value = 2.77e-06) between the absolute gain in bivariate GBLUP predictive ability and phenotypic correlation which increases to 0.62 in KE (p-value = 0.002) and 0.87 in PE (p-value = 1.89e-07) for maximum bivariate sERRBLUP.

Overall, genomic correlation and phenotypic correlation are shown to be the first and second factors affecting the absolute increase in bivariate model's predictive ability, respectively.

**Table 5.1:** The phenotypic correlation, GBLUP and sERRBLUP genomic correlations of the environments combinations with genotype overlap above 85 percent and the respective absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP and maximum bivariate sERRBLUP for the trait PH\_V4 in KE.

<b>Target environment</b>	<b>Secondary environment</b>	<b>Phenotypic correlation</b>	<b>Percentage of genotype overlap</b>	<b>GBLUP Genomic Correlation</b>	<b>Absolute increase in bivariate GBLUP Predictive ability</b>	<b>sERRBLUP Genomic Correlation</b>	<b>Absolute increase in maximum bivariate sERRBLUP Predictive ability</b>
<b>BBG</b>	<b>EIN</b>	0.82	100	0.99	0.25	0.99	0.35
	<b>ROG</b>	0.67	99.8	0.86	0.16	0.84	0.24
	<b>OLI</b>	0.66	95.4	0.82	0.15	0.74	0.23
<b>EIN</b>	<b>BBG</b>	0.82	99.8	0.99	0.25	1.00	0.36
	<b>ROG</b>	0.77	99.6	0.90	0.16	0.92	0.32
	<b>OLI</b>	0.71	95.2	0.90	0.15	0.90	0.26
<b>OLI</b>	<b>ROG</b>	0.71	100	0.91	0.19	0.91	0.28
	<b>EIN</b>	0.71	100	0.90	0.19	0.90	0.28
	<b>BBG</b>	0.66	100	0.82	0.16	0.81	0.24
<b>ROG</b>	<b>EIN</b>	0.77	100	0.90	0.20	0.93	0.31
	<b>BBG</b>	0.67	100	0.86	0.16	0.85	0.24
	<b>OLI</b>	0.71	95.7	0.91	0.16	0.92	0.26
<b>GOL</b>	<b>TOM</b>	0.69	99.5	0.99	0.21	0.87	0.22
	<b>EIN</b>	0.75	95.7	0.86	0.17	0.82	0.24
	<b>BBG</b>	0.69	95.2	0.82	0.16	0.75	0.21
	<b>ROG</b>	0.70	95.2	0.70	0.14	0.70	0.21
	<b>OLI</b>	0.65	86.2	0.76	0.13	0.66	0.18
<b>TOM</b>	<b>GOL</b>	0.69	99.5	0.99	0.31	0.96	0.34
	<b>EIN</b>	0.66	95.7	0.82	0.21	0.89	0.32
	<b>BBG</b>	0.58	95.2	0.81	0.21	0.80	0.27
	<b>ROG</b>	0.58	94.8	0.74	0.15	0.75	0.24
	<b>OLI</b>	0.50	85.7	0.78	0.12	0.79	0.20

**Table 5.2:** The phenotypic correlation, GBLUP and sERRBLUP genomic correlations of the environments combinations with genotype overlap above 85 percent and the respective absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP and maximum bivariate sERRBLUP for the trait PH\_V4 in PE.

Target environment	Secondary environment	Phenotypic correlation	Percentage of genotype overlap	GBLUP Genomic Correlation	Absolute increase in bivariate GBUP Predictive ability	sERRBLUP Genomic Correlation	Absolute increase in maximum bivariate sERRBLUP Predictive ability
BBG	EIN	0.78	100	0.95	0.27	0.94	0.36
	ROG	0.62	99.2	0.79	0.16	0.84	0.23
	OLI	0.60	99.2	0.60	0.13	0.70	0.22
EIN	BBG	0.78	100	0.95	0.27	0.97	0.36
	ROG	0.69	99.2	0.91	0.16	0.91	0.28
	OLI	0.66	99.2	0.83	0.13	0.83	0.27
OLI	EIN	0.66	100	0.83	0.26	0.86	0.34
	BBG	0.60	100	0.60	0.19	0.73	0.27
	ROG	0.65	99.7	0.97	0.26	0.99	0.33
ROG	EIN	0.69	100	0.91	0.19	0.90	0.24
	BBG	0.62	100	0.79	0.15	0.83	0.22
	OLI	0.65	99.7	0.97	0.17	0.92	0.23
GOL	TOM	0.54	100	0.65	0.06	0.53	0.08
	EIN	0.59	95.6	0.61	0.09	0.60	0.13
	BBG	0.55	95.6	0.52	0.06	0.52	0.08
	ROG	0.51	95.6	0.56	0.07	0.53	0.09
	OLI	0.46	95.6	0.52	0.06	0.43	0.08
TOM	GOL	0.54	100	0.65	0.13	0.95	0.17
	BBG	0.57	95.6	0.99	0.16	0.95	0.24
	EIN	0.68	95.6	0.99	0.24	0.96	0.29
	ROG	0.58	95.6	0.89	0.17	0.97	0.22
	OLI	0.57	95.6	0.86	0.16	0.88	0.21

## 5.5 Potential future research topics

In order to enhance the epistasis genomic prediction model's performance further investigations are required in general. In this regard, incorporating transcriptomic data in to the epistasis genomic prediction model is a potential approach to be tested for evaluating the epistasis model's prediction accuracy. It has been shown that a large scale incorporation of transcriptomic data into the genomic prediction model has the potential to enhance genomic prediction (Li *et al.*, 2019).

Moreover, in the context of a multi-trait model, incorporating weather data into epistasis genomic prediction models can be investigated in the context of the sERRBLUP model, since inclusion of weather data as  $G \times E$  has been shown to increase the genomic prediction model's accuracy (Hallauer *et al.*, 2010; Burgueño *et al.*, 2012; Crossa *et al.*, 2016; Mafouasson *et al.*, 2018; Gillberg *et al.*, 2019; Millet *et al.*, 2019).

Furthermore, detecting epistatic interactions through single-trait and multi-trait GWAS analysis is also a potential subject for future investigations. In the context of plant breeding, GWAS has been positively evaluated for various plants such as maize, wheat, rice, barley (Huang *et al.*, 2010; Weng *et al.*, 2011; M. Wang *et al.*, 2012; Van Inghelandt *et al.*, 2012; Yang *et al.*, 2014; Sukumaran *et al.*, 2015; Alqudah *et al.*, 2020). Epistatic analysis in GWAS studies can help detection of non-significant SNPs which interact with each other and may have a large effect on the trait (Koo *et al.*, 2015). Moreover, multi-trait analysis of GWAS has been shown to increase the number of detected associated loci compared to single-trait GWAS (Korte *et al.*, 2012; Turley *et al.*, 2018; Bonnemaier *et al.*, 2019; Rice *et al.*, 2020). Therefore, investigating epistasis GWAS through single-trait and multi-trait statistical settings is a potential field for future investigation in the context of epistasis models.

## 5.6 Outlooks and Conclusion

### 5.6.1 Outlook for epistasis models

In this work, we have developed and evaluated epistasis models in univariate and bivariate statistical settings for prediction across environments and across years in doubled haploid lines from European landraces Kemater and Petkuser in six locations in the year 2017 and in four locations in both years 2017 and 2018 in Germany and Spain for eight phenotypic traits.

ERRBLUP as the full epistasis model incorporating all pairwise SNP interactions and sERRBLUP as the selective epistasis model incorporating a subset of pairwise SNP interactions have been developed in this thesis. ERRBLUP was shown to be almost identical to GBLUP, since incorporating all pairwise SNP interactions into the genomic prediction model introduced a large number of unimportant variables in to the model producing the 'noise' which prevents a gain in ERRBLUP predictive ability compared to GBLUP. In contrast, sERRBLUP outperforms GBLUP

when the top ranked pairwise SNP interactions maintained in the model in both univariate and bivariate statistical settings across series of phenotypic traits and landraces in maize dataset for the majority of cases. Interaction selection based on effect variances in sERRBLUP was shown to be the preferable selection criterion compared to interaction selection based on absolute effect sizes due its robustness.

The increase in GBLUP predictive ability to maximum sERRBLUP predictive ability results from epistasis in bivariate statistical setting, while this increase is caused by both epistasis and borrowing information across environments in univariate statistical setting.

### **5.6.2 Outlook for EpiGP R-package**

In this thesis, we have found that the proposed sERRBLUP model's computing time display a quadratic growth by increasing the number of SNPs. Although our developed "EpiGP" R-package is able to process more than 30'000 SNPs and thus more than 450 Million SNP x SNP interactions in an efficient way, our proposed model can potentially lead to a considerable computational load by increasing the number of SNPs to hundreds of thousands. Therefore, reducing the number of SNPs by LD pruning is highly recommended which can potentially lead to higher predictive ability in addition to making the application of epistasis models feasible in an efficient computational time. Moreover, utilizing haplotype blocks in sERRBLUP model was shown to result in the very similar prediction accuracies as the ones obtained when utilizing pruned set of SNPs, while its computational load is significantly reduced compared to sERRBLUP model based on pruned set of SNPs.

### **5.6.3 Outlook for influential factors on the model's predictive ability**

Bivariate models were shown to be superior to univariate models under the cross validation scenario through which only the test set of the target environment (or year) is masked in each run of 5-fold cross validation with 5 replicates. With this, bivariate GBLUP was shown to be slightly better than maximum univariate sERRBLUP, and maximum bivariate sERRBLUP obtained the maximum predictive ability among all bivariate models in most cases.

Furthermore, in the context of bivariate models and univariate sERRBLUP model, the genotype overlap between the target environment (or year) and the secondary environment (or year) illustrates significant correlation with the respective models' predictive abilities. Under the assumption of a high level of genotype overlap, genomic correlation was shown to be the first and the most significant factor affecting the genomic prediction model's predictive ability. Phenotypic correlation and the trait's heritability were also the factors affecting the predictive ability of the respective models.

In this thesis, we further proposed genomic prediction models across multiple environments jointly as a successful approach through which the obtained predictive ability is as good as or better than the maximum predictive ability obtained across a single environment. This was potentially caused

by providing 100 percent genotype overlap between the target environment and the secondary environment. Besides, utilizing the average phenotypic values across all environments adjusts the impact of a single environment which helps to prevent the choice of a ‘wrong’ environment for training the model.

#### **5.6.4 Concluding remarks**

Genomic prediction of phenotypes can be a giant step toward increasing the accuracy of genomic prediction models by utilizing the sERRBLUP model especially through a bivariate statistical framework which can potentially increase the predictive ability the most and its computational load can be reduced significantly by utilizing haplotype blocks instead of pruned set of SNPs. The potential gain in bivariate sERRBLUP predictive ability is determined by the level of genotype overlap, genomic and phenotypic correlation between the target environment (or year) and the secondary environment (or year) for a highly heritable trait, when utilizing the full dataset of the secondary environment (or year) and the training set of the target environment (or year) for predicting the target environment’s test set (or year).

In fact, this scenario of genomic prediction of phenotypes in bivariate sERRBLUP model simulates the real scenarios in breeding programs in which the lines are recorded in multiple environments either during the same growing season or subsequent seasons. Moreover, the successful approach considered as prediction across multiple environments jointly is a potential approach in this context which can enhance the predictive ability through training the model on all available information across environments or across multiple growing seasons.

Overall, accurate genomic prediction of phenotypes is of great importance in plant breeding since genomic prediction is becoming a daily tool for selection purposes by plant breeders. Therefore, our proposed sERRBLUP model as a selective epistasis model increasing the prediction accuracy can revolutionize the genomic prediction of phenotypes especially through utilizing all available information across multiple environments or multiple growing seasons jointly to train the model in an efficient computational manner by EpiGp R-package. Moreover, efficiently utilizing all phenotypic records which have been collected at high costs over different environments and/or years is an important capacity of sERRBLUP which substantially enhances the prediction accuracy while decreasing the further costs in plant breeding. In general, the sERRBLUP model can potentially be utilized for other species such as animal and human genomic predictions where epistasis is a relevant gene action.

## 6 References

- Abendroth, L. J., Elmore, R. W., Boyer, M. J. and Marlay, S. K. (2011) ‘Corn Growth and Development’, *PMR 1009. Iowa State University of Science and Technology, Cooperative Extension Service, Ames, Iowa*.
- Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S. and Lawlor, T. J. (2010) ‘Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score’, *Journal of Dairy Science*, 93(2), pp. 743–752. doi: 10.3168/jds.2009-2730.
- Akdemir, D. and Godfrey, O. U. (2015) ‘EMMREML: Fitting Mixed Models with Known Covariance Structures’. Available at: <https://cran.r-project.org/package=EMMREML>.
- Akdemir, D. and Isidro-Sánchez, J. (2019) ‘Design of training populations for selective phenotyping in genomic prediction’, *Scientific Reports*, 9(1446). doi: 10.1038/s41598-018-38081-6.
- Albrecht, T., Wimmer, V., Auinger, H.-J., Erbe, M., Knaak, C., Ouzunova, M., Simianer, H. and Schön, C.-C. (2011) ‘Genome-based prediction of testcross values in maize’, *Theoretical and Applied Genetics*, 123(339), pp. 339–350. doi: 10.1007/s00122-011-1587-7.
- Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A. M., Hu, T. T., Jiang, R., Mulyati, N. W., Zhang, X., Amer, M. A., Baxter, I., Brachi, B., Chory, J., Dean, C., Debieu, M., de Meaux, J., Ecker, J. R., Faure, N., Kniskern, J. M., Jones, J. D. G., Michael, T., Nemri, A., Roux, F., Salt, D. E., Tang, C., Todesco, M., Traw, M. B., Weigel, D., Marjoram, P., Borevitz, J. O., Bergelson, J. and Nordborg, M. (2010) ‘Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines’, *Nature*, 465(7298), pp. 627–631. doi: 10.1038/nature08800.
- Auinger, H. J., Schönleben, M., Lehermeier, C., Schmidt, M., Korzun, V., Geiger, H. H., Piepho, H.-P., Gordillo, A., Wilde, P., Bauer, E. and Schön, C.-C. (2016) ‘Model training across multiple breeding cycles significantly improves genomic prediction accuracy in rye (*Secale cereale* L.)’, *Theoretical and Applied Genetics*, 129(11), pp. 2043–2053. doi: 10.1007/s00122-016-2756-5.
- Aulchenko, Y. S., Ripke, S., Isaacs, A. and van Duijn, C. M. (2007) ‘GenABEL: an R library for genome-wide association analysis’, *Bioinformatics*, 23(10), pp. 1294–1296. doi: 10.1093/bioinformatics/btm108.
- Aylor, D. L., Valdar, W., Foulds-Mathes, W., Buus, R. J., Verdugo, R. A., Baric, R. S., Ferris, M. T., Frelinger, J. A., Heise, M., Frieman, M. B., Gralinski, L. E., Bell, T. A., Didion, J. D., Hua, K., Nehrenberg, D. L., Powell, C. L., Steigerwalt, J., Xie, Y., Kelada, S. N. P., Collins, F. S., Yang, I. V., Schwartz, D. A., Branstetter, L. A., Chesler, E. J., Miller, D. R., Spence, J., Liu, E. Y., McMillan, L., Sarkar, A., Wang, J., Wang, W., Zhang, Q., Broman, K. W., Korstanje, R., Durrant, C., Mott, R., Iraqi, F. A., Pomp, D., Threadgill, D., Pardo-Manuel de Villena, F. and Churchill, G. A. (2011) ‘Genetic analysis of complex traits in the emerging Collaborative Cross’, *Genome Research*, 21(8), pp. 1213–1222. doi: 10.1101/gr.111310.110.
- Bajgain, P., Zhang, X. and Anderson, J. A. (2020) ‘Dominance and G×E interaction effects improve genomic prediction and genetic gain in intermediate wheatgrass (*Thinopyrum intermedium*)’, *The Plant Genome*, 13(1), p. e20012. doi: 10.1002/tpg2.20012.
- Barrett, J. C., Fry, B., Maller, J. and Daly, M. J. (2004) ‘Haploview: analysis and visualization of LD and haplotype maps’, *Bioinformatics*, 21(2), pp. 263–265. doi: 10.1093/bioinformatics/bth457.
- Bateson, W. (1909) *Mendel’s Principles of Heredity*. New York: Cambridge Univ. Press.

- Bernal-Vasquez, A. M., Möhring, J., Schmidt, M., Schönleben, M., Schön, C.-C. and Piepho, H.-P. (2014) 'The importance of phenotypic data analysis for genomic prediction - a case study comparing different spatial models in rye', *BMC Genomics*, 15(646). doi: 10.1186/1471-2164-15-646.
- Bernardo, R. and Yu, J. (2007) 'Prospects for Genomewide Selection for Quantitative Traits in Maize', *Crop Science*, 47(3), pp. 1082–1090. doi: 10.2135/cropsci2006.11.0690.
- Bielinski, S. J., Olson, J. E., Pathak, J., Weinshilboum, R. M., Wang, L., Lyke, K. J., Ryu, E., Targonski, P. V., Van Norstrand, M. D., Hathcock, M. A., Takahashi, P. Y., McCormick, J. B., Johnson, K. J., Maschke, K. J., Rohrer Vitek, C. R., Ellingson, M. S., Wieben, E. D., Farrugia, G., Morrisette, J. A., Kruckeberg, K. J., Bruflatt, J. K., Peterson, L. M., Blommel, J. H., Skierka, J. M., Ferber, M. J., Black, J. L., Baudhuin, L. M., Klee, E. W., Ross, J. L., Veldhuizen, T. L., Schultz, C. G., Caraballo, P. J., Freimuth, R. R., Chute, C. G. and Kullo, I. J. (2014) 'Preemptive genotyping for personalized medicine: design of the right drug, right dose, right time-using genomic data to individualize treatment protocol', *Mayo Clinic proceedings*, 89(1), pp. 25–33. doi: 10.1016/j.mayocp.2013.10.021.
- Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. and Kruglyak, L. (2013) 'Finding the sources of missing heritability in a yeast cross', *Nature*, 494(7436), pp. 234–237. doi: 10.1038/nature11867.
- Bloom, J. S., Kotenko, I., Sadhu, M. J., Treusch, S., Albert, F. W. and Kruglyak, L. (2015) 'Genetic interactions contribute less than additive effects to quantitative trait variation in yeast', *Nature Communications*, 6(8712). doi: 10.1038/ncomms9712.
- Brachi, B., Morris, G. P. and Borevitz, J. O. (2011) 'Genome-wide association studies in plants: the missing heritability is in the field', *Genome Biology*, 12(232). doi: 10.1186/gb-2011-12-10-232.
- Brown, A. A., Buil, A., Viñuela, A., Lappalainen, T., Zheng, H.-F., Richards, J. B., Small, K. S., Spector, T. D., Dermitzakis, E. T. and Durbin, R. (2014) 'Genetic interactions affecting human gene expression identified by variance association mapping', *eLife*, 3. doi: 10.7554/eLife.01381.001.
- Browning, S. R. and Browning, B. L. (2007) 'Rapid and accurate haplotype phasing and missing data inference for whole genome association studies by use of localized haplotype clustering', *Am J Hum Genet*, 81(5), pp. 1084–1097. doi: 10.1086/521987.
- Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., Ersoz, E., Flint-Garcia, S., Garcia, A., Glaubitz, J. C., Goodman, M. M., Harjes, C., Guill, K., Kroon, D. E., Larsson, S., Lepak, N. K., Li, H., Mitchell, S. E., Pressoir, G., Peiffer, J. A., Rosas, M. O., Rocheford, T. R., Romay, M. C., Romero, S., Salvo, S., Villeda, H. S., Sofia da Silva, H., Sun, Q., Tian, F., Upadaya, N., Ware, D., Yates, H., Yu, J., Zhang, Z., Kresovich, S. and McMullen, M. D. (2009) 'The Genetic Architecture of Maize Flowering Time', *Science*, 325(5941), pp. 714–718. doi: 10.1126/science.1174276.
- Burgueño, J., Campos, G. de los, Weigel, K. and Crossa, J. (2012) 'Genomic Prediction of Breeding Values when Modeling Genotype × Environment Interaction using Pedigree and Dense Molecular Markers', *Crop Science*, 52(2), pp. 707–719. doi: 10.2135/cropsci2011.06.0299.
- Burke, W. and Psaty, B. M. (2007) 'Personalized Medicine in the Era of Genomics', *JAMA*, 298(14), pp. 1682–1684. doi: 10.1001/jama.298.14.1682.
- Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J. and Thompson, R. (2018) 'ASReml-R Reference Manual Version 4'. VSN International Ltd., Hemel Hempstead.
- Calus, M. P. L. (2010) 'Genomic breeding value prediction: methods and procedures\*', *Animal*, 4(2), pp. 157–164. doi: 10.1017/S1751731109991352.



- Calus, M. P. L. and Vandenplas, J. (2018) ‘SNPrune: an efficient algorithm to prune large SNP array and sequence datasets based on high linkage disequilibrium’, *Genetics Selection Evolution*, 50(34). doi:10.1186/s12711-018-0404-z.
- Calus, M. P. L. and Veerkamp, R. F. (2011) ‘Accuracy of multi-trait genomic selection using different methods’, *Genetics Selection Evolution*, 43(26). doi: 10.1186/1297-9686-43-26.
- Carlborg, Ö. and Haley, C. S. (2004) ‘Epistasis: too often neglected in complex trait studies?’, *Nature Reviews Genetics*, 5, pp. 618–625. doi: 10.1038/nrg1407.
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M. and Lee, J. J. (2015) ‘Second-generation PLINK: rising to the challenge of larger and richer datasets’, *Gigascience*, 4(7). doi: 10.1186/s13742-015-0047-8.
- Cheverud, J. M. and Routman, E. J. (1995) ‘Epistasis and its contribution to genetic variance components’, *Genetics*, 139(3), pp. 1455–1461.
- Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., Shaw, P., Bayer, M., Thomas, W., Marshall, D., Hedley, P., Tondelli, A., Pecchioni, N., Francia, E., Korzun, V., Walther, A. and Waugh, R. (2012) ‘Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley’, *Nature Genetics*, 44(12), pp. 1388–1392. doi: 10.1038/ng.2447.
- Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., Vorsa, N., Iorizzo, M. and Zalapa, J. (2018) ‘Multivariate GBLUP Improves Accuracy of Genomic Selection for Yield and Fruit Weight in Biparental Populations of *Vaccinium macrocarpon* Ait’, *Frontiers in Plant Science*, 9(1310). doi: 10.3389/fpls.2018.01310.
- Crossa, J., Campos, G. de los, Pérez, P., Gianola, D., Burgueño, J., Araus, J. L., Makumbi, D., Singh, R. P., Dreisigacker, S., Yan, J., Arief, V., Banziger, M. and Braun, H.-J. (2010) ‘Prediction of Genetic Values of Quantitative Traits in Plant Breeding Using Pedigree and Molecular Markers’, *Genetics*, 186(2), pp. 713–724. doi: 10.1534/genetics.110.118521.
- Crossa, J., Pérez, P., Campos, G. de los, Mahuku, G., Dreisigacker, S. and Magorokosho, C. (2011) ‘Genomic Selection and Prediction in Plant Breeding’, *Crop Improvement*, 25(3), pp. 239–261. doi: 10.1080/15427528.2011.558767.
- Crossa, J., Jarquín, D., Franco, J., Pérez-Rodríguez, P., Burgueño, J., Saint-Pierre, C., Vikram, P., Sansaloni, C., Petrolí, C., Akdemir, D., Sneller, C., Reynolds, M., Tattaris, M., Payne, T., Guzman, C., Peña, R. J., Wenzl, P. and Singh, S. (2016) ‘Genomic Prediction of Gene Bank Wheat Landraces’, *G3: Genes/Genomes/Genetics*, 6(7), pp. 1819–1834. doi: 10.1534/g3.116.029637.
- Da, Y., Wang, C., Wang, S. and Hu, G. (2014) ‘Mixed Model Methods for Genomic Prediction and Variance Component Estimation of Additive and Dominance Effects Using SNP Markers’, *PLOS ONE*, 9(1). doi: 10.1371/journal.pone.0087666.
- Daetwyler, H. D., Villanueva, B., Bijma, P. and Woolliams, J. A. (2007) ‘Inbreeding in genome-wide selection’, *Journal of Animal Breeding and Genetics*, 124(6), pp. 369–376. doi: 10.1111/j.1439-0388.2007.00693.x.
- Daetwyler, H. D., Pong-Wong, R., Villanueva, B. and Woolliams, J. A. (2010) ‘The Impact of Genetic Architecture on Genome-Wide Evaluation Methods’, *Genetics*, 185(3), pp. 1021–1031. doi: 10.1534/genetics.110.116855.

- Daetwyler, H. D., Calus, M. P. L., Pong-Wong, R., Campos, G. de los and Hickey, J. M. (2013) ‘Genomic Prediction in Animals and Plants: Simulation of Data, Validation, Reporting, and Benchmarking’, *Genetics*, 193(2), pp. 347–365. doi: 10.1534/genetics.112.147983.
- Daetwyler, H. D., Villanueva, B. and Woolliams, J. A. (2008) ‘Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach’, *PLOS ONE*. Public Library of Science, 3(10), p. e3395. doi: 10.1371/journal.pone.0003395.
- Dekkers, J. C. M. (2007) ‘Prediction of response to marker-assisted and genomic selection using selection index theory’, *Journal of Animal Breeding and Genetics*, 124(6), pp. 331–341. doi: 10.1111/j.1439-0388.2007.00701.x.
- Desta, Z. A. and Ortiz, R. (2014) ‘Genomic selection: genome-wide prediction in plant improvement’, *Trends in Plant Science*, 19(9), pp. 592–601. doi: 10.1016/j.tplants.2014.05.006.
- Edwards, S. M., Buntjer, J. B., Jackson, R., Bentley, A. R., Lage, J., Byrne, E., Burt, C., Jack, P., Berry, S., Flatman, E., Poupard, B., Smith, S., Hayes, C., Gaynor, R. C., Gorjanc, G., Howell, P., Ober, E., Mackay, I. J. and Hickey, J. M. (2019) ‘The effects of training population design on genomic prediction accuracy in wheat’, *Theoretical and Applied Genetics*, 132(7), pp. 1943–1952. doi: 10.1007/s00122-019-03327-y.
- Elias, A. A., Robbins, K. R., Doerge, R. W. and Tuinstra, M. R. (2016) ‘Half a Century of Studying Genotype  $\times$  Environment Interactions in Plant Breeding Experiments’, *Crop Science*, 56(5), pp. 2090–2105. doi: 10.2135/cropsci2015.01.0061.
- Endelman, J. B. (2011) ‘Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP’, *The Plant Genome*, 4(3), pp. 250–255. doi: 10.3835/plantgenome2011.08.0024.
- Endelman, J. B. and Jannink, J.-L. (2012) ‘Shrinkage Estimation of the Realized Relationship Matrix’, *G3: Genes/Genomes/Genetics*, 2(11), pp. 1405–1413. doi: 10.1534/g3.112.004259.
- Erbe, M., Pimentel, E., Sharifi, A. R. and Simianer, H. (2010) ‘Assessment of cross-validation strategies for genomic prediction in cattle’, *9th World Congress of Genetics Applied to Livestock Production*, pp. 129–132.
- Falconer, D. S. and Mackay, T. F. C. (1996) *Introduction to Quantitative Genetics*. England: Longman. Essex Engl.
- Filho, J. de A., Guimarães, J., Silva, F. e, Resende, M. de, Muñoz, P., Kirst, M. and Jr, M. R. (2016) ‘The contribution of dominance to phenotype prediction in a pine breeding and simulated population’, *Heredity (Edinb)*, 117(1), pp. 33–41. doi: 10.1038/hdy.2016.23.
- Fisher, R. A. (1930) *The Genetical Theory of Natural Selection*. Oxford, England: Clarendon Press.
- Flint, J. and Mackay, T. F. C. (2009) ‘Genetic architecture of quantitative traits in mice, flies, and humans’, *Genome Research*, 19(5), pp. 723–733.
- Forsberg, S. K. G., Bloom, J. S., Sadhu, M. J., Kruglyak, L. and Carlborg, Ö. (2017) ‘Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast’, *Nature Genetics*, 49(4), pp. 497–503. doi: 10.1038/ng.3800.
- Fragomeni, B. de O., Misztal, I., Lourenco, D. L., Aguilar, I., Okimoto, R. and Muir, W. M. (2014) ‘Changes in variance explained by top SNP windows over generations for three traits in broiler chicken’, *Frontiers in Genetics*, 5(332). doi: 10.3389/fgene.2014.00332.

- Ganal, M. W., Durstewitz, G., Polley, A., Bérard, A., Buckler, E. S., Charcosset, A., Clarke, J. D., Graner, E.-M., Hansen, M., Joets, J., Le Paslier, M.-C., McMullen, M. D., Montalent, P., Rose, M., Schön, C.-C., Sun, Q., Walter, H., Martin, O. C. and Falque, M. (2011) ‘A Large Maize (*Zea mays* L.) SNP Genotyping Array: Development and Germplasm Genotyping, and Genetic Mapping to Compare with the B73 Reference Genome’, *PLOS ONE*. Public Library of Science, 6(12), p. e28334. doi: 10.1371/journal.pone.0028334.
- Gianola, D., de los Campos, G., Hill, W. G., Manfredi, E. and Fernando, R. (2009) ‘Additive Genetic Variability and the Bayesian Alphabet’, *Genetics*, 183(1), pp. 347–363. doi: 10.1534/genetics.109.103952.
- Gianola, D. (2013) ‘Priors in Whole-Genome Regression: The Bayesian Alphabet Returns’, *Genetics*, 194(3), pp. 573–596. doi: 10.1534/genetics.113.151753.
- Gianola, D., Fernando, R. L. and Stella, A. (2006) ‘Genomic-Assisted Prediction of Genetic Value With Semiparametric Procedures’, *Genetics*, 173(3), pp. 1761–1776. doi: 10.1534/genetics.105.049510.
- Gianola, D. and van Kaam, J. B. C. H. M. (2008) ‘Reproducing Kernel Hilbert Spaces Regression Methods for Genomic Assisted Prediction of Quantitative Traits’, *Genetics*, 178(4), pp. 2289–2303. doi: 10.1534/genetics.107.084285.
- Gianola, D., Perez-Enciso, M. and Toro, M. A. (2003) ‘On Marker-Assisted Prediction of Genetic Value: Beyond the Ridge’, *Genetics*, 163(1), pp. 347–365. doi: 10.1093/genetics/163.1.347.
- Goddard, M. E., Wray, N. R., Verbyla, K. and Visscher, P. M. (2009) ‘Estimating Effects and Making Predictions from Genome-Wide Marker Data’, *Statistical Science*. Institute of Mathematical Statistics, 24(4), pp. 517–529. doi: 10.1214/09-STS306.
- Goddard, M. E. and Hayes, B. J. (2009) ‘Mapping genes for complex traits in domestic animals and their use in breeding programmes’, *Nature Reviews Genetics*, 10(6), pp. 381–391. doi: 10.1038/nrg2575.
- Guo, G., Zhao, F., Wang, Y., Zhang, Y., Du, L. and Su, G. (2014) ‘Comparison of single-trait and multiple-trait genomic prediction models’, *BMC Genetics*, 15(1), p. 30. doi: 10.1186/1471-2156-15-30.
- Habier, D., Tetens, J., Seefried, F.-R., Lichtner, P. and Thaller, G. (2010) ‘The impact of genetic relationship information on genomic breeding values in German Holstein cattle’, *Genetics Selection Evolution*, 42(5). doi: 10.1186/1297-9686-42-5.
- Habier, D., Fernando, R. L., Kizilkaya, K. and Garrick, D. J. (2011) ‘Extension of the bayesian alphabet for genomic selection’, *BMC Bioinformatics*, 12(186). doi: 10.1186/1471-2105-12-186.
- Habier, D., Fernando, R. L. and Dekkers, J. C. M. (2007) ‘The impact of genetic relationship information on genome-assisted breeding values’, *Genetics*. Genetics Society of America, 177(4), pp. 2389–2397. doi: 10.1534/genetics.107.081190.
- Habier, D., Fernando, R. L. and Garrick, D. J. (2013) ‘Genomic BLUP Decoded: A Look into the Black Box of Genomic Prediction’, *Genetics*, 194(3), pp. 597–607. doi: 10.1534/genetics.113.152207.
- Hallauer, A. R., Carena, M. J. and Miranda Filho, J. B. (2010) *Quantitative genetics in maize breeding*, Springer. Berlin.
- Hammer, G., Messina, C., Wu, A. and Cooper, M. (2019) ‘Biological reality and parsimony in crop models—why we need both in crop improvement!’, *in silico Plants*, 1(1). doi: 10.1093/insilicoplants/diz010.

- Hassen, M. Ben, Bartholomé, J., Valè, G., Cao, T.-V. and Ahmadi, N. (2018) ‘Genomic Prediction Accounting for Genotype by Environment Interaction Offers an Effective Framework for Breeding Simultaneously for Adaptation to an Abiotic Stress and Performance Under Normal Cropping Conditions in Rice’, *G3 (Bethesda, Md.)*, 8(7), pp. 2319–2332. doi: 10.1534/g3.118.200098.
- Hayes, B., Bowman, P., Chamberlain, A., Verbyla, K. and Goddard, M. (2009) ‘Accuracy of genomic breeding values in multi-breed dairy cattle populations’, *Genetics Selection Evolution*, 41(51). doi: 10.1186/1297-9686-41-51.
- Hayes, B. and Goddard, M. (2008) ‘Technical note: Prediction of breeding values using marker-derived relationship matrices’, *Journal of Animal Science*, 86(9), pp. 2089–2092. doi: 10.2527/jas.2007-0733.
- Hayes, B. and Goddard, M. (2010) ‘Genome-wide association and genomic selection in animal breeding’, *Genome*. NRC Research Press, 53(11), pp. 876–883. doi: 10.1139/G10-076.
- Hayes, B., Visscher, P. and Goddard, M. (2009) ‘Increased accuracy of artificial selection by using the realized relationship matrix’, *Genetics Research*, 91(1), pp. 47–60. doi: 10.1017/S0016672308009981.
- He, D., Kuhn, D. and Parida, L. (2016) ‘Novel applications of multitask learning and multiple output regression to multiple genetic trait prediction’, *Bioinformatics*, 32(12), pp. i37–i43. doi: 10.1093/bioinformatics/btw249.
- He, D. and Parida, L. (2016) ‘Does encoding matter? A novel view on the quantitative genetic trait prediction problem’, *BMC Bioinformatics*, 17(272). doi: 10.1186/s12859-016-1127-1.
- He, D., Wang, Z. and Parida, L. (2015) ‘Data-driven encoding for quantitative genetic trait prediction’, *BMC Bioinformatics*, 16(S10). doi: 10.1186/1471-2105-16-S1-S10.
- Hemani, G., Knott, S. and Haley, C. (2013) ‘An evolutionary perspective on epistasis and the missing heritability.’, *PLoS genetics*, 9(2), p. e1003295. doi: 10.1371/journal.pgen.1003295.
- Henderson, C. R., Kempthorne, O., Searle, S. R. and von Krosigk, C. M. (1959) ‘The Estimation of Environmental and Genetic Trends from Records Subject to Culling’, *Biometrics*, 15(2), pp. 192–218. doi: 10.2307/2527669.
- Henderson, C. R. (1975) ‘Best Linear Unbiased Estimation and Prediction under a Selection Model’, *Biometrics*, 31(2), pp. 423–447. doi: 10.2307/2529430.
- Henderson, C. R. (1984) *Applications of linear models in animal breeding*. Guelph, Ontario, Canada: University of Guelph.
- Henderson, C. R. (1985) ‘Best Linear Unbiased Prediction of Nonadditive Genetic Merits in Noninbred Populations’, *Journal of Animal Science*, 60(1), pp. 111–117. doi: 10.2527/jas1985.601111x.
- Henderson, C. R. and Quaas, R. L. (1976) ‘Multiple Trait Evaluation Using Relatives’ Records’, *Journal of Animal Science*, 43(6), pp. 1188–1197. doi: 10.2527/jas1976.4361188x.
- Hickey, J. M., Chiurugwi, T., Mackay, I., Powell, W., Hickey, J. M., Chiurugwi, T., Mackay, I., Powell, W., Eggen, A., Kilian, A., Jones, C., Canales, C., Grattapaglia, D., Bassi, F., Atlin, G., Gorjanc, G., Dawson, I., Rabbi, I., Ribaut, J.-M., Rutkoski, J., Benzie, J., Lightner, J., Mwacharo, J., Parmentier, J., Robbins, K., Skot, L., Wolfe, M., Rouard, M., Clark, M., Amer, P., Gardiner, P., Hendre, P., Mrode, R., Sivasankar, S., Rasmussen, S., Groh, S., Jackson, V., Thomas, W., Beyene, Y. and Participants, I. G. S. in C. B. P. W. (2017) ‘Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery’, *Nature Genetics*, 49(9), pp. 1297–1303. doi: 10.1038/ng.3920.

- Hill, W. G. (2010) ‘Understanding and using quantitative genetic variation’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537), pp. 73–85. doi: 10.1098/rstb.2009.0203.
- Hill, W. G., Goddard, M. E. and Visscher, P. M. (2008) ‘Data and theory point to mainly additive genetic variance for complex traits’, *PLoS Genetics*, 4(2). doi: 10.1371/journal.pgen.1000008.
- Hölker, A. C., Mayer, M., Presterl, T., Bolduan, T., Bauer, E., Ordas, B., Brauner, P. C., Ouzunova, M., Melchinger, A. E. and Schön, C.-C. (2019) ‘European maize landraces made accessible for plant breeding and genome-based studies’, *Theoretical and Applied Genetics*, 132(12), pp. 3333–3345. doi: 10.1007/s00122-019-03428-8.
- Hu, Z., Li, Y., Song, X., Han, Y., Cai, X., Xu, S. and Li, W. (2011) ‘Genomic value prediction for quantitative traits under the epistatic model’, *BMC Genet*, 12(15). doi: 10.1186/1471-2156-12-15.
- Huang, A., Xu, S. and Cai, X. (2014) ‘Whole-genome quantitative trait locus mapping reveals major role of epistasis on yield of rice’, *PloS one*. Public Library of Science, 9(1), p. e87330. doi: 10.1371/journal.pone.0087330.
- Huang, W., Richards, S., Carbone, M. A., Zhu, D., Anholt, R. R. H., Ayroles, J. F., Duncan, L., Jordan, K. W., Lawrence, F., Magwire, M. M., Warner, C. B., Blankenburg, K., Han, Y., Javaid, M., Jayaseelan, J., Jhangiani, S. N., Muzny, D., Onger, F., Perales, L., Wu, Y.-Q., Zhang, Y., Zou, X., Stone, E. A., Gibbs, R. A. and Mackay, T. F. C. (2012) ‘Epistasis dominates the genetic architecture of *Drosophila* quantitative traits’, *Proceedings of the National Academy of Sciences of the United States of America*, 109(39), pp. 15553–15559. doi: 10.1073/pnas.1213423109.
- Huq, A., Akter, S., Nou, I. S., Kim, H. T., Jung, Y. J. and Kang, K. K. (2016) ‘Identification of functional SNPs in genes and their effects on plant phenotypes’, *Journal of Plant Biotechnology*. 2016/03/31. The Korean Society of Plant Biotechnology, 43(1), pp. 1–11. doi: 10.5010/JPB.2016.43.1.1.
- Jarquín, D., Howard, R., Crossa, J., Beyene, Y., Gowda, M., Martini, J. W. R., Covarrubias Pazarán, G., Burgueño, J., Pacheco, A., Grondona, M., Wimmer, V. and Prasanna, B. M. (2020) ‘Genomic Prediction Enhanced Sparse Testing for Multi-environment Trials’, *G3: Genes/Genomes/Genetics*, 10(8), pp. 2725–2739. doi: 10.1534/g3.120.401349.
- Jia, Y. and Jannink, J.-L. (2012) ‘Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy’, *Genetics*, 192(4), pp. 1513–1522. doi: 10.1534/genetics.112.144246.
- Jiang, J., Zhang, Q., Ma, L., Li, J., Wang, Z. and Liu, J. (2015) ‘Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model’, *Heredity*, 115(1), pp. 29–36. doi: 10.1038/hdy.2015.9.
- Jiang, Y. and Reif, J. C. (2015) ‘Modeling Epistasis in Genomic Selection’, *Genetics*, 201(2), pp. 759–768. doi: 10.1534/genetics.115.177907.
- Jiang, Y. and Reif, J. C. (2020) ‘Efficient Algorithms for Calculating Epistatic Genomic Relationship Matrices’, *Genetics*, 216(3), pp. 651–669. doi: 10.1534/genetics.120.303459.
- Jiang, Y., Schmidt, R. H. and Reif, J. C. (2018) ‘Haplotype-Based Genome-Wide Prediction Models Exploit Local Epistatic Interactions Among Markers’, *G3: Genes/Genomes/Genetics*, 8(5), pp. 1687–1699. doi: 10.1534/g3.117.300548.
- Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., Campbell, M. S., Stein, J. C., Wei, X., Chin, C.-S., Guill, K., Regulski, M., Kumari, S., Olson, A., Gent, J., Schneider, K. L., Wolfgruber, T. K., May, M. R., Springer, N. M., Antoniou, E., McCombie, W. R., Presting, G. G., McMullen, M., Ross-Ibarra, J., Dawe, R. K., Hastie, A., Rank, D. R. and Ware, D. (2017) ‘Improved maize reference genome with single-molecule technologies’, *Nature*, 546, pp. 524–527. doi: 10.1038/nature22971.

- Jones, B. (2012) 'Predicting phenotypes', *Nature Reviews Genetics*, 13(450). doi: 10.1038/nrg3267.
- Kang, M. S. and Gorman, D. P. (1989) 'Genotype  $\times$  Environment Interaction in Maize', *Agronomy Journal*, 81(4), pp. 662–664. doi: 10.2134/agronj1989.00021962008100040020x.
- Karaman, E., Cheng, H., Firat, M. Z., Garrick, D. J. and Fernando, R. L. (2016) 'An Upper Bound for Accuracy of Prediction Using GBLUP', *PLOS ONE*, 11(8), p. e0161054. doi: 10.1371/journal.pone.0161054.
- Koning, N. B. J., Van Ittersum, M. K., Becx, G. A., Van Boekel, M. A. J. S., Brandenburg, W. A., Van Den Broek, J. A., Goudriaan, J., Van Hofwegen, G., Jongeneel, R. A., Schiere, J. B. and Smies, M. (2008) 'Long-term global availability of food: continued abundance or new scarcity?', *NJAS - Wageningen Journal of Life Sciences*, 55(3), pp. 229–292. doi: 10.1016/S1573-5214(08)80001-2.
- Kranis, A., Gheyas, A. A., Boschiero, C., Turner, F., Yu, L., Smith, S., Talbot, R., Pirani, A., Brew, F., Kaiser, P., Hocking, P. M., Fife, M., Salmon, N., Fulton, J., Strom, T. M., Haberer, G., Weigend, S., Preisinger, R., Gholami, M., Qanbari, S., Simianer, H., Watson, K. A., Woolliams, J. A. and Burt, D. W. (2013) 'Development of a high density 600K SNP genotyping array for chicken', *BMC Genomics*, 14(59). doi: 10.1186/1471-2164-14-59.
- Lee, S. H., van der Werf, J. H. J., Hayes, B. J., Goddard, M. E. and Visscher, P. M. (2008) 'Predicting Unobserved Phenotypes for Complex Traits from Whole-Genome SNP Data', *PLOS Genetics*. Public Library of Science, 4(10), p. e1000231. doi: 10.1371/journal.pgen.1000231.
- Lee, S. H. and van der Werf, J. H. J. (2016) 'MTG2: an efficient algorithm for multivariate linear mixed model analysis based on genomic information', *Bioinformatics*, 32(9), pp. 1420–1422. doi: 10.1093/bioinformatics/btw012.
- Legarra, A., Christensen, O. F., Aguilar, I. and Misztal, I. (2014) 'Single Step, a general approach for genomic selection', *Livestock Science*, 166, pp. 54–65. doi: 10.1016/j.livsci.2014.04.029.
- Lello, L., Raben, T. G., Yong, S. Y., Tellier, L. C. A. M. and Hsu, S. D. H. (2019) 'Genomic Prediction of Complex Disease Risk', *bioRxiv*, p. 506600. doi: 10.1101/506600.
- de Leon, N., Jannink, J.-L., Edwards, J. W. and Kaeppler, S. M. (2016) 'Introduction to a Special Issue on Genotype by Environment Interaction', *Crop Science*, 56(5), pp. 2081–2089. doi: 10.2135/cropsci2016.07.0002in.
- Liu, Y., Wang, D., He, F., Wang, J., Joshi, T. and Xu, D. (2019) 'Phenotype Prediction and Genome-Wide Association Study Using Deep Convolutional Neural Network of Soybean', *Frontiers in Genetics*, 10(1091). doi: 10.3389/fgene.2019.01091.
- Lorenzana, R. E. and Bernardo, R. (2009) 'Accuracy of genotypic value predictions for marker-based selection in biparental plant populations', *Theoretical and Applied Genetics*, 120(1), pp. 151–161. doi: 10.1007/s00122-009-1166-3.
- de los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., Weigel, K. and Cotes, J. M. (2009) 'Predicting Quantitative Traits With Regression Models for Dense Molecular Markers and Pedigree', *Genetics*, 182(1), pp. 375–385. doi: 10.1534/genetics.109.101501.
- de los Campos, G., Vazquez, A. I., Fernando, R., Klimentidis, Y. C. and Sorensen, D. (2013) 'Prediction of Complex Human Traits Using the Genomic Best Linear Unbiased Predictor', *PLoS Genetics*, 9(7). doi: 10.1371/journal.pgen.1003608.

- de los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D. and Calus, M. P. L. (2013) ‘Whole-genome regression and prediction methods applied to plant and animal breeding’, *Genetics*, 193(2), pp. 327–345. doi: 10.1534/genetics.112.143313.
- de Los Campos, G., Gianila, D., Rosa, G., Weigel, K. and Crossa, J. (2010) ‘Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods’, *Genetics Research*, 92(4), pp. 295–308. doi: 10.1017/S0016672310000285.
- de los Campos, G., Gianola, D. and Allison, D. B. (2010) ‘Predicting genetic predisposition in humans: the promise of whole-genome markers’, *Nature Reviews Genetics*, 11(12), pp. 880–886. doi: 10.1038/nrg2898.
- Lynch, M. and Walsh, B. (1998) *Genetics and Analysis of Quantitative Traits*. Sunderland: Sinauer Associates.
- Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Gibbs, R. A., *et al.* (2012) ‘The *Drosophila melanogaster* Genetic Reference Panel’, *Nature*, 482(7384), pp. 173–178. doi: 10.1038/nature10811.
- Mackay, T. F. C. (2014) ‘Epistasis and Quantitative Traits: Using Model Organisms to Study Gene-Gene Interactions’, *Nat Rev Genet.*, 15(1), pp. 22–33. doi: 10.1038/nrg3627.
- Mackay, T. F. C., Stone, E. A. and Ayroles, J. F. (2009) ‘The genetics of quantitative traits: challenges and prospects’, *Nature Reviews Genetics*, 10(8), pp. 565–577. doi: 10.1038/nrg2612.
- Mafouasson, N. H., Gracen, V., Yeboah, A. M., Ntsomboh-Ntsefong, G., Tandzi, N. L. and Mutengwa, S. C. (2018) ‘Genotype-by-Environment Interaction and Yield Stability of Maize Single Cross Hybrids Developed from Tropical Inbred Lines’, *Agronomy*, 8(5). doi: 10.3390/agronomy8050062.
- Mäki-Tanila, A. and Hill, W. G. (2014) ‘Influence of Gene Interaction on Complex Trait Variation with Multilocus Models’, *Genetics*, 198(1), pp. 355–367. doi: 10.1534/genetics.114.165282.
- Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorf, L. A., Hunter, D. J., McCarthy, M. I., Ramos, E. M., Cardon, L. R., Chakravarti, A., Cho, J. H., Guttmacher, A. E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C. N., Slatkin, M., Valle, D., Whittemore, A. S., Boehnke, M., Clark, A. G., Eichler, E. E., Gibson, G., Haines, J. L., Mackay, T. F. C., McCarroll, S. A. and Visscher, P. M. (2009) ‘Finding the missing heritability of complex diseases’, *Nature*, 461(7265), pp. 747–753. doi: 10.1038/nature08494.
- Martini, J. W. R., Wimmer, V., Erbe, M. and Simianer, H. (2016) ‘Epistasis and covariance: how gene interaction translates into genomic relationship’, *Theoretical and Applied Genetics*, 129(5), pp. 963–976. doi: 10.1007/s00122-016-2675-5.
- Martini, J. W. R., Gao, N., Cardoso, D. F., Wimmer, V., Erbe, M., Cantet, R. J. C. and Henner, S. (2017) ‘Genomic prediction with epistasis models: on the marker-coding-dependent performance of the extended GBLUP and properties of the categorical epistasis model (CE)’, *BMC Bioinformatics*, 18(3). doi: 10.1186/s12859-016-1439-1.
- Martini, J. W. R., Rosales, F., Ha, N.-T., Heise, J., Wimmer, V. and Kneib, T. (2019) ‘Lost in Translation: On the Problem of Data Coding in Penalized Whole Genome Regression with Interactions’, *G3: Genes/Genomes/Genetics*, 9(4), pp. 1117–1129. doi: 10.1534/g3.118.200961.
- Martini, J. W. R., Toledo, F. H. and Crossa, J. (2020) ‘On the approximation of interaction effect models by Hadamard powers of the additive genomic relationship’, *Theoretical population biology*, 132, pp. 16–23. doi: 10.1016/j.tpb.2020.01.004.

- Meuwissen, T. H. E., Odegard, J., Andersen-Ranberg, I. and Grindflek, E. (2014) ‘On the distance of genetic relationships and the accuracy of genomic prediction in pig breeding’, *Genetics Selection Evolution*, 46(49). doi: 10.1186/1297-9686-46-49.
- Meuwissen, T. H. E., Hayes, B. J. and Goddard, M. E. (2001) ‘Prediction of total genetic value using genome-wide dense marker maps’, *Genetics*, 157(4), pp. 1819–1829. doi: doi.org/10.1093/genetics/157.4.1819.
- Millet, E. J., Kruijjer, W., Coupel-Ledru, A., Alvarez Prado, S., Cabrera-Bosquet, L., Lacube, S., Charcosset, A., Welcker, C., van Eeuwijk, F. and Tardieu, F. (2019) ‘Genomic prediction of maize yield across European environmental conditions’, *Nature Genetics*, 51(6), pp. 952–956. doi: 10.1038/s41588-019-0414-y.
- Momen, M., Mehrgardi, A. A., Sheikhi, A., Kranis, A., Tusell, L., Morota, G., Rosa, G. J. M. and Gianola, D. (2018) ‘Predictive ability of genome-assisted statistical models under various forms of gene action’, *Scientific Reports*, 8(12309). doi: 10.1038/s41598-018-30089-2.
- Montesinos-López, A., Montesinos-López, O. A., Gianola, D., Crossa, J. and Hernández-Suárez, C. M. (2018) ‘Multi-environment Genomic Prediction of Plant Traits Using Deep Learners With Dense Architecture’, *G3: Genes/Genomes/Genetics*, 8(12), pp. 3813–3828. doi: 10.1534/g3.118.200740.
- Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Toledo, F. H., Pérez-Hernández, O., Eskridge, K. M. and Rutkoski, J. (2016) ‘A Genomic Bayesian Multi-trait and Multi-environment Model’, *G3: Genes/Genomes/Genetics*, 6(9), pp. 2725–2744. doi: 10.1534/g3.116.032359.
- Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Gianola, D., Hernández-Suárez, C. M. and Martín-Vallejo, J. (2018) ‘Multi-trait, Multi-environment Deep Learning Modeling for Genomic-Enabled Prediction of Plant Traits’, *G3 (Bethesda, Md.)*. Genetics Society of America, 8(12), pp. 3829–3840. doi: 10.1534/g3.118.200728.
- Morgil, H. (2020) ‘Single Nucleotide Polymorphisms (SNPs) in Plant Genetics and Breeding’, in. Rijeka: IntechOpen. doi: 10.5772/intechopen.91886.
- Mrode, R. A. (2014) *Linear models for the prediction of animal breeding values*. Third edit. Wallingford, OXon, UK: CAB International. doi: 10.1079/9780851990002.0000.
- Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E. and Buckler, E. S. (2009) ‘Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design’, *The Plant Cell*, 21(8), pp. 2194–2202. doi: 10.1105/tpc.109.068437.
- Nelson, R. M., Pettersson, M. E. and Carlborg, Ö. (2013) ‘A century after Fisher: time for a new paradigm in quantitative genetics’, *Trends in Genetics*, 29(12), pp. 669–676. doi: 10.1016/j.tig.2013.09.006.
- Nordborg, M. and Weigel, D. (2008) ‘Next-generation genetics in plants’, *Nature*, 456(7223), pp. 720–723. doi: 10.1038/nature07629.
- Ober, U., Ayroles, J. F., Stone, E. A., Richards, S., Zhu, D., Gibbs, R. A., Stricker, C., Gianola, D., Schlather, M., Mackay, T. F. C. and Simianer, H. (2012) ‘Using whole-genome sequence data to predict quantitative trait phenotypes in *Drosophila melanogaster*’, *PLoS genetics*. 2012/05/03. Public Library of Science, 8(5), p. e1002685. doi: 10.1371/journal.pgen.1002685.
- Ober, U., Huang, W., Magwire, M., Schlather, M., Simianer, H. and Mackay, T. F. C. (2015) ‘Accounting for Genetic Architecture Improves Sequence Based Genomic Prediction for a *Drosophila* Fitness Trait’, *PLoS ONE*. Public Library of Science, 10(5), p. e0126880. doi: 10.1371/journal.pone.0126880.



- Park, T. and Casella, G. (2008) ‘The Bayesian Lasso’, *Journal of the American Statistical Association*, 103(482), pp. 681–686. doi: 10.1198/016214508000000337.
- Pérez, P., de Los Campos, G., Crossa, J. and Gianola, D. (2010) ‘Genomic-Enabled Prediction Based on Molecular Markers and Pedigree Using the Bayesian Linear Regression Package in R’, *The Plant Genome*, 3(2), pp. 106–116. doi: 10.3835/plantgenome2010.04.0005.
- Pérez, P. and de los Campos, G. (2014) ‘Genome-wide regression and prediction with the BGLR statistical package’, *Genetics*. Genetics Society of America, 198(2), pp. 483–495. doi: 10.1534/genetics.114.164442.
- Persa, R., Bernardeli, A. and Jarquin, D. (2020) ‘Prediction Strategies for Leveraging Information of Associated Traits under Single- and Multi-Trait Approaches in Soybeans’, *Agriculture*, 10(308). doi: 10.3390/agriculture10080308.
- Pook, T., Schlather, M., de los Campos, G., Mayer, M., Schoen, C. C. and Simianer, H. (2019) ‘HaploBlocker: Creation of subgroup specific haplotype blocks and libraries’, *GENETICS*, 212(4), pp. 1045–1061. doi: 10.1534/genetics.119.302283.
- Pook, T., Mayer, M., Geibel, J., Weigend, S., Cavero, D., Schoen, C. C. and Simianer, H. (2020) ‘Improving Imputation Quality in BEAGLE for Crop and Livestock Data’, *G3: Genes/Genomes/Genetics*, 10(1), pp. 177–188. doi: 10.1534/g3.119.400798.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., Bakker, P. I. W. de, Daly, M. J. and Sham, P. C. (2007) ‘PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses’, *American Journal of Human Genetics*, 81(3), pp. 559–575. doi: 10.1086/519795.
- Rönnegård, L. and Shen, X. (2016) ‘Genomic prediction and estimation of marker interaction effects’, *bioRxiv*, 38935. doi: 10.1101/038935.
- Roorkiwal, M., Jarquin, D., Singh, M. K., Gaur, P. M., Bharadwaj, C., Rathore, A., Howard, R., Srinivasan, S., Jain, A., Garg, V., Kale, S., Chitikineni, A., Tripathi, S., Jones, E., Robbins, K. R., Crossa, J. and Varshney, R. K. (2018) ‘Genomic-enabled prediction models using multi-environment trials to estimate the effect of genotype × environment interaction on prediction accuracy in chickpea’, *Sci Rep.*, 8(11701). doi: 10.1038/s41598-018-30027-2.
- Runcie, D. and Cheng, H. (2019) ‘Pitfalls and Remedies for Cross Validation with Multi-trait Genomic Prediction Methods’, *G3: Genes/Genomes/Genetics*, 9(11), pp. 3727–3741. doi: 10.1534/g3.119.400598.
- Schaeffer, L. R. (2006) ‘Strategy for applying genome-wide selection in dairy cattle’, *Journal of Animal Breeding & Genetics*, 123(4), pp. 218–223.
- Schlather, M. (2020) ‘Efficient Calculation of the Genomic Relationship Matrix’, *bioRxiv*. doi: 10.1101/2020.01.12.903146.
- Schrag, T. A., Schipprack, W. and Melchinger, A. E. (2019a) ‘Across-years prediction of hybrid performance in maize using genomics’, *Theoretical and Applied Genetics*, 132, pp. 933–946.
- Schrag, T. A., Schipprack, W. and Melchinger, A. E. (2019b) ‘Across-years prediction of hybrid performance in maize using genomics’, *Theoretical and Applied Genetics*. Springer Verlag, 132(4), pp. 933–946. doi: 10.1007/s00122-018-3249-5.
- Schrauf, M. F., Martini, J. W. R., Simianer, H. and de Los Campos, G. (2020) ‘Phantom Epistasis in Genomic Selection: On the predictive ability of epistatic models’, *G3: Genes/Genomes/Genetics*, 10(9), pp. 3137–3145. doi: 10.1534/g3.120.401300.

- Schulthess, A. W., Zhao, Y., Longin, C. F. H. and Reif, J. C. (2018) 'Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (*Triticum aestivum* L.)', *Theoretical and Applied Genetics*, 131(3), pp. 685–701. doi: 10.1007/s00122-017-3029-7.
- Shin, J. and Lee, S. H. (2020) 'GxEsum: genotype-by-environment interaction model based on summary statistics', *bioRxiv*. doi: 10.1101/2020.05.31.122549.
- Slatkin, M. (2008) 'Linkage disequilibrium--understanding the evolutionary past and mapping the medical future', *Nature reviews. Genetics*, 9(6), pp. 477–485. doi: 10.1038/nrg2361.
- Stich, B. and Ingheland, D. Van (2018) 'Prospects and Potential Uses of Genomic Prediction of Key Performance Traits in Tetraploid Potato', *Frontiers in Plant Science*, 9(159). doi: 10.3389/fpls.2018.00159.
- Tan, B., Grattapaglia, D., Martins, G. S., Ferreira, K. Z., Sundberg, B. and Ingvarsson, P. K. (2017) 'Evaluating the accuracy of genomic prediction of growth and wood traits in two Eucalyptus species and their F(1) hybrids', *BMC plant biology*. BioMed Central, 17(110). doi: 10.1186/s12870-017-1059-6.
- Tester, M. and Langridge, P. (2010) 'Breeding Technologies to Increase Crop Production in a Changing World', *Science*, 327(5967), pp. 818–822. doi: 10.1126/science.1183700.
- Thompson, R. and Meyer, K. (1986) 'A review of theoretical aspects in the estimation of breeding values for multi-trait selection', *Livestock Production Science*, 15(4), pp. 299–313. doi: 10.1016/0301-6226(86)90071-0.
- Tsai, H.-Y., Janss, L. L., Andersen, J. R., Orabi, J., Jensen, J. D., Jahoor, A. and Jensen, J. (2020) 'Genomic prediction and GWAS of yield, quality and disease-related traits in spring barley and winter wheat', *Scientific Reports*, 10(3347). doi: 10.1038/s41598-020-60203-2.
- Unterseer, S., Bauer, E., Haberer, G., Seidel, M., Knaak, C., Ouzunova, M., Meitinger, T., Strom, T. M., Fries, R., Pausch, H., Bertani, C., Davassi, A., Mayer, K. F. and Schön, C.-C. (2014) 'A powerful tool for genome analysis in maize: 584 development and evaluation of the high density 600 k SNP genotyping array', *BMC Genomics*, 15(823). doi: 10.1186/1471-2164-15-823.
- Utz, H. F., Melchinger, A. E. and Schön, C. C. (2000) 'Bias and Sampling Error of the Estimated Proportion of Genotypic Variance Explained by Quantitative Trait Loci Determined From Experimental Data in Maize Using Cross Validation and Validation With Independent Samples', *Genetics*, 154(4), pp. 1839–1849. doi: 10.1093/genetics/154.4.1839.
- VanRaden, P. (2007) 'Efficient estimation of breeding values from dense genomic data', *Journal of Dairy Science*, 90, pp. 374–375.
- VanRaden, P. (2008) 'Efficient methods to compute genomic predictions', *Journal of Dairy Science*, 91(11), pp. 4414–4423. doi: 10.3168/jds.2007-0980.
- Vazquez, A. I., de los Campos, G., Klimentidis, Y. C., Rosa, G. J. M., Gianola, D., Yi, N. and Allison, D. B. (2012) 'A Comprehensive Genetic Approach for Improving Prediction of Skin Cancer Risk in Humans', *Genetics*, 192(4), pp. 1493–1502. doi: 10.1534/genetics.112.141705.
- Velazco, J. G., Jordan, D. R., Mace, E. S., Hunt, C. H., Malosetti, M. and van Eeuwijk, F. A. (2019) 'Genomic Prediction of Grain Yield and Drought-Adaptation Capacity in Sorghum Is Enhanced by Multi-Trait Analysis', *Frontiers in Plant Science*, 10(997). doi: doi.org/10.3389/fpls.2019.00997.
- Vojgani, E., Pook, T., Martini, J. W. R., Hölker, A. C., Mayer, M., Schön, C.-C. and Simianer, H. (2020) 'Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments', *bioRxiv*. doi: 10.1101/2020.10.08.331074.

- Vojgani, E., Pook, T. and Simianer, H. (2019) ‘EpiGP: Epistatic relationship matrix based genomic prediction of phenotypes’. Available at: <https://github.com/evojpgani/EpiGP>.
- Vojgani, E., Pook, T. and Simianer, H. (2021) ‘Phenotype Prediction under Epistasis’, in *Epistasis: Methods and Protocols*. Springer. doi: 10.1007/978-1-0716-0947-7\_8.
- Walsh, B. and Lynch, M. (2018) *Evolution and Selection of Quantitative Traits*. Oxford, United Kingdom : Oxford University Press.
- Wang, D., El-Basyoni, I. S., Baenziger, P. S., Crossa, J., Eskridge, K. M. and Dweikat, I. (2012) ‘Prediction of genetic values of quantitative traits with epistatic effects in plant breeding populations’, *Heredity*, 109(5), pp. 313–319. doi: 10.1038/hdy.2012.44.
- Wang, J., Zhou, Z., Zhang, Zhe, Li, H., Liu, D., Zhang, Q., Bradbury, P. J., Buckler, E. S. and Zhang, Zhiwu (2018) ‘Expanding the BLUP alphabet for genomic prediction adaptable to the genetic architectures of complex traits’, *Heredity*, 121(6), pp. 648–662. doi: 10.1038/s41437-018-0075-0.
- Wentzell, A. M., Rowe, H. C., Hansen, B. G., Ticconi, C., Halkier, B. A. and Kliebenstein, D. J. (2007) ‘Linking metabolic QTLs with network and cis-eQTLs controlling biosynthetic pathways’, *PLoS genetics*. 2007/08/01. Public Library of Science, 3(9), pp. 1687–1701. doi: 10.1371/journal.pgen.0030162.
- Wimmer, V., Lehermeier, C., Albrecht, T., Auinger, H.-J., Wang, Y. and Schön, C.-C. (2013) ‘Genome-Wide Prediction of Traits with Different Genetic Architecture Through Efficient Variable Selection’, *Genetics*, 195(2), pp. 573–587. doi: 10.1534/genetics.113.150078.
- Windhausen, V. S., Atlin, G. N., Hickey, J. M., Crossa, J., Jannink, J.-L., Sorrells, M. E., Raman, B., Cairns, J. E., Tarekegne, A., Semagn, K., Beyene, Y., Grudloyma, P., Technow, F., Riedelsheimer, C. and Melchinger, A. E. (2012) ‘Effectiveness of Genomic Prediction of Maize Hybrid Performance in Different Breeding Populations and Environments’, *G3 (Bethesda)*, 2(11), pp. 1427–1436. doi: 10.1534/g3.112.003699.
- Wolc, A., Stricker, C., Arango, J., Settar, P., Fulton, J. E., O’Sullivan, N. P., Preisinger, R., Habier, D., Fernando, R., Garrick, D. J., Lamont, S. J. and Dekkers, J. C. M. (2011) ‘Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model’, *Genetics Selection Evolution*, 43(5). doi: 10.1186/1297-9686-43-5.
- Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E. and Visscher, P. M. (2013) ‘Pitfalls of predicting complex traits from SNPs’, *Nature Reviews Genetics*, 14(7), pp. 507–515. doi: 10.1038/nrg3457.
- Wray, N. R., Goddard, M. E. and Visscher, P. M. (2007) ‘Prediction of individual genetic risk to disease from genome-wide association studies’, *Genome research*. 2007/09/04. Cold Spring Harbor Laboratory Press, 17(10), pp. 1520–1528. doi: 10.1101/gr.6665407.
- Wright, S. (1931) ‘Evolution in Mendelian populations’, *Genetics*, 16(2), pp. 97–159.
- Xu, Y., Li, P., Zou, C., Lu, Y., Xie, C., Zhang, X., Prasanna, B. M. and Olsen, M. S. (2017) ‘Enhancing genetic gain in the era of molecular breeding’, *Journal of Experimental Botany*, 68(11), pp. 2641–2666. doi: 10.1093/jxb/erx135.
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E. and Visscher, P. M. (2010) ‘Common SNPs explain a large proportion of the heritability for human height’, *Nature Genetics*, 42(7), pp. 565–569. doi: 10.1038/ng.608.

Zou, H. and Hastie, T. (2005) ‘Regularization and variable selection via the elastic net’, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), pp. 301–320. doi: 10.1111/j.1467-9868.2005.00503.x.

Zuk, O., Hechter, E., Sunyaev, S. R. and Lander, E. S. (2012) ‘The mystery of missing heritability: Genetic interactions create phantom heritability’, *Proceedings of the National Academy of Sciences*, 109(4), pp. 1193–1198. doi: 10.1073/pnas.1119675109.

## 7 List of Figures

**Fig. 2.1:** Comparison of predictive ability of GBLUP (open circle), ERRBLUP (black filled circle) and sERRBLUP for 50 to 0.1% of SNP interactions maintained in the model. Red and green dots represent the sERRBLUP predictive ability (with standard errors) when the SNP interaction selections are based on the estimated effects and estimated effects variances, respectively. The vertical arrows represent the predictions  $\pm$  their standard deviations. .... 40

**Fig. 2.2:** Comparison of predictive ability of GBLUP (open circle), ERRBLUP (black filled circle) and sERRBLUP for 95 to 0.1% of SNP interactions maintained in the model. Green dots represent the sERRBLUP predictive ability for prediction within environment 3, while black, blue and red dots represent the predictive ability with relationship matrices determined by variable selection in environments 1, 2 and 4. .... 41

**Fig. 2.3:** Comparison of computational time for sERRBLUP by utilizing 15 cores (red bars) and the method of Martini (Martini et al., 2016) with a reduced epistatic model (blue bars) for the top 95% of interactions maintenance in the model in different numbers of SNPs. .... 43

**Fig. 3.1:** Basic scheme of uni- and bivariate sERRBLUP across environments. All pairwise SNP interaction effects and their variances are estimated from all data in environment 1, and effects are ordered either by absolute effect size or effect variance (A). Then, an epistatic relationship matrix for all lines is constructed from the top ranked subset of interaction effects (B) which in the univariate model is used in environment 2 (C) to predict phenotypes of the test set (green) from the respective training set (red, D). In the bivariate model, this information is combined with the complete data from environment 1 (blue, E) to predict the test set. .... 56

**Fig. 3.2:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments (solid colored lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH\_V4. In each panel, the solid lines' color indicates the environment in which the relationship matrices were determined by variable selection. .... 59

**Fig. 3.3:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH\_V4. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 60

**Fig. 3.4:** Percentage of increase in prediction accuracy from univariate GBLUP within environments to the maximum prediction accuracy of univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances in KE (left side) and in PE (right side). The average percentage of increase in prediction accuracy for each trait and environments are displayed in rows and columns, respectively. .... 62

**Fig. 3.5:** Percentage of increase in prediction accuracy from the maximum bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP when the SNP interaction selections are based on estimated effects variances in KE (left side) and in PE (right side). The average percentage of increase in prediction accuracy for each trait and environments are displayed in rows and columns, respectively. .63

**Fig. 3.6:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (gray open circle), univariate sERRBLUP using a single environment for selecting the SNP interactions (gray open circles) and univariate sERRBLUP using all 5 environments jointly (filled black circles and solid line) for the SNP interaction selection based on estimated effects variances for trait PH\_V4 in KE (left side) and PE (right side). ..... 65

**Fig. 3.7:** Predictive ability for bivariate GBLUP (black dashed horizontal line), bivariate ERRBLUP and bivariate sERRBLUP (filled black circles) for the SNP interaction selection based on estimated effects variances using all 5 environments jointly for trait PH-V4 in KE (left side) and PE (right side). In each panel, gray horizontal line and first gray open circles refer to maximum bivariate GBLUP and maximum bivariate ERRBLUP, and the gray open circles at the top 5, 1, 0.1, 0.01, 0.001, 0.0001 quantiles refer to bivariate sERRBLUP using a single environment as an additional environment. .... 66

**Fig. 4.1:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH-V4 in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 116

**Fig. 4.2:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH-V4 in PE based on pruned sets of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 117

**Fig. 4.3:** Regression of the absolute increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP on the respective sERRBLUP genomic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits. In each panel, the overall linear regression line (gray solid line) with the regression coefficient (***b***) and R-squared (***R*<sup>2</sup>**) are shown. .... 118

**Fig. 4.4:** Regression of the absolute increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP on the phenotypic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits. In each panel, the overall linear regression line (gray solid line) with the regression coefficient (***b***) and R-squared (***R*<sup>2</sup>**) are shown. .... 120

**Fig. 4.5:** Percentage of change in prediction accuracy from bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP based on pruned set of SNPs in KE (left side plot) and in PE (right side plot). The average percentage of change in prediction accuracy for each trait and environment is displayed in all rows and columns, respectively. .... 121

**Fig. 5.1:** Predictive ability of EG-BLUP within environment (black circle) as a full epistasis model and reduced EG-BLUP across environments (colored circles and solid lines) when interaction selections are based on eGWAS (left) and rrGBLUP (right) for wheat grain yield. In each panel, the solid lines' color indicates the environment in which the relationship matrices were determined by variable selection. 159

**Fig. 5.2:** Predictive ability of EG-BLUP within environment (black circle) as a full epistasis model and reduced EG-BLUP across all three environments jointly when interaction selections are based on eGWAS. Orange, red, green and blue lines colors indicate the predictive abilities in environment 1, 2, 3, and 4, when the other three environments are used jointly for variable selection. Vertical lines represent the

range of accuracies obtained when selection of sub-networks was in a single environment rather than combining three environments and predict the fourth based on eGWAS..... 160

**Fig. 5.3:** Comparison of predictive abilities obtained from univariate GBLUP when the panel of 501'124 SNPs is used (red) and when the pruned set of 25'437 SNPs is used (green) in KE for prediction within environments across series of all traits. .... 162

**Fig. 5.4:** Comparison of predictive abilities obtained from univariate GBLUP when the panel of 501'124 SNPs is used (red) and when the pruned set of 30'212 SNPs is used (green) in PE for prediction within environments across series of all traits. .... 163

**Fig. 5.5:** Comparison of bivariate GBLUP predictive abilities obtained through masking test sets in both target and additional environment (blue bars) and masking only a test set of the target environment (red bars) in cross validation scenario for trait PH\_V4 in 2017 in both KE (left) and PE (right). In each panel, the dashed horizontal line indicates the univariate GBLUP predictive ability for prediction within each environment. .... 165

**Fig. 5.6:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH-V4 in 2017 when the test sets in both target and additional environments are masked in cross validation scenario. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 166

**Fig. 5.7:** Absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP (squares) and maximum bivariate sERRBLUP (circles) versus the percentage of genotypes overlap between the target environment and the additional environment in 2017 in KE (left) and in PE (right) for trait PH\_V4. In each panel, colored circles and squares represent the additional environment used in bivariate model to predict the target environment with the R-squared representing the correlation coefficient between the respective increase in bivariate GBLUP (***RGBLUP2***) and bivariate sERRBLUP (***RsERRBLUP2***) predictive abilities and the percentage of genotype overlap..... 169





## 8 List of Tables

<b>Table 3.1:</b> Phenotypic correlation across all environments for the trait PH_V4 in KE (blue numbers above diagonal) and PE (red numbers below diagonal) which are highly significant ( $p\_values < 0.001$ ).....	50
<b>Table 3.2:</b> Number of KE (blue numbers above diagonal) and PE (red numbers below diagonal) phenotyped lines in each pair of environments for trait PH_V4.....	55
<b>Table 4.1:</b> Phenotypic trait description and the mean, minimum, maximum and standard deviation of the BLUEs for each phenotypic trait in KE and PE landraces in the years 2017 and 2018.....	112
<b>Table 4.2:</b> Number of KE and PE lines phenotyped in each location for the years 2017 (blue numbers) and 2018 (red numbers) for trait PH_V4. ....	113
<b>Table 4.3:</b> Genomic correlation between 2017 and 2018 in each environment for trait PH_V4 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively.....	119
<b>Table 5.1:</b> The phenotypic correlation, GBLUP and sERRBLUP genomic correlations of the environments combinations with genotype overlap above 85 percent and the respective absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP and maximum bivariate sERRBLUP for the trait PH_V4 in KE. ....	172
<b>Table 5.2:</b> The phenotypic correlation, GBLUP and sERRBLUP genomic correlations of the environments combinations with genotype overlap above 85 percent and the respective absolute increase in predictive ability from univariate GBLUP to bivariate GBLUP and maximum bivariate sERRBLUP for the trait PH_V4 in PE. ....	173



## 9 List of Supplementary Figures

- Fig. S. 3.1:** Comparison of pre estimated genetic and residual variances and covariances of converged bivariate sERRBLUP model (top 5%) based on the full dataset (dashed horizontal lines) and estimated genetic and residual variances and covariances of converged bivariate sERRBLUP (top 5%) based on training set in each run of 5-fold cross validation with 5 replicates (colored bars) for predicting BBG when the additional environment is OLI in KE for trait PH\_V4. .... 81
- Fig. S. 3.2:** The difference between the mean predictive ability of only the converged folds and the mean predictive ability of all folds in 5-fold cross validation with 5 replicates vs. the number of the folds (1 to 24) which did not converge across all traits in all combinations for both KE and PE. .... 82
- Fig. S. 3.3:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait EV\_V3 in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection. .... 83
- Fig. S. 3.4:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait EV\_V4 in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection. .... 84
- Fig. S. 3.5:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait EV\_V6 in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection. .... 85
- Fig. S. 3.6:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait PH-V6 in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection. .... 86
- Fig. S. 3.7:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait PH-final in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection. .... 87
- Fig. S. 3.8:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait FF in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection. .... 88

**Fig. S. 3.9:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances (solid colored lines) for trait RL in KE (left side plots) and PE (right side plots). In each panel, the solid lines' color indicates the environment in which the relationship matrices determined by variable selection. .... 89

**Fig. S. 3.10:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments (solid colored lines) when SNP interaction selections are based on estimated effects sizes (left side) and estimated effects variances (right side) for trait PH-V4 in KE. In each panel, the solid lines' color indicates the environment in which the relationship matrices were determined by variable selection. .... 90

**Fig. S. 3.11:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (black filled circle) and univariate sERRBLUP across environments (solid colored lines) when SNP interaction selections are based on estimated effects sizes (left side) and estimated effects variances (right side) for trait PH-V4 in PE. In each panel, the solid lines' color indicates the environment in which the relationship matrices were determined by variable selection. .... 91

**Fig. S. 3.12:** Comparison of predictive ability of univariate GBLUP within environments (filled squares) and the maximum predictive ability of univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects sizes (circles) and estimated effects variances (triangles) for trait PH-V4 in KE (left side plot) and in PE (right side plot). The colors dark blue, orange, purple, red, light blue and green represent the environments BBG, EIN, OLI, ROG, GOL and TOM, respectively. The circles' and triangles' colors indicate the environment which had the maximum predictive ability for this respective target environment..... 92

**Fig. S. 3.13:** Predictive ability for univariate GBLUP within Bernburg (dashed horizontal line), univariate ERRBLUP within Bernburg (black filled circle) and univariate sERRBLUP when the SNP interaction selections are based on estimated effects variances (blue solid line) and estimated effect sizes (red solid line) within Bernburg for trait PH-V4 in KE. .... 92

**Fig. S. 3.14:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait EV-V3. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 93

**Fig. S. 3.15:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait EV-V4. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 94

**Fig. S. 3.16:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait EV-V6. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 95

**Fig. S. 3.17:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH-V6. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 96

**Fig. S. 3.18:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait PH-final. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 97

**Fig. S. 3.19:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait FF. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 98

**Fig. S. 3.20:** Predictive ability for bivariate GBLUP (open squares), bivariate ERRBLUP (open circles) and bivariate sERRBLUP (filled circles and solid lines) when SNP interaction selections are based on estimated effects variances in KE (left side) and PE (right side) for trait RL. In each panel, the solid lines' color indicates the additional environment used to predict the target environment. .... 99

**Fig. S. 3.21:** Comparison of predictive ability of univariate GBLUP within environments (filled squares) and the maximum predictive ability of univariate sERRBLUP across EIN when the SNP interaction selections are based on estimated effects variances (orange circles) for trait PH-V4 in KE (left side plot) and in PE (right side plot). The colors dark blue, purple, red, light blue and green represent the environments BBG, OLI, ROG, GOL and TOM, respectively. .... 100

**Fig. S. 3.22:** Absolute increase in prediction accuracy from univariate GBLUP within environments to the maximum prediction accuracy of univariate sERRBLUP across environments when the SNP interaction selections are based on estimated effects variances in KE (left side plot) and in PE (right side plot). The average of absolute increase in prediction accuracy for each trait and environments are display in rows and columns, respectively. .... 101

**Fig. S. 3.23:** Absolute increase in prediction accuracy from maximum bivariate GBLUP the maximum prediction accuracy of bivariate sERRBLUP when the SNP interaction selections are based on estimated effects variances in KE (left side plot) and in PE (right side plot). The average of absolute increase in prediction accuracy for each trait and environments are display in rows and columns, respectively. ... 102

**Fig. S. 3.24:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (gray open circle), univariate sERRBLUP using a single environment for selecting the SNP interactions (gray open circles) and univariate sERRBLUP using all 5 environments jointly (filled black circles and solid line) for the SNP interaction selection based on estimated effects sizes (left side) and estimated effects variances (right side) for trait PH-V4 in KE. .... 103

**Fig. S. 3.25:** Predictive ability for univariate GBLUP within environment (dashed horizontal line), univariate ERRBLUP within environment (gray open circle), univariate sERRBLUP using a single environment for selecting the SNP interactions (gray open circles) and univariate sERRBLUP using all 5 environments jointly (filled black circles and solid line) for the SNP interaction selection based on estimated effects sizes (left side) and estimated effects variances (right side) for trait PH-V4 in PE. .... 104

**Fig. S. 3.26:** The comparison between the maximum predictive abilities of bivariate GBLUP within all six environments and the maximum predictive abilities of univariate sERRBLUP across environments for all traits in KE (left side plot) and PE (right side plot). In each plot, the diagonal line (red line) and the overall linear regression line (black line) with the regression formula are shown. .... 105

**Fig. S. 3.27:** The correlation between all eight traits' heritabilities and predictive abilities of univariate GBLUP within environments (left side) and maximum predictive abilities of univariate sERRBLUP across

environments (right side) in both landraces. The black lines indicate the overall linear regression lines. .... 106

**Fig. S. 3.28:** The correlation between all eight traits' heritabilities and the maximum predictive abilities of bivariate GBLUP (left side) and maximum predictive abilities of bivariate sERRBLUP (right side) for all environments in both landraces. The black lines indicate the overall linear regression lines. .... 106

**Fig. S. 4.1:** Comparison of pre estimated genetic and residual variances and covariances of converged bivariate sERRBLUP (top 10%) based on the full dataset (dashed horizontal lines) and estimated genetic and residual variances and covariances of converged bivariate sERRBLUP (top 10%) based on training set in each run of 5-fold cross validation with 5 replicates (colored bars) for predicting EIN in 2018 when the additional environment is EIN in 2017 in KE for trait PH-V4. .... 134

**Fig. S. 4.2:** The difference between the mean predictive ability of only the converged folds and the mean predictive ability of all folds in 5-fold cross validation with 5 replicates virus the number of the folds which did not converged across all traits in all combinations for both KE and PE in bivariate GBLUP, ERRBLUP, sERRBLUP. .... 135

**Fig. S. 4.3:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V3 in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 136

**Fig. S. 4.4:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V3 in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 137

**Fig. S. 4.5:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V4 in KE based on pPruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 138

**Fig. S. 4.6:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V4 in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 139

**Fig. S. 4.7:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V6 in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 140

**Fig. S. 4.8:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait EV\_V6 in PE based on pruned set of SNPs (left) and

haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 141

**Fig. S. 4.9:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH\_V6 in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 142

**Fig. S. 4.10:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH\_V6 in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 143

**Fig. S. 4.11:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH\_final in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 144

**Fig. S. 4.12:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait PH\_final in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 145

**Fig. S. 4.13:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait FF in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 146

**Fig. S. 4.14:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait FF in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 147

**Fig. S. 4.15:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait RL in KE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 148

**Fig. S. 4.16:** Predictive ability for univariate GBLUP within 2018 (orange and red dashed horizontal line), bivariate GBLUP (green and blue dashed horizontal line), bivariate ERRBLUP (open circle) and bivariate sERRBLUP (filled circles and solid line) for trait RL in PE based on pruned set of SNPs (left) and haplotype blocks (right). In each plot, the sERRBLUP maximum indicates the maximum predictive ability obtained from bivariate sERRBLUP. .... 148

**Fig. S. 4.17:** Regression of the absolute increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP on the respective sERRBLUP genomic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits. In each panel, the overall linear regression line (gray solid line) with the regression coefficient (***b***) and R-squared (***R*<sup>2</sup>**) are shown..... 149

**Fig. S. 4.18:** Regression of the absolute increase in predictive ability from bivariate GBLUP to maximum bivariate sERRBLUP on the difference between the GBLUP genomic correlation and maximum sERRBLUP genomic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits, when utilizing pruned set of SNPs. In each panel, the overall linear regression line with the regression coefficient (***b***) and R-squared (***R*<sup>2</sup>**) are shown. The colors green, light blue, pink, red, orange, purple, yellow and dark blue represent the phenotypic traits EV\_V3, EV\_V4, EV\_V6, PH\_V4, PH\_V6, PH\_final, FF and RL, respectively..... 150

**Fig. S. 4.19:** Regression of the absolute increase in predictive ability from univariate GBLUP to maximum bivariate sERRBLUP on the phenotypic correlation between 2017 and 2018 in KE (left) and in PE (right) for all studied traits. In each panel, the overall linear regression line (gray solid line) with the regression coefficient (***b***) and R-squared (***R*<sup>2</sup>**) are shown..... 151

**Fig. S. 4.20:** Absolute change in prediction accuracy from bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP based on pruned set of SNPs in KE (left side plot) and in PE (right side plot). The average absolute change in prediction accuracy for each trait and environment is displayed in all rows and columns, respectively..... 152

**Fig. S. 4.21:** Absolute change in prediction accuracy from bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP based on haplotype blocks in KE (left side plot) and in PE (right side plot). The average absolute change in prediction accuracy for each trait and environment is displayed in all rows and columns, respectively. .... 153

**Fig. S. 4.22:** Absolute change in prediction accuracy from bivariate GBLUP to the maximum prediction accuracy of bivariate sERRBLUP in PE vs. KE. The black line represents the overall linear regression line. .... 154

**Fig. S. 4.23:** Maximum sERRBLUP predictive ability based on pruned set of SNPs versus Maximum sERRBLUP predictive ability based on haplotype blocks across all traits in all environments and both KE and PE. ***b*** represents the regression coefficient..... 155



## 10 List of Supplementary Table

<b>Table S. 3.1:</b> Phenotypic traits descriptions, locations, number of lines, mean, minimum, maximum and standard deviation of phenotypic traits in each location for KE (blue numbers) and PE (red numbers)...	71
<b>Table S. 3.2:</b> The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait PH-V4 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.....	73
<b>Table S. 3.3:</b> The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait EV_V3 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.....	74
<b>Table S. 3.4:</b> The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait EV_V4 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.....	75
<b>Table S. 3.5:</b> The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait EV_V6 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.....	76
<b>Table S. 3.6:</b> The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait PH_V6 for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.....	77
<b>Table S. 3.7:</b> The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait PH_final for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set.....	78
<b>Table S. 3.8:</b> The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait FF for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set. ....	79
<b>Table S. 3.9:</b> The percentage of bivariate sERRBLUP model convergence in 5-fold cross validation with 5 replicates for trait RL for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set. ....	80
<b>Table S. 3.10:</b> The predictive ability of RKHS, univariate GBLUP within environments, univariate ERRBLUP within environments and maximum univariate sERRBLUP across environments (blue numbers) and the maximum predictive ability of bivariate GBLUP, bivariate ERRBLUP and bivariate sERRBLUP (red numbers) in KE and PE for the trait PH_V4. ....	80

<b>Table S. 4.1:</b> The mean, minimum, maximum and standard deviation of BLUEs of phenotypic traits in each location for KE (blue numbers) and PE (red numbers) in 2017 and 2018. ....	125
<b>Table S. 4.2:</b> The traits heritabilities in 2017, 2018 and both years jointly in KE (blue numbers) and PE (red numbers). ....	127
<b>Table S. 4.3:</b> The percentage of bivariate models convergence in 5-fold cross validation with 5 replicates based on pruned set of SNPs for both KE and PE (black percentages), only KE (blue percentages) and only PE (red percentages). The starts represent the non-convergence of pre estimated variance components based on the full set. ....	127
<b>Table S. 4.4:</b> The number of epistasis interactions maintained in the model based on haplotype blocks and pruned set of SNPs for the each proportions of interactions in KE and PE. ....	129
<b>Table S. 4.5:</b> GBLUP predictive ability based on pruned set of SNPs for prediction in 2018 with training the model either on 2018 data or the average phenotypic values of 2017 and 2018 in each environment for series of phenotypic traits in KE. ....	129
<b>Table S. 4.6:</b> GBLUP predictive ability based on pruned set of SNPs for prediction in 2018 with training the model either on 2018 data or the average phenotypic values of 2017 and 2018 in each environment for series of phenotypic traits in PE. ....	130
<b>Table S. 4.7:</b> Genomic correlation between 2017 and 2018 in each environment for trait EV_V3 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively. ....	130
<b>Table S. 4.8:</b> Genomic correlation between 2017 and 2018 in each environment for trait EV_V4 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively. ....	131
<b>Table S. 4.9:</b> Genomic correlation between 2017 and 2018 in each environment for trait EV_V6 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively. ....	131
<b>Table S. 4.10:</b> Genomic correlation between 2017 and 2018 in each environment for trait PH_V6 for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively. ....	131
<b>Table S. 4.11:</b> Genomic correlation between 2017 and 2018 in each environment for trait PH_final for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively. ....	132
<b>Table S. 4.12:</b> Genomic correlation between 2017 and 2018 in each environment for trait FF for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively. ....	132
<b>Table S. 4.13:</b> Genomic correlation between 2017 and 2018 in each environment for trait RL for KE (blue numbers) and PE (red numbers). The blue and red bold numbers with stars indicate which proportion of	

interactions in bivariate sERRBLUP maximized the predictive ability based on pruned set of SNPs in each environment for KE and PE, respectively..... 132

**Table S. 4.14:** Phenotypic correlation between 2017 and 2018 in each environment for KE and PE..... 133



# Curriculum Vitae

## Personal Details

**Name** Elaheh Vojgani  
**Date / Place of Birth** 10<sup>th</sup> of May, 1991, Tehran, Iran

## Education

2017 – 2020  
Ph.D. student in Agricultural sciences  
Georg-August Universität Göttingen, Germany  
Thesis: "Accounting for Epistasis in Genomic Phenotype Prediction"  
Disputation data: 22.01.2021

2012 – 2014  
Master of Science in Mathematical Statistics  
University of Tehran, Tehran, Tehran Province, Iran  
Thesis: "Non-parametric survival analysis of infectious disease data"  
Dissertation Grade: 19.40/20

2008 – 2012  
Bachelor of Science in Statistics  
University of Tehran, Tehran, Tehran Province, Iran  
Thesis: "Influence of foeniculum vulgare and lepidium perfoliatum on milk of dairy cow Project"  
Dissertation Grade: 19.75/20

2004 – 2008  
Diploma in Mathematics & Physics  
Tehran, Iran

## Employment

February 2017- November 2020  
Scientific Assistant (Ph.D. Student) to Prof. Dr. Henner Simianer  
Department of Animal Breeding  
University of Göttingen, Germany  
Project: MAZE – Accessing the genomic and functional  
Diversity of maize to improve quantitative traits

April 2015 – June 2015  
Mathematics and Statistics Tutor  
Mathematics and Statistics Institute - Tehran, Tehran Province

January 2013 – March 2013

Statistician

Statistical Center of Iran, Tehran, Tehran Province, Iran

Project: Review of internal emigration flows and comparing the characteristics of immigrants and non-immigrants in Iran in the 2006 to 2011 time period

September 2012 – March 2013

Statistical Researcher

Statistical Research and Training Center of Iran, Tehran, Tehran Province, Iran

Project: K-means clustering of households in terms of the expected income

## Publications

Vojgani, E., Pook, T. and Simianer, H. (2019) EpiGP: Epistatic relationship matrix based genomic prediction of phenotypes. R-Package available at: <https://github.com/evoigani/EpiGP>

Vojgani, E., Pook, T. and Simianer, H. (2019) Phenotype Prediction under Epistasis. Published as a chapter of a book “Epistasis: Methods and Protocols”. KC, W. (ed.) Springer

Vojgani, E., Pook, T., Martini, J. W. R., Hoelker, A. C., Mayer, M., Schoen, C.-C. and Simianer, H. (2020) Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments. Accepted in the journal TAG (Theoretical and Applied Genetics).

Vojgani, E., Pook, T., Hoelker, A. C., Mayer, M., Schoen, C.-C. and Simianer, H. (2020) Bivariate genomic prediction of phenotypes by selecting epistatic interactions across years based on haplotype blocks and pruned sets of SNPs. Submitted to the journal TAG (Theoretical and Applied Genetics)

## Talks

10<sup>th</sup> October 2019

TAST2019 Symposium. Munich, Germany

Vojgani, E., Martini, J. W. R., Pook, T., Simianer, H. (2019) Accounting for epistasis improves genomic prediction of phenotypes within and across environments.

8<sup>th</sup> October 2019

XXIVth EUCARPIA Maize and Sorghum Conference. Munich, Germany

Vojgani, E., Pook, T., Simianer, H. (2019) Accounting for epistasis improves genomic prediction of phenotypes within and across environments.

10<sup>th</sup> September 2019

Cibreed Workshop. Goettingen, Germany

Vojgani, E., Pook, T., Simianer, H. (2019) Accounting for epistasis improves genomic prediction of phenotypes within and across environments.

7<sup>th</sup> - 8<sup>th</sup> October 2018

MAZE Projekttreffen, KWS, Einbeck, Germany

Vojgani, E., Martini, J. W. R., Pook, T., Simianer, H. (2018) Genomic prediction within and across environments.

22<sup>nd</sup> – 23<sup>rd</sup> October 2017

Universität Hohenheim, Stuttgart, Germany

MAZE Projekttreffen

Vojgani, E., Martini, J. W. R., Simianer, H. (2017) Accounting for epistatic interaction in phenotype prediction across environments.

## Posters

10<sup>th</sup> October 2019

TAST2019 Symposium. Munich, Germany

Vojgani, E., Martini, J. W. R., Pook, T., Simianer, H. (2019) Accounting for epistasis improves genomic prediction of phenotypes within and across environments.

13<sup>th</sup> - 15<sup>th</sup> March 2019

PLANT2030 Statusseminar. Potsdam, Germany

Vojgani, E., Martini, J. W. R., Pook, T., Schoen, C.-C., Simianer, H. (2019) Accounting for epistasis improves genomic prediction of phenotypes within and across environments.

28<sup>th</sup> February - 2<sup>nd</sup> March 2018

German Plant Breeding Conference 2018. Wernigerode, Germany

Vojgani E., Martini, J. W. R., Simianer, H. (2018) Accounting for epistatic interaction in phenotype prediction across environments.

5<sup>th</sup> - 7<sup>th</sup> February 2018

PLANT2030 Statusseminar. Potsdam, Germany

Vojgani, E., Martini, J. W. R., Simianer, H. (2018) Accounting for epistatic interaction in phenotype prediction across environments.

**Courses**

25<sup>th</sup> June - 2<sup>nd</sup> July 2019

Herrsching, Germany

TUM Summer School 2019 - Selection and Breeding

3<sup>rd</sup> December - 7<sup>th</sup> December 2018

Freising, Germany

TUM JULIA Workshop

8<sup>th</sup> - 9<sup>th</sup> June 2017

Einbeck, Germany

Plant 2030 Academy workshop

February 2015 - February 2016

Toefl Courses. Tehran, Iran

**Research Interest**

Statistical Genetics

Genomic prediction

Statistical Modeling

Data Analysis

Biostatistics

**Honors & Awards**

Ranked 1<sup>th</sup> in Statistics of University of Tehran, in B.Sc.

Offered to the Master Program for fall 2012 in the school of Mathematics, Statistics and Computer Science at the University of Tehran, Tehran, Iran. Entrance Exam for Mathematical Statistics Waived as an Award for being among the Top-10% students.