Institute of Plant Breeding, Seed Science and Population Genetics
Applied Genetics and Plant Breeding
Prof. Dr. Albrecht E. Melchinger
University of Hohenheim

Gibberella ear rot resistance in European maize: genetic analysis by complementary mapping approaches and improvement with genomic selection

Dissertation
submitted in fulfillment of the regulation to acquire the degree
“Doktor der Agrarwissenschaften”
(Dr. sc. agr./Ph.D. in Agricultural Sciences)
to the
Faculty of Agricultural Sciences

by

Sen Han

from

Shandong, China

2022

This thesis was accepted as a doctoral dissertation in fulfillment of the requirements for the degree “Doktor der Agrarwissenschaften” (Dr. sc. agr. / Ph.D. in Agricultural Sciences) by the Faculty of Agricultural Sciences at the University of Hohenheim, on May 22nd, 2019.

Date of oral examination: April 26th, 2022

Examination Committee:
Chairman of Committee: Prof. Dr. Uwe Ludewig
1. examiner and reviewer: Prof. Dr. Albrecht E. Melchinger
2. examiner and reviewer: Prof. Dr. Hermann Bürstmayr
3. examiner: apl. Prof. Dr. Thomas Miedaner

Contents

Abbreviations

BC	backcross
DH	doubled-haploid
DON	deoxynivalenol
DS	days to silking
GBLUP	genomic best linear unbiased prediction
GEBV	genomic estimated breeding value
GER	Gibberella ear rot
GP	genomic prediction
GS	genomic selection
GWAS	genome-wide association studies
IBD	identical by descent
IBS	identical by state
kb	kilobase pair
LD	linkage disequilibrium
MAGIC	multi-parent advanced generation intercross
MAS	marker-assisted selection
NAM	nested association mapping population
NIRS	near-infrared spectroscopy
PS	prediction set
QTL	quantitative trait locus/loci
SNP	single-nucleotide polymorphism
SSR	simple-sequence repeats
TS	training set
ZON	zearalenone

General introduction

Gibberella ear rot (GER) of maize (*Zea mays* L.) is a devastating fungal disease in Europe and North America, since it reduces grain yield and contaminates grains with mycotoxins (Martin et al. 2011; Zila et al. 2013; Kebede et al. 2016). The most prevalent fungal species in Europe is *Fusarium graminearum* Schwabe, with symptoms as red or pink mold covering the tip and ear (Fig. 1, Bolduan et al. 2009). To protect public health, the European Union enacted a law regulating maximum concentrations for certain mycotoxins in food (European Commission 2006). In European countries such as Austria, the level of GER resistance is one criteria for registration of new maize varieties, which further underlines the importance of containing GER in maize. As agronomic practices and post-harvest measures have limited effect to contain the infection, development of genetically resistant cultivars is the most effective and ecologically friendly approach (Martin et al. 2011; Zila et al. 2013). From quantitative genetics perspective, large genetic variation and high heritabilities of complex GER resistance-related traits observed in elite European maize germplasm suggest prospects of successful accumulation of resistance alleles through breeding, whilst retaining good agronomic performance (Bolduan et al. 2009; Martin et al. 2011, 2012).

Fig. 1 *Fusarium* artificially infected ears (left) and non-infected ear (right) of maize in the breeding program of the University of Hohenheim, Stuttgart, Germany

To date, three selection approaches have been widely and routinely implemented in plant breeding sector to select genotypes with superior agronomic characteristics such as yield, disease resistance and drought tolerance, namely phenotypic selection, marker-assisted selection (MAS) and genomic selection (GS, Melchinger et al. 1998; Riedelsheimer et al. 2013; Technow et al. 2014a). In brief, phenotypic selection entails that direct field evaluations of various traits for genotypes (*i.e.*, lines or hybrids) are conducted in multiple locations of interest and across multiple years in order to select outperforming genotypes that perform stably across different environments. In comparison, without field testing, MAS and GS indirectly screen individuals solely based upon molecular markers (*e.g.*, SNP markers), for which associations with phenotypic variation of target traits across environments have been identified beforehand with respective statistical analysis in a training population. The major difference between MAS and GS resides in the number of involved markers. MAS relies on usually a few markers with statistically significant and large effects on phenotypic variation, termed QTL (quantitative trait locus/loci, depending on the context). In comparison, GS utilizes all genome-wide markers regardless of the effect size to estimate the breeding value of individuals (termed genomic prediction GP), upon which selection is subsequently performed (Fig. 2, Meuwissen et al. 2001; Heffner et al. 2009).

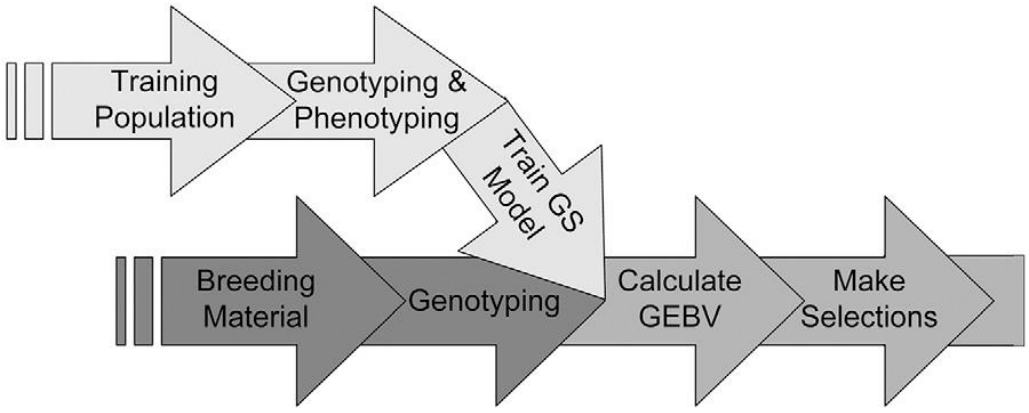

Fig. 2 Diagram of genomic selection (GS) processes starting from construction of the training population for building the GS model, followed by calculating the genomic estimated breeding value (GEBV) for untested but genotyped breeding materials with the obtained GS model, based on which selection is subsequently performed (taken from Heffner et al. 2009)

The concept of expected genetic gain per unit time is a standard criterion to evaluate and compare different breeding and selection schemes in the era of molecular breeding (Falconer and Mackay 1996; Heffner et al. 2010; Xu et al. 2017). It is defined as: $R = i\ r\ \sigma_A/t$, where R is the response to selection or selection gain or genetic gain per unit time, i is the selection intensity (mean deviation of selected individuals in units of the standard deviation of the selection criterion), r is the correlation between the selection criterion and the true breeding value, σ_A is the standard deviation of breeding values (Falconer and Mackay, 1996) or genetic standard deviation (the square root of the additive genetic variance), and t is the time length to complete one breeding cycle. In the context of phenotypic selection, r equals to the square root of narrow-sense heritability (h), whereas with respect to MAS or GS, r is the prediction accuracy, *i.e.*, the correlation between true breeding values and QTL-based estimated breeding values or genomic estimated breeding values (GEBVs). A large number of studies reported higher selection accuracy for GS than MAS, regardless of the trait thanks to the deployment of genome-wide markers, and results in higher genetic gain, whereas MAS can be more cost-effective for traits with high heritabilities and large-effect QTL identified (Heffner et al. 2009; Jannink et al. 2010; Riedelsheimer et al. 2013). Compared with phenotypic selection, MAS and GS are able to accelerate the breeding process through shortening the generation interval and/or largely increasing the selection intensity. Previous empirical studies identified QTL with significant effects for GER resistance in European maize germplasm, which implies feasibility and efficiency of applying MAS and/or GS to improve this complex trait (Martin et al. 2011, 2012; Riedelsheimer et al. 2013). Moreover, MAS and GS are less resource-consuming than field testing, considering the continuous drop of genotyping costs. This is particularly appealing for disease resistance breeding that requires laborious field disease evaluation and oftentimes expensive laboratory analyses.

Multi-parental QTL mapping and genome-wide association studies

To implement MAS, genetic analysis of relevant traits are prerequisites, including identification of QTL, precise estimation of QTL effects and positions in the genome as well as investigation of epistasis (Melchinger et al. 1998; Bernardo 2008). Linkage mapping (*i.e.*, QTL mapping) has been a routine approach

in the past decades to dissect the genetic architecture of complex traits in crops (Holland 2007). The classical scheme of bi-parental QTL mapping comprises three steps. First, a bi-parental population is generated by crossing two phenotypically and genetically divergent inbred parental lines and, afterwards, F1 progenies are selfed for 6 to 7 generations to obtain a relatively large number (>100) of homozygous inbred lines with high phenotypic variation of the target trait and segregation at a large number of loci. Depending on the crop and research interest, the target trait can be line per-se performance or testcross performance in relation to testers from the opposite heterotic pool (Melchinger et al. 1998). Alternatively for maize, pure-breeding inbred lines can be developed much faster and more efficient with the doubled-haploid (DH) technology that has been largely adopted in public and commercial maize breeding programs (Prigge and Melchinger 2012; Melchinger et al. 2013; Melchinger et al. 2016; Chaikam et al. 2019), by which homozygous lines are developed within two generations. Second, the segregating population is evaluated for phenotypic performance in the field or greenhouse and in parallel genotyped with molecular markers, *e.g.*, SSR, SNP. If the target trait is testcross performance, the population is crossed with one or multiple testers to produce testcross hybrids that are subsequently evaluated. Third, associations between phenotypic variation and segregating markers are investigated with proper statistical methods (*e.g.*, composite interval mapping) to detect QTL, localize these on the genetic map and estimate their effects.

This approach has often been criticized for several drawbacks: (i) The allelic diversity and the number of segregating loci in the mapping population are relatively limited due to narrow representation of the breeding germplasm by only two parental lines (Xu 1998; Liu and Zeng 2000). (ii) Different QTL and QTL effects for the same trait were observed between different mapping populations (Beavis 1998; Melchinger et al. 1998). (iii) The population size, that has large influence on the power of QTL detection, is usually constrained because of practical reasons. To mitigate these limitations, multi-parental QTL mapping was proposed to detect QTL jointly from multiple bi-parental families that, as a whole, possess higher allelic diversity, more complex genetic background and larger population size (Jansen et al. 2003; Blanc et al. 2006; Bink et al. 2012). These families are routinely tested and genotyped in practical breeding programs (Bardol et al. 2013). Alternatively, they can be generated in research programs with sophisticated

mating designs, *e.g.*, nested association mapping population (NAM, Yu et al. 2008), diallel design (Blanc et al. 2006) and the multi-parent advanced generation intercross design (MAGIC, Huang et al. 2015). Depending on the number of parents shared between families, they are inter-connected at various levels.

While the set of QTL segregates in only one family for the bi-parental QTL mapping approach, the multi-parental QTL mapping approach captures QTL that segregate in several families and enables testing for QTL × genetic background (family) interactions (Blanc et al. 2006). According to Li et al. (2011) and Ogut et al. (2015), the two mapping approaches are complementary: for a given mapping population size, the bi-parental QTL mapping approach is generally more powerful to identify rare but large-effect QTL that have low frequencies in the germplasm and segregate in a limited number of families, whereas the multi-parental QTL mapping approach is more capable of detecting small-effect QTL if shared by a large number of families. It is of high interest to investigate which approach is more suitable for dissecting the genetic architecture of GER resistance traits in European maize germplasm.

Five different biometric models can be applied for QTL mapping with multiple families in plants, which differ in the underlying assumptions on transferability of QTL effects over parents or families (Table 1): (i) The classical single-family model can be applied within each family, assuming that the marker effect is specific to each family (Model 1, Blanc et al. 2006). (ii) The QTL allele effect is very much impacted by family genetic background and varies between different families and, therefore, it is modeled as being nested within family in the joint analysis (Model 2, disconnected model, Jannink and Jansen 2001; Blanc et al. 2006). (iii) The QTL allele in each parent has a unique effect that persists across different progeny families sharing this parent and, thus, the number of QTL alleles at one QTL locus equals to the number of inbred parental lines (Model 3, connected model, Jannink and Jansen 2001; Blanc et al. 2006). (iv) Genome segments of different parents that trace back to the same ancestor line (*i.e.*, are identical by decent IBD) have the same alleles and effects, regardless of family genetic background (Model 4, LDLA model, Jansen et al. 2003; Bardol et al. 2013; Giraud et al. 2014). (v) Genome segments of different parents that have identical nucleotide sequences (*i.e.*, are identical by state IBS) carry the same alleles with effects identically

expressed in different families (Model 5, LDLA-1-marker model, Bardol et al. 2013; Yu et al. 2008; Würschum et al. 2012). Since the number of parameters to be estimated at each locus is reduced from Model 2 to Model 5, the QTL detection power is expected to increase (Rebai and Goffinet 1993, 2000). However, in experimental studies, the performance ranking of these models depends on traits and populations (Blanc et al. 2006; Steinhoff et al. 2011; Bardol et al. 2013; Giraud et al. 2014).

Table 1 Overview of the five biometric models for detecting QTL with single or multiple families in Han et al. (2016)

Model	Model name	QTL set	Number of alleles at a QTL	Reference(s)
1	Single-family model	One per family	Two per family	Blanc et al. (2006)
2	Disconnected model	Joint	Two × number of families	Blanc et al. (2006)
3	Connected model	Joint	Number of parents	Blanc et al. (2006)
4	Linkage disequilibrium and linkage model (LDLA)	Joint	Number of clustered ancestral alleles in IBD	Bardol et al. (2013), Giraud et al. (2014)
5	LDLA-1-marker model	Joint	Number of marker alleles in IBS, *i.e.*, two (SNP) or at least two (SSR)	Würschum et al. (2012), Giraud et al. (2014)

Different from linkage mapping, QTL detection with genome-wide association studies (GWAS) is conducted on a large collection of genetically diverse lines that can be chosen from various sources or genetic backgrounds. These lines are phenotyped in multiple environments and genotyped with a large number of markers across the entire genome. In comparison with linkage mapping populations, the main merits of GWAS populations include larger allelic diversity, higher recombination frequency between adjacent markers and being closer to the situation encountered in practice (Yan et al. 2011). Therefore, one could expect that, in relation to linkage mapping, GWAS has greater power for QTL detection, higher map

resolution and better transferability of QTL results across different genetic backgrounds. However, since GWAS populations comprises diverse genetic materials, the presence of population structure and cryptic relatedness could cause confounding effects, leading to spurious positive signals in screening QTL across the genome with statistical models (Sillanpää 2011). This is one of the major challenges in implementing GWAS and for its solution, the unified mixed-model approach has been developed and widely applied with various options of corrections for confounding effects (Table 2 and Formula 1, Yu et al. 2006; Stich et al. 2008). Investigating the genetic architecture of GER resistance traits with various GWAS models could enlighten the importance of the genetic background on the expression of these traits and also assess the effectiveness of these mixed models for correcting confounding effects. To our knowledge, our study is the first GWAS conducted for GER resistance traits in European maize germplasm.

Table 2 GWAS models with various levels of corrections for population structure and cryptic relatedness

Model	Correction for population structure	Correction for cryptic relatedness
ANOVA	No	No
Q3 and Q10 models	Take the first three and ten principal components as covariates, respectively	No
K model	No	Use a realized kinship matrix estimated from markers as variance-covariance matrix of random genotype effect
Q3 + K and Q10 + K models	Take the first three and ten principal components as covariates, respectively	Use a realized kinship matrix estimated from markers as variance-covariance matrix of random genotype effect

All these mixed models can be expressed as

$$\boldsymbol{y} = \mathbf{1}\mu + \boldsymbol{Q\beta} + \boldsymbol{S}\alpha + \boldsymbol{Zu} + \boldsymbol{e}, \quad (1)$$

where **y** is a vector of adjusted entry means; μ is the overall mean; $\boldsymbol{\beta}$ is a vector of effects of fixed covariates correcting for population structure and Q is a design matrix relating **y** to $\boldsymbol{\beta}$; α is the effect of the SNP under test and $\boldsymbol{S}$ is a vector referring to the number of minor alleles (*i.e.*, 0, 1, 2) of each genotype at this SNP locus; $\boldsymbol{u} \sim N(0, \boldsymbol{A}\sigma_g^2)$ is a vector of random polygenic background effects, $\boldsymbol{A}$ is the genomic relationship matrix of the lines and σ_g^2 is the additive genetic variance in the respective heterotic pool; $\boldsymbol{Z}$ is a design matrix of 1s and 0s relating **y** to $\boldsymbol{u}$; and $\boldsymbol{e}$ is the vector of residuals. In the ANOVA model, $\boldsymbol{Q\beta}$ and $\boldsymbol{Zu}$ were excluded, whereas $\boldsymbol{Q\beta}$ was not included in the K model.

Genomic selection and training set design

In recent years, plant breeding technologies have undergone a revolutionary transformation in academia and industry, thanks to the dramatic reduction of genotyping cost as well as advancements of statistical models and computational capability. It centres on development and implementation of GS, which mainly concerns prediction for performance of individuals solely based on their genotypes (*i.e.*, GP) without field testing (Fig. 2, Meuwissen et al. 2001; Jannink et al. 2010). Compared with MAS that mainly relies on major QTL with large effects, GS leveraging all genome-wide markers has the merit of further accounting for QTL with minor effects and, therefore, has higher prediction accuracy (Meuwissen et al. 2001; Schopp et al. 2017). It was illustrated in empirical studies to be a powerful approach for predicting complex traits, for instance, *Fusarium* ear rot resistance in maize (Riedelsheimer et al 2013), *Fusarium* head blight resistance in wheat (Michel et al. 2021; Rutkoski et al. 2012) and barley (Lorenz et al. 2012). In addition, thanks to the *in vivo* haploid induction technology and oil content-based haploid sorting system, large DH populations can be effectively and routinely generated in maize breeding programs (Prigge and Melchinger 2012; Melchinger et al. 2013; Melchinger et al. 2015; Hu et al. 2016). This substantially reduces the time

duration and costs for developing a large number of inbred lines and, meanwhile, makes it possible to largely increase selection intensity together with the implementation of GS.

To perform prediction for a set of untested individuals (*i.e.*, prediction set (PS)), a training set (TS) consisting of a number of phenotyped and genotyped individuals should be put in place beforehand in order to derive a prediction equation that essentially explores the association between phenotypic variation and genotypic variation at DNA marker level. One important question in GS, and generally in marker-based prediction such as MAS, is how to design the TS for the purpose of obtaining a high prediction accuracy (*i.e.*, correlation between predicted and true breeding values). Major influential factors include the size of the TS, the number of families in the TS as well as the genomic relatedness between TS and PS (Riedelsheimer et al. 2013; Lehermeier et al. 2014). One approach often applied in practice for increasing the TS size is to combine different groups or families of small to mediate size, which are historically and routinely generated in breeding programs and already phenotyped and genotyped (Lorenz et al. 2012; Technow et al. 2013). This approach has been evaluated across a wide range of genetic materials and traits with animals and plants, however prediction accuracy was not always increased (Hayes et al. 2009; De Roos et al. 2009; Lund et al. 2014; Technow et al. 2013; Guo et al. 2014; Lehermeier et al. 2014; Lorenz and Smith 2015).

To explain this inconsistency, four factors and their interdependency were studied in the literature, however to our knowledge a systematic and comprehensive investigation of all these factors is still warranted. These factors include (i) similarity of the segregating QTL across populations that corresponds to inter-population genetic relationship at the QTL level and therefore positively impacts prediction accuracy across populations (Goddard 2009; de los Campos et al. 2013; Wientjes et al. 2015; Schopp et al. 2017), (ii) consistency of the linkage disequilibrium (LD) pattern (*i.e.*, both extent and phase of LD) across populations, because LD between markers reflect that between markers and unknown QTL upon which GP relies, and inconsistent LD pattern between populations leads to low or even negative prediction accuracy of GP across populations (Goddard et al. 2006; De Roos et al. 2008), (iii) genomic relationship between the

TS and PS that is associated with inter-population similarity of segregating QTL, consistency of the LD pattern and resemblance of genetic backgrounds, (iv) statistical models that rely on different assumptions about the marker effects, including the genomic best linear unbiased prediction (GBLUP) model, the Bayesian variable selection models (*e.g.*, BayesB) and the multi-trait GBLUP model (Habier et al. 2007, 2013; Lehermeier et al. 2015; Schulz-Streeck et al. 2012; Olson et al. 2012; Zhou et al. 2013). A still open question was whether increasing the training set size by combining dent and flint materials can improve the prediction accuracy for GER resistance traits in European maize germplasm.

Objectives

The overall aim of this dissertation was to dissect the genetic architecture of GER resistance related traits in European maize germplasm with complementary QTL mapping approaches, including bi-parental QTL mapping, multi-parental QTL mapping and GWAS, and to investigate effective implementation of QTL-based and genomic selection approaches by better designing the TS. Furthermore, inexpensive measurements of mycotoxin concentrations such as GER severity visual rating scores and near-infrared spectroscopy (NIRS) measurement were studied as proxy for the expensive assays of mycotoxins themselves. In particular, our objectives were to

(1) Detect QTL and estimate QTL effects with five flint interconnected bi-parental families using classical single-family model and four multi-family models,
(2) Examine various TS compositions for QTL-based prediction that have different levels of relatedness with the PS and different population sizes,
(3) Identify QTL with GWAS in two diversity panels of European dent and flint DH lines,
(4) Provide insights in optimizing TS composition and size for GP by comparing scenarios with and without including lines from the opposite pool,
(5) Evaluate the reliability of GER severity visual scoring and NIRS measurements for predicting mycotoxin concentrations in grain maize.

2. Choice of Models for QTL Mapping with Multiple Families and Design of the Training Set for Prediction of *Fusarium* Resistance Traits in Maize

Sen Han[1], H. Friedrich Utz[1], Wenxin Liu[2], Tobias Schrag[1], Michael Stange[1,†], Tobias Würschum[3], Thomas Miedaner[3], Eva Bauer[4], Chris-Carolin Schön[4], Albrecht E. Melchinger[1]

[1] Institute of Plant Breeding, Seed Science and Population Genetics (350a), University of Hohenheim, D-70593 Stuttgart, Germany;

[2] Crop Genetics and Breeding Department, China Agricultural University, Beijing 100193, China;

[3] State Plant Breeding Institute (720), University of Hohenheim, D-70593 Stuttgart, Germany;

[4] Department of Plant Breeding, Center of Life and Food Sciences Weihenstephan, Technische Universität München, D-85350 Freising, Germany;

† *Present address:* Strube Research GmbH & Co. KG, Hauptstraße 1, D-38387 Söllingen, Germany

Theor Appl Genet 129(2):431-44

Doi: 10.1007/s00122-015-2637-3

Article published: Feb. 2016

The original publication is available at:

http://link.springer.com/article/10.1007%2Fs00122-015-2637-3

Abstract

Recently, QTL mapping has shifted from analysis of single families to multiple, connected families and several biometric models have been proposed. Using a high-density consensus map consisting of 2,472 marker loci, QTL mapping was performed with five connected bi-parental families with 639 maize flint doubled-haploid (DH) lines for ear rot resistance and analyzed traits DON, Gibberella ear rot severity (GER), and days to silking (DS). Five biometric models were compared, which are different in the assumption about the number and effects of alleles at QTL. Model 2 to 5 performing joint analyses across all families and using linkage and/or linkage disequilibrium (LD) information detected all and even further QTL than Model 1 (single family analyses) and generally explained a higher proportion p_G of the genotypic variance for all three traits. For DON and GER, QTL were mostly family-specific, but several QTL for DS were shared by multiple families. Many QTL had large additive effects and most alleles increasing resistance originated from a resistant parent. Interactions between detected QTL and genetic background (family) were rarely identified and effects were comparatively small. Detailed analysis of three fully connected families obtained higher p_G values for Model 3 or 4 than for Model 2 and 5, irrespective of the size N_{TS} of the training set (TS). In conclusion, we recommend Model 3 and 4 for QTL-based prediction with larger families. Including a sufficiently large number of full sibs in the TS resulted in increase of QTL-based prediction accuracy (r_{VS}) for various scenarios differing in the composition of the TS.

3. Genomic prediction and GWAS of Gibberella ear rot resistance traits in dent and flint lines of a public maize breeding program

Sen Han[1, †], Thomas Miedaner[2], H. Friedrich Utz[1], Wolfgang Schipprack[1], Tobias A. Schrag[1], Albrecht E. Melchinger[1,]

[1] Institute of Plant Breeding, Seed Science and Population Genetics (350a), University of Hohenheim, D-70593 Stuttgart, Germany;

[2] State Plant Breeding Institute (720), University of Hohenheim, D-70593 Stuttgart, Germany;

†*Present address:* Monsanto Holland B.V., Leeuwenhoekweg 52, 2661 CZ Bergschenhoek, The Netherlands

Euphytica

Doi: 10.1007/s10681-017-2090-2

Article published: Jan. 2018

The original publication is available at:

https://link.springer.com/article/10.1007/s10681-017-2090-2

Abstract

Gibberalla ear rot (GER) is a large threat to European maize, reducing grain yield and contaminating grain with mycotoxins. Genomic prediction (GP) has great potential to accelerate GER resistance breeding. However, a large training set (TS) comprising both phenotyped and genotyped individuals is essential for GP to obtain high prediction accuracy (ρ), which imposes major challenges for small-size populations. A possible solution would be combining small-size populations. However, genetic heterogeneities between populations with regards to segregating QTL, linkage disequilibrium (LD) pattern and genomic relationships can impair ρ of this approach. In this study, genetic architectures were investigated for traits GER severity, deoxynivalenol concentration (DON) and days to silking (DS) with genome-wide association studies (GWAS) performed separately within the European dent and flint diversity panels. Moreover, we assessed the consistency of LD pattern between heterotic pools as well as genomic relationship within and across pools. Subsequently, we compared four GP approaches with cross-validation, which composed the TS with lines from single or combined pools and deployed statistical models assuming marker effects identical or different but correlated between pools. For DON, two and six QTL were identified within the dent and flint pool, respectively; however, none was in common. The LD pattern was consistent between heterotic pools for marker pairs less than 10 kb apart. GP between pools resulted in low or even negative ρ. Combined-pool GP did not yield higher accuracy than within-pool GP regardless of statistical models. Our findings underline the importance of assessment of genetic heterogeneities between populations prior to implementing GP using a combined TS.

4. Prediction of deoxynivalenol and zearalenone concentrations in Fusarium graminearum inoculated backcross populations of maize by symptom rating and near-infrared spectroscopy

Thomas Miedaner[1], Sen Han[2], Bettina Kessel[3], Milena Ouzunova[3], Tobias Schrag[2], Friedrich H. Utz[2], Albrecht E. Melchinger[2]

[1] University of Hohenheim, State Plant Breeding Institute, Fruwirthstr. 21, 70599 Stuttgart, Germany

[2] University of Hohenheim, Institute of Plant Breeding, Seed Science and Population Genetics, Fruwirthstr. 21, 70599 Stuttgart, Germany

[3] KWS SAAT AG, Grimsehlstr. 31, 37555 Einbeck, Germany

Plant Breeding

Doi: 10.1111/pbr.12297

Article published: Aug. 2015

The original publication is available at:

https://onlinelibrary.wiley.com/doi/abs/10.1111/pbr.12297

Abstract

Gibberella ear rot (GER) caused by *Fusarium graminearum* is a destructive disease in maize in temperate regions causing not only yield reduction but also mycotoxin contamination, especially by deoxynivalenol (DON) and zearalenone (ZON). In EU, both mycotoxins are subjected to strict legislative limits. The objective of this study was to evaluate whether prediction of DON and ZON concentrations is feasible either by GER severity visual rating score or by near-infrared spectroscopy (NIRS). We analyzed 80 and 102 lines developed by backcrossing resistant doubled-haploid lines selected from segregating populations to one resistant and one susceptible parent, respectively. Artificial infection was conducted at three locations in Germany and France. Traits evaluated included GER severity visual rating score defined as percentage of affected ear as well as DON and ZON concentrations measured by immunoassays. All samples were additionally measured with NIRS. We observed that both backcross (BC) populations had substantial difference in their means. Genotypic variances within each BC population were significantly larger than zero ($P < 0.01$) in all circumstances. Within each BC population, DON and ZON concentrations measured by immunoassays were significantly ($P < 0.01$) correlated with each other and with GER severity visual rating score ($0.6 \leq r \leq 0.9$, $P < 0.01$). DON concentration measured by an immunoassay and by NIRS also had correlation of $r = 0.87$ and $r = 0.90$ for the BC population with the resistant and susceptible parent, respectively. In conclusion, DON and ZON concentrations could be reliably predicted by GER severity visual rating score. Additional NIRS analysis of DON concentration might be useful for the positively selected fraction of genotypes.

5. General discussion

Given a fixed budget and time duration, plant breeders are mainly concerned with the question: how to improve current breeding schemes in order to obtain higher selection gain? To answer this question in the context of GER resistance breeding for European maize with the dent and flint heterotic pools, in-depth studies were conducted in this dissertation on classical phenotypic selection as well as two modern molecular breeding technologies, namely MAS and GS. With the deployment of DNA markers, the latter two have revolutionized plant breeding practices in the past decades, particularly for improvement of polygenic and complex traits such as yield, quality and disease resistances controlled by many genes and significantly influenced by environmental conditions. More specifically, the aim was to answer the following questions: (i) Can GER resistance be potentially improved via breeding in European maize germplasm? (ii) Do mycotoxin concentrations measured by immunoassays correlate well with GER severity visual rating scores and NIRS measurements? (iii) Does multi-parental QTL mapping perform better than classical bi-parental QTL mapping in terms of QTL detection and QTL effect estimation for GER resistance related traits? (iv) Can additional QTL be identified with GWAS in diversity panels of dent and flint lines? (v) How to better design the training population for the multi-parental QTL mapping approach and GS with respect to composition and population size, in order to achieve high prediction accuracy?

Gibberella ear rot, genetic variation, heritability and genetic correlation

Fusarium species are common pathogens on maize ears and stalks, reducing yield and contaminating grains with mycotoxins that jeopardize safety of human food and animal feed. Therefore, the Europe Union has set up a maximum level of intake for mycotoxins and *Fusarium* resistance is considered for new maize variety registration in European countries such as Austria. One of the most predominant species in Central Europe is *F. graminearum*, one major pathogen causing Gibberella ear rot (Pfordt et al. 2020). Its primary ear infection pathway is through the silk channel during the first 6 to 10 days after silking that normally takes place between beginning and mid of July in Central Europe. Studies have shown that the range of local *Fusarium* species is primarily impacted by weather conditions during vegetation period

and, particularly, GER pathogens, including *F. graminearum* and *F. culmorum*, are favored by high precipitation and low temperature during silking (Pfordt et al. 2020). In this dissertation, an aggressive isolate of *F. graminearum* (IFA66) was applied to artificially inoculate our materials, which were tested in multiple locations. Relevant traits such as visual rating of GER severity, mycotoxin deoxynivalenol (DON) and zearalenone (ZON) concentrations in grain as well as days to silking (DS) were measured and subject to phenotypic and genotypic data analysis.

In this dissertation, three different sets of materials were tested and analyzed for different objectives. In Han et al. (2016), five flint inter-connected DH families sharing parents were generated in an incomplete half-diallel design with four inbred lines developed by the University of Hohenheim, among which UH006 (R2) is highly resistant against GER, UH007 (R1) is moderately resistant, and UH009 (S1) and D152 (S2) are highly susceptible (Bolduan et al. 2009). The family size had an average of 128 and ranged from 43 to 204. Consistent with expectation from GER resistance levels of the parents, family R1R2 had the lowest means ($\bar{X}$) of GER and DON, followed by the family R2S1. This result confirms the high interest of using R1 and R2 lines as resistant donors for developing new resistant lines. For DS, the differences in $\bar{X}$ between families were small, suggesting similar maturities of these families. The observed significant ($P < 0.01$) genotypic variances for all traits in all families and generally high heritabilities with an average of 0.78 for DON and GER are in line with previous studies on GER resistance in European maize Germplasm (Martin et al. 2011, 2012; Bolduan et al. 2009; Riedelsheimer et al. 2013). These indicate good prospects of resistance breeding and phenotypic selection against GER across different environments in European maize germplasm. In addition, the extremely high genetic correlations between DON and GER ($r_g \geq 0.96$) for almost all families imply that visual evaluation of GER severity with low cost is a reliable indicator for DON, which enables large-scale screening of materials and increasing selection pressure. Moreover, the moderately negative genetic correlations ($-0.66 \leq r_g \leq -0.21$) for DS with GER and DON show that early-flowering genotypes have higher chance to be infected by GER and accumulate more mycotoxins than late-flowering genotypes.

In the study of Han et al. (2018), two diversity panels consisting of 130 European dent lines and 114 European flint lines, respectively, were evaluated in two years (2010 and 2012) at two locations in Germany. These lines with diverse genetic background and compositions represent the elite maize breeding materials of the University of Hohenheim as detailed by Westhues et al. (2017). For GER and DON, dent and flint lines were not significantly ($P \leq 0.05$) different in means, although dent lines flowered on average about one week later than the flint lines and both traits showed slightly negative correlations with DS within each pool. Similar to Han et al. (2016), these two diversity panels showed significant genotypic variances for all traits, moderately high heritabilities for GER and DON ($0.73 \leq h^2 \leq 0.78$) and tight genetic correlations between GER and DON ($r_g \geq 0.92$). Interestingly, estimates of the genomic correlations between both pools were around zero for DON and DS, and slightly higher for GER, which can be explained by the large genetic divergence between the two pools and the history of separate and reciprocal development of dent and flint lines for the purpose of exploring heterosis in a hybrid breeding scheme (Reif et al. 2005).

In the study of Miedaner et al. (2015), a population of 200 inbred lines was tested, including two backcross families derived from the recurrent parent lines UH007 (resistant) and UH009 (susceptible), respectively. In line with expectation, the means of GER, DON and ZON concentrations were significantly different between the two BC families, which can be explained by UH007 and UH009 carrying different QTL alleles for GER resistance. In both BC populations, the large and positive genetic correlations between mycotoxin concentrations measured with immunoassays and GER severity visual rating scores and NIRS measurements imply that mycotoxins can be indirectly and reliably reduced by directional selection for GER severity ratings and NIRS measurements that are less expensive and less laborious.

Multi-parental QTL mapping, GWAS and genetic architecture of GER resistance in European maize

The classical approach for analyzing the genetic architecture of quantitative traits in maize is QTL mapping with bi-parental populations, which are derived from crossing two highly diverse parents and

therefore embrace a large number of segregating QTL as well as large phenotypic variance of the target trait (Edwards et al. 1987; Lander and Botstein 1989). However, several problems were reported in the literature for this approach: (i) Estimated QTL effects are often highly inflated (Utz and Melchinger 1994; Beavis 1998; Xu 2003). (ii) QTL detected in one population are oftentimes not confirmed in other populations (Melchinger et al. 1998). (iii) The power of QTL detection is limited, when the number of lines in the mapping population is low (Schön et al. 2004). To overcome these constrains, the multi-parental QTL mapping approach was investigated in this dissertation by conducting QTL mapping in a joint population of multiple bi-parental families, which has higher allelic diversity, more complex genetic background and larger population size than a single bi-parental family. It also has practical advantages, because normally this type of family is routinely generated, genotyped and evaluated in commercial breeding programs and no extra cost is needed to apply this approach on already available data to obtain potentially better QTL results. In addition, cross-validation proposed by Utz et al. (2000) to use separate sets of materials for QTL detection, corresponding to model selection, and QTL effects estimation was applied and adapted to the multi-parental population scenario in the present study to obtain unbiased estimation of QTL effects and compare various biometric models.

Using a high-density consensus map with 2,472 marker loci, Han et al. (2016) compared five different biometric models with the aim to shed light on the choice of the best fitting model for this approach as well as to study the genetic architecture of GER resistance related traits including digenic and QTL × genetic background epistasis. Among these models, Model 1 was the classical bi-parental QTL mapping model, namely the composite interval QTL mapping model. It was computed separately for each family, except the smallest family R2S2 with only 43 lines, whereas Model 2 to 5 were computed jointly for all families with different assumptions and parameterizations of QTL alleles (Table 1). The analyzed traits included GER, DON and DS. For GER and DON, we identified with Model 1 only a proportion of the QTL detected previously with the same families (Martin et al 2011, 2012), because a more stringent significance level (α = 2% *vs.* 15%) in permutation tests was applied. All detected QTL were adjacent to the flanking markers of QTL reported previously, indicating that basically the same QTL were identified.

One interesting finding with Model 1 was that QTL detected in each family for DON were family-specific and rarely consistent across families, except one common QTL for DON between families R1R2 and R2S1 that explained a high proportion of the genotypic variance (> 20%) and inherited the favorable resistance allele from the same highly resistant parent line R2 (UH006). The results of QTL for GER showed a similar picture as for DON. This is in agreement with the multi-parental QTL mapping study of Blanc et al. (2006) that congruent QTL between different populations often have large effects and originate from a shared parental line. In comparison, QTL for DS were more commonly shared between families, suggesting different genetic architecture for DS and GER or DON. One explanation similar to the hypothesis proposed by Kemper et al. (2015) is that the adaptation and maturity related trait DS has undergone a longer period of artificial selection and thus QTL with large effects are most likely fixed and the remaining small-effect QTL are segregating and commonly shared between different germplasm of similar maturity, whereas the history of GER resistance breeding is shorter and large-effect QTL are still present but favorable alleles have low frequency and are mainly embraced by a limited number of materials.

One major challenge in conducting QTL mapping jointly on multiple families is parameterization of QTL allele effects in biometric models based on different assumptions (Table 1). Accordingly, the expression of a QTL allele can be determined by its nucleobase or nucleotide sequence (Model 5), ancestral origin or haplotype (Model 4), donor parental line (Model 3) and/or interactions with population genetic background (Model 2). From a statistical point of view, the number of parameters to be estimated at each locus was reduced from Model 2 to Model 5 and, therefore, QTL detection power was expected to increase given the same size of the mapping population. With all 5 families included in the joint analysis, Model 2 to Model 5 detected all QTL of all traits previously identified by Model 1 as well as additional QTL, all together explaining a high proportion of genotypic variances (34.4 - 58.3%). This showed clear advantages of the joint analysis over the single-family analysis that are attributable to a larger mapping population, more segregating loci and more replicates of QTL genotypes, particularly when QTL are more commonly shared between families such as DS. To ensure unbiased estimation of QTL effects and a fair comparison of biometric models, cross-validation was further applied in this study, in which detection of QTL and

estimation of QTL effects were conducted in a so-called training set (TS) with the same size for all models and the same number of randomly selected lines from each of three inter-connected families (*i.e.,* R1R2, R1S1, R2S1), except for a single-family analysis Model 1 with lines from only one family, and the obtained QTL results were afterwards deployed to predict the performance of lines in a so-called prediction set (PS) that had the same size for each model and was formed in the same way with the remaining lines. Prediction accuracy for each family in the PS was calculated and compared for these five models. Different from expectation and the finding of Ogut et al. (2015), with relatively small sizes of TS and PS (81 and 48, respectively), prediction for almost all traits and families was less accurate with joint analysis models than with the single-family analysis Model 1. This implies that given a fixed budget for developing a fixed and relatively limited number of lines for a mapping population, it is better to construct a classical bi-parental population, especially for traits with rare but large-effect QTL in the germplasm such as GER and DON. When increasing TS and PS sizes to 180 and 108, respectively, prediction accuracy of joint analysis models was increased for all traits and families as expected, and Model 3 generally outperformed other models except Model 1. In relation to Model 2, this suggested that the genetic background is less influential for QTL effect expression of the studied traits, which is in agreement with our results of low proportions of the genotypic variance explained by the detected digenic epistasis and QTL × genetic background (family) interactions. The reason why Model 4 was generally not better performing than Model 3 as expected could be that the advantage of reduction of parameters in the model thanks to clustering of ancestral haplotypes was less obvious with only three parents involved. It could be different in a more practical context with a large number of populations generated from a higher number of parental lines. The bi-allelic model, Model 5, did not outperform Model 3 either, which is in line with Lu et al. (2012) and Bardol et al. (2013) that a larger proportion of the genetic variance was captured by multi-allelic models than bi-allelic models.

To dissect the genetic architecture of GER resistance related traits and validate previously detected QTL in more complex and diverse genetic backgrounds, we applied GWAS separately in European dent and flint diversity panels of elite inbred lines from the public breeding program of the University of Hohenheim. Two and six QTL for DON were detected within the dent and flint pool, respectively. Similar

to Han et al. (2016), QTL for DON were all specific to each heterotic pool and not commonly shared between the two pools. This finding is in line with the almost zero genomic correlation estimated for DON between the two pools as well as large genetic distances observed between pools in this study. No overlapping QTL for DON was identified in comparison with those detected from European flint segregating populations by Han et al. (2016) and Martin et al. (2012). Moreover, we did not identify common QTL for DON with the linkage mapping study of a large Canadian RIL maize population (Kebede et al. 2016) as well as a GWAS study with US maize diversity panels (Zila et al. 2013, 2014). Although no QTL was detected for GER, the patterns of Manhattan plots for GER were similar to those for DON, including those regions with significant SNPs for DON. This is consistent with our finding in Han et al. (2016) that many congruent QTL were detected between GER and DON. For DS, we did not identify QTL in the dent and flint pools, which could be due to medium population sizes of these two pools that constrain the power of QTL detection.

In conclusion, the multi-parental QTL mapping models did not perform better than the classical bi-parental QTL mapping model for all traits, when given the same population size and compared with cross-validation. In comparison with QTL results for DON, GER and DS in Han et al. (2016), no common QTL was detected in Han et al. (2018) by performing GWAS within the dent and flint diversity panels. Nevertheless, a few new QTL for DON were detected in Han et al. (2018) for each heterotic pool.

Genetic heterogeneity and training population design for QTL-based and genomic prediction

The underlying genetic composition of the training population, which is used for QTL detection and QTL effect estimation in MAS and establishment of a prediction model in GS, could be different from that of the targeted prediction population at various levels, depending on which materials, families or populations are included in the training population. As revealed by our studies (Han et al. 2016, 2018), some traits such as GER and DON have QTL that are more specific to certain families/pools or even a limited number of donor lines, whereas QTL for traits like DS are more often shared between families and have higher allele frequency in the germplasm. This shows that different traits have different genetic architecture and,

moreover, the genetic composition in terms of presence of QTL can be less consistent between families or heterotic pools for some traits. In addition, Han et al. (2018) thoroughly studied the genetic heterogeneity between the European flint and dent pools and reported that: (i) the flint pool was clearly genetically separated from the dent germplasm based on population structure analysis with SNP markers, which agrees well with previous studies on European maize (Reif et al. 2005; Fischer et al. 2008), (ii) the LD extent (r^2), corresponding to correlations between adjacent markers, had high and similar values in the two heterotic pools, but the signs of r (*i.e.*, LD phase) for the same marker pairs in the two pools were mostly different for marker pairs > 10 kb apart. This result is consistent with Technow et al. (2013) and Lehermeier et al. (2014), who both reported highly inconsistent LD phases between the European dent and flint pools. This means that even when the same QTL are present in the two pools, their effects estimated by markers can be often reverse between the flint and dent lines. These findings pointed out potential factors that impact prediction accuracy of MAS and GS.

The materials and the various levels of inter-connectedness among the five bi-parental families in our study Han et al. (2016) enabled a comprehensive and systematic investigation of a better design of the training population for QTL-based prediction (*i.e.*, MAS) for GER resistance related traits. With the connected model (Model 3) and cross-validation, we evaluated various scenarios of composition of the TS for predicting performance of the PS that comprised one single family. While the population size was the same for these TS compositions, they were different in the number of families included, ranging from 1 to 4, and in the relatedness between the TS and PS, including full-sib (highly related), half-sib (related), unrelated as well as a mixture of these relatedness. In accordance with the GS studies of Riedelsheimer et al. (2013), Lehermeier et al. (2014) and Foiada et al. (2015), we observed that irrespective of traits, TS sizes and the number of families in the TS, higher prediction accuracy can be generally achieved, when the TS is largely composed with materials more closely related to the PS such as full sibs. Furthermore, generally speaking, the prediction accuracy increased with enlarging the TS size, and the rate of increase was larger when more closely related materials were added to the TS. These findings imply that the relatedness between the TS and PS largely and positively impacts the prediction accuracy, which can be explained by (i) more

related materials share more QTL, particularly for rare QTL that segregate in a limited number of families or materials such as GER and DON, (ii) LD pattern including LD extent and phase is more consistent between related materials than less related ones, and (iii) related materials have similar genetic backgrounds, which could have a significant effect on QTL expression because of digenic epistasis and/or QTL × genetic background interactions as observed in this study. Therefore, a golden rule for achieving high selection gain by MAS in practical breeding programs is to warrant a high prediction accuracy by including more related materials in the TS as well as increasing the TS size. Nevertheless, one should keep in mind that the prediction accuracy also depends on the genetic architecture of a particular trait and genetic composition of specific materials, and incorporating more related materials in the TS may not always result in higher prediction accuracy, as we observed similar prediction accuracies across all scenarios of TS compositions for DS in family R1R2. This can be explained by our observation that QTL for DS were more commonly shared between families, and R1S1 harbored both rare and common QTL for DS.

Similar to QTL-based prediction or MAS, TS composition plays an important role in determining the accuracy of GP and, therefore, has substantial impact on its implementation in plant breeding programs (Riedelsheimer and Melchinger 2013). In hybrid maize breeding, parental inbred lines are separately developed within each of the two heterotic pools, for instance flint and dent pools in Europe, with the aim to maximize and utilize heterosis in hybrids between them (Melchinger 1999). However, within-pool line development may limit the size of the TS, when populating the TS with lines from only one pool, and therefore constrain prediction accuracy within each pool. One solution for increase of the TS size is composing TS with lines from both pools, however genetic divergence and heterogeneity between heterotic pools may counterbalance its effectiveness, as reflected by the findings in Han et al. (2016). To investigate this point further for European flint and dent diversity panels, Han et al. (2018) compared with cross-validation four different GP approaches for GER resistance related traits, namely within-pool prediction (WP), across-pool prediction (WP), combined-identical prediction (CI) and combined-different prediction (CD). For the WP and AP approaches, the TS and PS comprised lines from the same (WP) or different (AP) heterotic pools. For the CI and CD approaches, the TS and PS were composed in the same way with the TS

including lines from combined pools and the PS consisting of lines from one of the two pools. However, the CI approach assumed marker effects to be identical across pools, whereas the CD approach assumed different but correlated marker effects between pools. In agreement with previous GP studies for various traits across genetically different populations in maize (Technow et al. 2013; Lehermeier et al. 2014; Schopp et al. 2017) and cattle (Harris et al. 2008; Hayes et al. 2009; Weber et al. 2012), the AP approach yielded very low or even negative prediction accuracy for DON and GER, which could be explained by our GWAS finding that the flint and dent heterotic pools did not have common QTL for DON and the same explanation holds true for GER that is expected to have similar genetic architecture as DON. Furthermore, even though the two pools may share some small-effect QTL for DON and GER that were perhaps not identified by our GWAS due to for instance intermediate population sizes of the two diversity panels, the largely inconsistent LD phase as well as highly separate genetic backgrounds between pools can lead to very different QTL effects in the two genetically divergent pools and, therefore, low transferability of GP model equations across pools. Nevertheless, in line with Lehermeier et al. (2014), we observed higher prediction accuracy for the AP approach for DS, suggesting that the flint and dent pools share more common QTL for DS than for DON and GER.

For all traits and TS sizes, the CI and CD approaches had generally lower or similar prediction accuracy compared with the WP approach, even when the TS sizes were doubled for the CI and CD approaches by adding lines from the opposite pool. This is in agreement with the low genomic correlations between the flint and dent pools for all traits estimated in this study, showing that incorporating genetically distant lines from the opposite pool does not necessarily accumulate useful information in the TS. When increasing the TS size, the prediction accuracy was increased as expected for all traits and GP approaches, however, the increase was marginal for DON and GER compared to DS. Possible reasons are (i) the added materials in the TS were more genetically distant from the PS for DON and GER than for DS, and/or (ii) the effective population size for DON and GER is smaller than for DS (Technow et al. 2014b; Technow et al. 2013; Lorenz et al. 2012).

For the AP, WP and CI approaches, we compared two frequently-used biometric models for GP, namely the GBLUP model and one of the Bayesian variable selection models, the BayesB model (Meuwissen et al. 2001). According to Habier et al. (2007, 2013), the Bayesian models generally outperform the GBLUP model, if the TS is composed of multiple populations that are not closely related. The GBLUP model, which assumes that the effects of the genome-wide markers are taken from a normal distribution with equal variance, is very efficient in capturing relationship information, whereas the Bayesian models rely more on LD information persisting across less related materials (Weber et al. 2012; De Roos et al. 2009). Contrary to expectation, we observed that the BayesB model did not yield higher predication accuracy than the GBLUP model for all traits and TS sizes when applied for the AP, WP and CI approaches, except for the WP approach with the TS size ≤ 92. Furthermore, the CD approach using the multi-trait model (*i.e.*, MG-GBLUP model) was applied to explore and account for the potential correlations of marker effects between pools in the combined-pool TS. Different from our expectation, no significant difference in the prediction accuracy was observed for the CI and CD approaches, which agrees well with low genetic correlations estimated for all traits between the flint and dent pools in our study (Wientjes et al. 2016). A further explanation is that the investigated TS sizes were medium and the power was low to estimate a large number of population-specific parameters for estimation of inter-pool genetic correlations in the CD approach.

In conclusion, in order to achieve high prediction accuracy for QTL-based prediction (*i.e.*, MAS) and GS, the TS should be composed of materials closely related with the PS and the TS size should be large, irrespective of the trait and statistical model.

6. Summary

During the last decades, implementation of molecular markers such as single nucleotide polymorphisms (SNPs) has transformed plant breeding practices from conventional phenotypic selection to marker-assisted selection (MAS) and genomic selection (GS) that are more precise, faster and less resource-consuming. In this dissertation, we investigated these three selection approaches for improving the polygenic trait Gibberella ear rot (GER) resistance in maize (*Zea mays* L.), which is an important fungal disease in Europe and North America leading to reduced grain yield and grain contaminated with mycotoxins such as deoxynivalenol (DON) and zearalenone (ZON).

Three different sets of materials were evaluated in multiple environments and analyzed for different objectives. In the first study, five flint doubled-haploid (DH) families (with size 43 to 204) inter-connected at various levels through common parents, were generated in an incomplete half-diallel design with four parental lines developed by the University of Hohenheim. Significant genotypic variances and generally high heritabilities were observed for all three traits (*i.e.*, GER, DON and days to silking (DS)) in all families, implying good prospects for resistance breeding and phenotypic selection against GER across different environments in European maize germplasm. Genetic correlations were extremely tight between DON and GER and moderately negative for DS with DON or GER, suggesting that indirect selection against GER would be efficient to reduce DON, but maturity should be considered in GER resistance breeding. Using a high-density consensus map with 2,472 marker loci, we compared classical bi-parental mapping of QTL (quantitative trait locus/loci) with multi-parental QTL mapping conducted with joint families and using four different biometric models. Multi-parental QTL mapping models identified all and even further QTL than the bi-parental QTL mapping model conducted within each family. Interestingly, QTL for DON and GER were mostly family-specific, yet multiple families had several common QTL for DS. Many QTL displayed large additive effects and most favorable alleles originated from the highly resistant parent. Interactions between detected QTL and genetic background (family) were rare and had comparatively small effects.

Multi-parental QTL mapping models generally did not yield higher prediction accuracy than the bi-parental QTL mapping model for all traits.

In the second study, two diversity panels consisting of 130 elite European dent and 114 flint lines, respectively, from the University of Hohenheim were evaluated and subject to a genome-wide association study within each pool. Similar to the first study, highly significant genotypic and genotype × environment interaction variances were observed for GER, DON and DS. Heritabilities were moderately high for GER and DON and high for DS in both pools. Estimated genomic correlations between pools were close to zero for DON and DS, and slightly higher for GER. The detected QTL for DON were all specific to each heterotic pool and none of them was in common with previously detected QTL. Furthermore, no QTL was detected for GER and DS in both pools. Genomic prediction (GP) across pools yielded low or even negative prediction accuracy for all traits. When the training set (TS) size was increased by combining lines from both heterotic pools, the combined-pool GP approaches had no higher prediction accuracy than the within-pool GP approach. Different from expectation, method BayesB did not outperform genomic best linear unbiased prediction (GBLUP).

In the third study, we analyzed two backcross (BC) families derived from a resistant and a susceptible recurrent parent. Both BC populations differed substantially in their means for all traits, suggesting that the two recurrent parents have different QTL alleles for GER resistance. Relatively high correlations were observed between DON and ZON concentrations measured by immunoassays and GER visual severity scoring and NIRS (near-infrared spectroscopy) within each BC population. Thus, the mycotoxin content in grain can reliably be reduced by directional selection for GER severity and NIRS measurements that are less expensive and less laborious.

In conclusion, GER resistance in European maize germplasm can be effectively improved through breeding with resistant donor lines. GER visual severity scoring and NIRS measurements were found to be reliable predictors for DON and ZON concentrations in grain. We observed that QTL for GER and DON are mostly specific to a few families or a limited number of materials, whereas QTL for DS are more

commonly shared between families. The multi-parental QTL mapping approach is complementary to the classical bi-parental QTL mapping in that the latter has generally higher power to identify rare but large-effect QTL for traits such as GER and DON, whereas the former is superior in detecting common but small-effect QTL for traits such as DS. Composing the TS with materials more closely related to the prediction set and increasing the TS size generally resulted in higher prediction accuracy for MAS and GS, irrespective of the trait and statistical model.

7. Zusammenfassung

In den letzten Jahrzehnten hat die Anwendung molekularer Markern wie z.B. Single Nucleotide Polymorphismen (SNPs) die praktische Pflanzenzüchtung von der konventionellen phänotypischen Selektion hin zur markergestützten Selektion (MAS) und genomischen Selektion (GS) verändert, da letztere Methoden oft präziser, schneller und weniger ressourcenintensiv sind. In dieser Dissertation untersuchten wir diese drei Selektionsansätze zur Verbesserung des polygenen Merkmals der Gibberella-Kolbenfäule (GER)-Resistenz bei Mais (Zea mays L.), einer in Europa und Nordamerika bedeutenden Pilzkrankheit, die zu verringertem Kornertrag und zu kontaminiertem Erntegut führt, belastet mit Mykotoxinen wie Deoxynivalenol (DON) und Zearalenone (ZON).

In dieser Arbeit wurden drei verschiedene Materialsätze von Maisinzuchtlinien in mehreren Umwelten evaluiert und für unterschiedliche Ziele analysiert. In der ersten Studie wurden fünf Doppelhaploiden(DH)-Familien (Größe 43 bis 204), die durch gemeinsame Eltern untereinander verwandt waren, in einem unvollständigen Halb-Diallel-Design mit vier von der Universität Hohenheim entwickelten Flint-Elternlinien generiert. Signifikante genotypische Varianzen und allgemein hohe Heritabilitäten wurden für alle drei untersuchten Merkmale (GER, DON und Tage bis zur weiblichen Blüte (DS)) in allen Familien beobachtet, was gute Erfolgsaussichten für Resistenzzüchtung und phänotypische Selektion gegen GER in verschiedenen Umwelten im europäischen Zuchtmaterial verspricht. Die genetischen Korrelationen zwischen DON und GER waren extrem eng und für DS mit DON oder GER moderat negativ. Dies deutet darauf hin, dass eine indirekte Selektion gegen GER effizient wäre, um die DON-Konzentration zu reduzieren, jedoch die Reife bei der GER-Resistenzzüchtung in Betracht gezogen werden sollte. Unter Verwendung einer hochdichten Konsensus-Karte mit 2.472 Marker Loci verglichen wir die klassische bi-parentale Kartierung von QTL (quantitative trait locus/loci) mit der multi-parentalen QTL-Kartierung, wozu alle Familien and vier verschiedene biometrische Modelle herangezogen wurden. Mit multi-parentale QTL-Kartierung-Modellen wurden alle und sogar noch weitere QTL identifiziert als mit dem bi-parentale QTL-Kartierung-Modell, das jeweils innerhalb jeder Familie verwendet wurde. Interessanterweise waren die

QTL für DON und GER meist familienspezifisch, während für DS verschiedene Familien mehrere gemeinsame QTL aufwiesen. Viele QTL zeigten bedeutende additive Gen-Effekte und die günstigsten Allele stammten vom hochresistenten Elter. Interaktionen zwischen den detektierten QTL und dem genetischen Hintergrund (Familie) waren selten und hatten vergleichsweise geringe Effekte. Multi-parentale QTL-Mapping-Modelle ergaben meist keine höhere Vorhersagegenauigkeit als bi-parentale QTL-Kartierung-Modelle für alle Merkmale.

In der zweiten Studie wurden zwei Diversitäts-Panels bestehend aus 130 europäischen Elite Dent- bzw. 114 Flint-Linien der Universität Hohenheim evaluiert und innerhalb jedes Pools einer genomweiten Assoziationsstudie unterzogen. Ähnlich wie in der ersten Studie wurden für GER, DON und DS hochsignifikante genotypische Varianzen und Genotyp × Umwelt-Interaktionsvarianzen beobachtet. Die Heritabilitäten waren mittelhoch für GER und DON und hoch für DS in beiden Pools. Die geschätzten genomischen Korrelationen zwischen den Pools waren für DON und DS nahe Null und für GER etwas höher. Die nachgewiesenen QTL für DON waren alle spezifisch für den jeweiligen Pool und verschieden von zuvor detektierten QTL. Darüber hinaus wurde in beiden Pools kein QTL für GER und DS gefunden. Die genomische Vorhersage (GP) über Pools hinweg ergab eine geringe oder gar negative Vorhersagegenauigkeit für alle Merkmale. Selbst wenn die Größe des Training-Satzes durch Kombinieren von Linien aus beiden heterotischen Pools erhöht wurde, hatten die GP-Ansätze mit kombinierten Pools keine höhere Vorhersagegenauigkeit als der GP-Ansatz innerhalb eines Pools. Anders als erwartet war beim Vergleich der statistischen Verfahren die BayesB Methode der genomischen besten linearen unverzerrten Vorhersage (GBLUP) hinsichtlich der Vorhersagegenauigkeit nicht überlegen.

In der dritten Studie analysierten wir zwei Rückkreuzungsfamilien (BC), die von einem resistenten und einem anfälligen rekurrenten Elter stammten. Beide BC-Populationen unterschieden sich erheblich in ihren Mittelwerten für alle Merkmale, was darauf hindeutet, dass die beiden rekurrenten Eltern unterschiedliche QTL-Allele für GER-Resistenz besitzen. Zwischen DON- und ZON-Konzentrationen Relativ wurde hohe Korrelationen beobachtet. Neben Immunoassays wurde auch visuelle Bonituren für den Befallsgrad von GER sowie NIRS(Nah-Infrarot-Spektroskopie)-Messungen am Erntegut innerhalb jeder

BC-Population durchgeführt. Danach stellt eine Selektion basierend auf visuellen Bonituren des Befalls von GER oder NIRS-Messungen am Erntegut einen effizienten und kostengünstigen Ansatz zur Reduzierung des Mykotoxin-Gehalts dar.

Zusammenfassend lässt sich aus den Ergebnissen ableiten, dass die GER-Resistenz im europäischen Mais durch Züchtung mit Hilfe resistenter Donoren effektiv verbessert werden kann. Die visuelle Bonitur des GER-Befallsgrades und NIRS-Messungen am vermahlenen Erntegut erwiesen sich als zuverlässige Indikatoren für die Rangierung von Genotypen hinsichtlich der DON- und ZON-Konzentration in Maiskörnern. Die detektierten QTL für GER und DON waren meist spezifisch für einige wenige Familien, während für DS häufiger gemeinsame QTL in mehreren Familien vorlagen. Der Ansatz der multi-parentalen QTL-Kartierung ergänzt die klassische bi-parentale QTL-Kartierung insofern, als letztere meist eine höhere Aussagekraft hat, um seltene QTL-Allele mit großem Effekt für Merkmale wie GER und DON zu detektieren, während erstere bei der Detektion von häufig vorkommenden QTL-Allelen mit geringem Effekt für Merkmale wie DS überlegen ist. Unabhängig vom Merkmal und dem verwendeten statistischen Verfahren konnte die Vorhersagegenauigkeit für MAS und GS durch eine engere Verwandtschaft zwischen dem zur Selektion vorgesehenen Zuchtmaterial und dem Trainingssatz und einer Vergrößerung desselben verbessert werden.

8. References

Bardol N, Ventelon M, Mangin B, et al (2013) Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (Zea mays L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. Theor Appl Genet 126:2717–2736. doi: 10.1007/s00122-013-2167-9

Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: Paterson AH (ed) Molecular Dissection of Complex Traits. CRC press, New York, pp 145–162

Bernardo R (2008) Molecular markers and selection for complex traits in plants : Learning from the last 20 years. Crop Sci 1649–1664. doi: 10.2135/cropsci2008.03.0131

Bink MCAM, Totir LR, ter Braak CJF, et al (2012) QTL linkage analysis of connected populations using ancestral marker and pedigree information. Theor Appl Genet 124:1097–1113. doi: 10.1007/s00122-011-1772-8

Blanc G, Charcosset A, Mangin B, et al (2006) Connected populations for detecting quantitative trait loci and testing for epistasis: An application in maize. Theor Appl Genet 113:206–224. doi: 10.1007/s00122-006-0287-1

Bolduan C, Miedaner T, Schipprack W, et al (2009) Genetic variation for resistance to ear rots and mycotoxins contamination in early European maize inbred lines. Crop Sci 49:2019–2028. doi: 10.2135/cropsci2008.12.0701

Chaikam V, Molenaar W, Melchinger AE, Boddupalli PM (2019) Doubled haploid technology for line development in maize: Technical advances and prospects. Theor. Appl. Genet 132: 3227–3243

de los Campos G, Vazquez AI, Fernando R, Klimentidis YC, Sorensen D (2013) Prediction of complex human traits using the genomic best linear unbiased predictor. PLoS Genet 9: e1003608. doi:10.1371/journal.pgen.1003608

De Roos APW, Hayes BJ, Goddard ME (2009) Reliability of genomic predictions across multiple populations. Genetics 183:1545–1553. doi: 10.1534/genetics.109.104935

De Roos APW, Hayes BJ, Spelman RJ, Goddard ME (2008) Linkage disequilibrium and persistence of phase in Holstein–Friesian, Jersey and Angus cattle. Genetics 179:1503–1512. doi: 10.1534/genetics.107.084301

Edwards, MD, Stuber CW, Wendel, JF (1987) Molecular-marker-facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116: 113 -125.

European Commission (2006) Commission recommendation (EC) No 576/2006 on the presence of deoxynivalenol, zearalenone, ochratoxin A, T-2 and HT-2 and fumonisins in products intended for animal feeding. https://eur-lex.

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. London: Longmans Green, Harlow, Essex, UK

Fischer S, Möhring J, Schön CC, Piepho HP, Klein D, Schipprack W, Utz HF, Melchinger AE, Reif JC (2008) Trends in genetic variance components during 30 years of hybrid maize breeding at the University of Hohenheim. Plant Breed 451:446–451. doi: 10.1111/j.1439-0523.2007.01475.x

Foiada F, Westermeier P, Kessel B, et al (2015) Improving resistance to the European corn borer: a comprehensive study in elite maize using QTL mapping and genome-wide prediction. Theor Appl Genet 128:875–891. doi: 10.1007/s00122-015-2477-1

Giraud H, Lehermeier C, Bauer E, et al (2014) Linkage disequilibrium with linkage analysis of multiline crosses reveals different multiallelic QTL for hybrid performance in the flint and dent heterotic groups of maize. Genetics 198:1717–1734. doi: 10.1534/genetics.114.169367

Goddard M (2009) Genomic selection : prediction of accuracy and maximisation of long term response. Genetica 136:245–257. doi: 10.1007/s10709-008-9308-0

Goddard ME, Hayes B, McPartlan H, Chamberlain AJ (2006) Can the same genetic markers be used in multiple breeds? 8th World Congr Genet Appl to Livest Prod August 13-18, 2006, Belo Horizonte, MG, Bras 4–7.

Guo Z, Tucker DM, Basten CJ, et al (2014) The impact of population structure on genomic prediction in stratified populations. Theor Appl Genet 127:749–762. doi: 10.1007/s00122-013-2255-x

Habier D, Fernando RL, Dekkers JCM (2007) The impact of genetic relationship information on genome-assisted breeding values. Gentetics 177:2389-2397. doi: 10.1534/genetics.107.081190

Habier D, Fernando RL, Garrick DJ (2013) Genomic BLUP decoded : A look into the black box of genomic prediction. Genetics 194:597–607. doi: 10.1534/genetics.113.152207

Han S, Miedaner T, Utz HF et al (2018) Genomic prediction and GWAS of Gibberella ear rot resistance traits in dent and flint lines of a public maize breeding program. Euphytica 214:1–20. https://doi.org/10.1007/s10681-017-2090-2

Han S, Utz HF, Liu W, Schrag TA, Stange M, Würschum T, Miedaner T, Bauer E, Schön CC, Melchinger AE (2016) Choice of models for QTL mapping with multiple families and design of the training set for prediction of *Fusarium* resistance traits in maize. Theor Appl Genet 129:431–444. doi: 10.1007/s00122-015-2637-3

Harrris BL, Johnson DL, Spelman RJ (2008) Genomic selection in New Zealand and the implications for national genetic evaluation. Proc Interbull Meet Niagara Falls, Canada 325–330.

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009) Invited review : Genomic selection in dairy cattle : Progress and challenges. J Dairy Sci 92:433–443. doi: 10.3168/jds.2008-1646

Heffner EL, Lorenz AJ, Jannink JL, Sorrells ME (2010) Plant breeding with genomic selection: gain per unit time and cost. Crop Sci 50: 1681–1690.

Heffner EL, Sorrells ME, Jannink J-L. (2009) Genomic selection for crop improvement. Crop Sci. 2009;49:1–12. doi: 10.2135/cropsci2008.08.0512

Holland JB (2007) Genetic architecture of complex traits in plants. Curr Opin Plant Biol 10:156–161. doi: 10.1016/j.pbi.2007.01.003

Hu H, Schrag TA, Peis R et al (2016) The genetic basis of haploid induction in maize identified with a novel genome-wide association method. Genetics 202:1267–1276

Huang, BE, Verbyla KL, Verbyla AP, et al. CR (2015) MAGIC populations in crops: current status and future prospects. Theor Appl Genet 128: 999-1017

Jannink JL, Jansen R (2001) Mapping epistatic quantitative trait loci with one-dimensional genome searches. Genetics 157:445–454.

Jannink JL, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. Brief Funct Genomics 9:166–77. doi: 10.1093/bfgp/elq001

Jansen RC, Jannink JL, Beavis WD (2003) Mapping Quantitative Trait Loci in Plant Breeding Populations. Crop Sci 43:829. doi: 10.2135/cropsci2003.0829

Kebede AZ, Woldemariam T, Reid LM, Harris LJ (2016) Quantitative trait loci mapping for Gibberella ear rot resistance and associated agronomic traits using genotyping-by-sequencing in maize. Theor Appl Genet 129:17-29. doi: 10.1007/s00122-015-2600-3

Kemper KE, Hayes BJ, Daetwyler HD, Goddard ME (2015) How old are quantitative trait loci and how widely do they segregate? J Anim Breed Genet 132:121–134. doi: 10.1111/jbg.12152

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps [published erratum appears in Genetics 1994 Feb;136(2):705]. Genetics 121:185–199.

Lehermeier C, Krämer N, Bauer E, et al (2014) Usefulness of multi-parental populations of maize (Zea mays L.) for genome-based prediction. Genetics 198:3–16. doi: 10.1534/genetics.114.161943

Lehermeier C, Schön CC, de los Campos G (2015) Assessment of genetic heterogeneity in structured plant populations using multivariate whole-genome regression models. Genetics 201:323–337. doi: 10.1534/genetics.115.177394

Lorenz AJ, Smith KP (2015) Adding genetically distant individuals to training populations reduces genomic prediction accuracy in barley. Crop Sci 55:2657–2667 doi: 10.2135/cropsci2014.12.0827

Lorenz AJ, Smith KP, Jannink J (2012) Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. Crop Sci 52:1609–1621. doi: 10.2135/cropsci2011.09.0503

Lund MS, Su G, Janss L, Guldbrandtsen B (2014) Invited review : Genomic evaluation of cattle in a multi-breed context. Livest Sci 166:101–110. doi: 10.1016/j.livsci.2014.05.008

Li H, Bradbury P, Ersoz E, et al (2011) Joint QTL linkage mapping for multiple-cross mating design sharing one common parent. PLoS One. doi: 10.1371/journal.pone.0017573

Liu Y, Zeng ZB (2000) A general mixture model approach for mapping quantitative trait loci from diverse cross designs involving multiple inbred lines. Genet Res 75:345–355. doi: 10.1017/S0016672300004493

Lu Y, Xu J, Yuan Z, et al (2012) Comparative LD mapping using single SNPs and haplotypes identifies QTL for plant height and biomass as secondary traits of drought tolerance in maize. Mol Breed 30:407–418. doi: 10.1007/s11032-011-9631-5

Martin M, Miedaner T, Dhillon BS, et al (2011) Colocalization of QTL for gibberella ear rot resistance and low mycotoxin contamination in early European maize. Crop Sci 51:1935–1945. doi: 10.2135/cropsci2010.11.0664

Martin M, Miedaner T, Schwegler DD, et al (2012) Comparative quantitative trait loci mapping for Gibberella ear rot resistance and reduced deoxynivalenol contamination across connected maize populations. Crop Sci 52:32–43. doi: 10.2135/cropsci2011.04.0214

Melchinger AE (1999) Genetic diversity and heterosis. J.G. Coors and S. Pandey (eds.) Genetics and Exploitation of Heterosis in Crops. ASA - CSSA, Madison, WI, USA: 99-118

Melchinger AE, Brauner PC, Böhm J, Schipprack W (2016) In vivo haploid induction in maize: comparison of different testing regimes for measuring haploid induction rates. Crop Sci 56:1127–1135

Melchinger AE, Schipprack W, Mi X, Mirdita V (2015) Oil content is superior to oil mass for identification of haploid seeds in maize produced with high-oil inducers. Crop Sci 55(1):188–195

Melchinger AE, Schipprack W, Würschum T, Chen S, Technow F (2013) Rapid and accurate identification of in vivo induced haploid seeds based on oil content provides a new tool for maize genetics and breeding. Sci Rep 3: 2129

Melchinger AE, Utz HF, Schön CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403. doi: 10.1016/1369-5266(88)80015-3

Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–29.

Michel S, Wagner C, Nosenko T, Steiner B, Samad-Zamini M, Buerstmayr M, Mayer K, Buerstmayr H (2021): Merging genomics and transcriptomics for predicting Fusarium head blight resistance in wheat GENES-BASEL. 2021; 12(1), 114

Miedaner T, Han S, Kessel B, et al (2015) Prediction of deoxynivalenol and zearalenone concentrations in *Fusarium graminearum* inoculated backcross populations of maize by symptom rating and near-infrared spectroscopy. Plant Breed 009:n/a–n/a. doi: 10.1111/pbr.12297

Ogut F, Bian Y, Bradbury PJ, Holland JB (2015) Joint-multiple family linkage analysis predicts within-family variation better than single-family analysis of the maize nested association mapping population. Heredity (Edinb) 114:552–563. doi: 10.1038/hdy.2014.123

Olson KM, VanRaden PM, Tooker ME (2012) Multibreed genomic evaluations using purebred Holsteins, Jerseys, and Brown Swiss. J Dairy Sci 95:5378–5383. doi: 10.3168/jds.2011-5006

Pfordt A, Ramos Romero L, Schiwek S et al (2020) Impact of environmental conditions and agronomic practices on the prevalence of Fusarium species associated with ear- and stalk rot in maize. Pathogens 9:236. https://doi.org/10. 3390/pathogens9030236

Prigge V, Melchinger AE (2012) Production of haploids and doubled haploids in maize. Methods Mol Biol 877:161–172

Rebai A, Goffinet B (1993) Power of tests for QTL detection using replicated progenies derived from a diallel cross. Theor Appl Genet 86:1014–1022. doi: 10.1007/BF00211055

Rebai A, Goffinet B (2000) More about quantitative trait locus mapping with diallel designs. Genet Res 75:243–247

Reif JC, Hamrit S, Heckenberger M, Schipprack W, Maurer HP, Bohn M, Melchinger AE (2005) Trends in genetic diversity among European maize cultivars and their parental components during the past 50 years. Theor Appl Genet 111:838–45. doi: 10.1007/s00122-005-0004-5

Riedelsheimer C, Endelman JB, Stange M, Sorrells ME, Jannink JL, Melchinger AE (2013) Genomic predictability of interconnected biparental maize populations. Genetics 194:493–503. doi: 10.1534/genetics.113.150227

Riedelsheimer C, Melchinger AE (2013) Optimizing the allocation of resources for genomic selection in one breeding cycle. Theor Appl Genet 126: 2835–2848. doi:10.1007/s00122-013-2175-9

Rutkoski J, Benson J, Jia Y, et al (2012) Evaluation of genomic prediction methods for Fusarium head blight resistance in wheat. Plant Genome J 5:51. doi: 10.3835/plantgenome2012.02.0001

Schopp P, Müller D, Technow F, Melchinger AE (2017) Accuracy of genomic prediction in synthetic populations depending on the number of parents, relatedness and ancestral linkage disequilibrium. Genetics 205:441-454. doi: 10.1534/genetics.116.193243

Schulz-Streeck T, Ogutu JO, Karaman Z, Knaak C, Piepho HP (2012) Genomic selection using multiple populations. Crop Sci 52:2453–2461. doi: 10.2135/cropsci2012.03.0160

Schön CC, Utz HF, Groh S, et al (2004) Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. Genetics 167:485–498. doi: 10.1534/genetics.167.1.485

Sillanpää MJ (2011) Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. Heredity (Edinb) 106:511–519. doi: 10.1038/hdy.2010.91

Steinhoff J, Liu W, Maurer HP, et al (2011) Multiple-Line Cross Quantitative Trait Locus Mapping in European Elite Maize. Crop Sci 51:2505. doi: 10.2135/cropsci2011.03.0181

Stich B, Möhring J, Piepho HP, Heckenberger M, Buckler ES, Melchinger AE (2008) Comparison of mixed-model approaches for association mapping. Genetics 178:1745–54. doi: 10.1534/genetics.107.079707

Technow F, Bürger A, Melchinger AE (2013) Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. G3:Genes Genomes Genet 3:197–203. doi: 10.1534/g3.112.004630

Technow F, Schrag TA, Schipprack W, Bauer E, Simianer H, Melchinger AE (2014a) Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. Genetics 197: 1343–1355. doi: 10.1534/genetics.114.165860

Technow F, Schrag TA, Schipprack W, Melchinger AE (2014b) Identification of key ancestors of modern germplasm in a breeding program of maize. Theor Appl Genet 127:2545-2553.

Utz HF, Melchinger AE (1994) Comparison of different approaches to interval mapping of quantitative trait loci. In: Ooijen JW van, Jansen J (ed), Biometrics plant Breed Appl Mol markers Wageningen:the Netherlands, 6–8 July 1994. 1994, 195–204 ST.

Utz HF, Melchinger AE, Schön CC (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 154:1839–1849.

Weber KL, Thallman RM, Keele JW, et al (2012) Accuracy of genomic breeding values in multibreed beef cattle populations derived from deregressed breeding values and phenotypes. J Anim Sci 90:4177–4190. doi: 10.2527/jas2011-4586

Westhues M, Schrag TA, Heuer C, Thaller G, Utz HF, Schipprack W, Thiemann A, Seifert F, Ehret A, Schlereth A, Stitt M, Nikoloski Z, Willmitzer L, Schön CC, Scholten S, Melchinger AE (2017) Omics-based hybrid prediction in maize Theor Appl Genet 130:1927–1939 doi: 10.1007/s00122-017-2934-0

Wientjes YCJ, Bijma P, Veerkamp RF, Calus MPL (2016) An equation to predict the accuracy of genomic values by combining data from multiple traits, populations, or environments. Genetics 202:799–823. doi: 10.1534/genetics.115.183269

Wientjes YCJ, Calus MPL, Goddard ME, Hayes BJ (2015) Impact of QTL properties on the accuracy of multi-breed genomic prediction. Genet Sel Evol 47:42. doi: 10.1186/s12711-015-0124-6

Würschum T, Liu W, Gowda M, et al (2012) Comparison of biometrical models for joint linkage association mapping. Heredity (Edinb) 108:332–340. doi: 10.1038/hdy.2011.78

Xu S (1998) Mapping quantitative trait loci using multiple families of line crosses. Genetics 148: 517–524

Xu S (2003) Theoretical basis of the Beavis effect. Genetics 165: 2259 - 2268

Xu Y, Li P, Zou C, Lu Y, et al. (2017) Enhancing genetic gain in the era of molecular breeding, Journal of Experimental Botany, Volume 68, Issue 11, Pages 2641–2666, https://doi.org/10.1093/jxb/erx135

Yan J, Warburton, M.; Crouch, J.H (2011) Association mapping for enhancing maize (Zea mays L.) genetic improvement. Crop Sci 51 (2): 433-449. doi: 10.2135/cropsci2010.04.0233

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539–551. doi: 10.1534/genetics.107.074245

Yu J, Pressoir G, Briggs WH, et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–8. doi: 10.1038/ng1702

Zhou L, Ding X, Zhang Q, Wang Y, Lund MS, Sun G (2013) Consistency of linkage disequilibrium between Chinese and Nordic Holsteins and genomic prediction for Chinese Holsteins using a joint reference population. Genet Sel Evol 45:7. doi: 10.1186/1297-9686-45-7

Zila CT, Ogut F, Romay MC, Gardner CA, Buckler ES, Holland JB (2014). Genome-wide association study of Fusarium ear rot disease in the U.S.A. maize inbred line collection. BMC Plant Biol. 14:372. doi: 10.1186/s12870-014-0372-6

Zila CT, Samayoa LF, Santiago R, Butrón A, Holland JB (2013) A genome-wide association study reveals genes associated with Fusarium ear rot resistance in a maize core diversity panel. G3 Genes Genomes Genet 3:2095–2104. doi: 10.1534/g3.113.007328

9. Acknowlegements

First and foremost, I would like to thank my academic supervisor Prof. Dr. Albrecht E. Melchinger for giving me the opportunity to pursue my Ph.D. in his distinguished working group. I am sincerely grateful for his suggestions, guidance and constant support throughout my thesis. He is a lifelong role model to me for his passion and persistence for scientific research, enthusiasm in training young scientists, hard-working spirit and exceptional contributions to worldwide scientific communities.

Many thanks to Prof. Dr. H. Friedrich Utz, with whom I enjoyed very much discussions about QTL mapping, phenotypic analysis, statistical models and his softwares. He was always friendly, available and supportive. It was fun to work with him. To me, he is not only a supervisor but also a friend.

I would like to thank Prof. Dr. Thomas Miedaner for being my second supervisor for my Ph.D. thesis. He shared with me his great expertise on *Fusarium* resistance breeding in maize, and brought me back to the right track to complete my Ph.D. thesis.

I am grateful to Dr. Tobias Schrag for his contributions to my thesis and insightful discussions on various topics. He was always patient, supportive and friendly to me. He is a great example for being well-organized, details-oriented and multi-tasked.

Many thanks to Dr. Marco Bink and Prof. Dr. Fred van Eeuwijk for hosting me at Wageningen University and Research Center, the Netherlands. It was a very nice experience to interact with scientists such as Dr. Martin Boer and Ph.D. students there.

I thank Dr. Wenxin Liu, Dr. Michael Stange, Dr. Wolfgang Schipprack and Prof. Dr. Tobias Würschum as well as other co-authors for their contributions to our publications and interesting discussions. Many thanks to Prof. Dr. Hermann Bürstmayr for reviewing my thesis.

I am thankful to group 350a members Helga Kösling, Beate Devezi-Savula, Birgit Schilling, Franz Mauch, Jochen Jesse and Margit Lieb. Without their help and support, my Ph.D. research would have been very difficult to conduct.

Many thanks to Frank Technow, Christian Riedelsheimer, Anna Pucher, Willmar Leiser, Dr. Xuefei Mi, Haixiao Hu, Matthias Westhues, Willem Molenaar, Jose Marulanda, Philipp Boeven, Pedro Brauner-Correa, Felix Sattler, Valheria Castiblanco, Pascal Schopp, Juliane Renner, Dominik Müller, Patrick Thorwarth, Angela Maria Bernal Vasquez, Heike Molenaar, Dominik Losert and Linda Homann, who had been great mentors and companies throughout my Ph.D. and private life.

I sincerely thank my family in China, my grandma, my parents and my sister, for their constant encouragement and support. I am greatly indebted to my wife Xueli Zhao and our son Antong Han, without whom my Ph.D. journey would not have been possible.

10. Curriculum vitae

Name	Sen Han
Birth	October 8th 1985 in Shandong, China

Work Experience

2018 - present	Corn Breeder and Station Manager, RAGT2n, Deggendorf (Bavaria), Germany
2017 - 2018	Statistical Scientist, Monsanto Company, Rotterdam, the Netherlands
2015 - 2016	Project Manager, Bioeconomy Research Center, University of Hohenheim, Stuttgart, Germany
2012 - 2012	Research Intern, DuPont Pioneer, Eschbach, Germany

University Education

2013 - 2021	Ph.D. candidate, Institute of Plant Breeding, Seed Science and Population Genetics, Prof. Dr. Albrecht E. Melchinger research group, University of Hohenheim, Stuttgart, Germany
2010 - 2013	M.Sc. in Crop Sciences (major in Plant Breeding and Seed Sciences), University of Hohenheim, Stuttgart, Germany
2005 - 2009	B.E. in Food Science and Engineering, Sichuan University, Chengdu, China

School Education

2002 - 2005	Shouguang High School, Shouguang, China
1998 - 2002	Luocheng No. 2 Middle School, Shouguang, China
1993 - 1998	Hanjia and Zhongxin Primary School, Shouguang, China

_____

Sen Han

Deggendorf, 16.10.2021

11. Declaration

Annex 3

Declaration in lieu of an oath on independent work

according to Sec. 18(3) sentence 5 of the University of Hohenheim's Doctoral Regulations for the Faculties of Agricultural Sciences, Natural Sciences, and Business, Economics and Social Sciences

1. The dissertation submitted on the topic

   Gibberella ear rot resistance in European maize: genetic analysis by complementary mapping approaches and improvement with genomic selection

   is work done independently by me.
2. I only used the sources and aids listed and did not make use of any impermissible assistance from third parties. In particular, I marked all content taken word-for-word or paraphrased from other works.
3. I did not use the assistance of a commercial doctoral placement or advising agency.
4. I am aware of the importance of the declaration in lieu of oath and the criminal consequences of false or incomplete declarations in lieu of oath. I confirm that the declaration above is correct. I declare in lieu of oath that I have declared only the truth to the best of my knowledge and have not omitted anything.

_____ _____

Place, Date Signature