Aus dem Institut für
Pflanzenzüchtung, Saatgutforschung und Populationsgenetik
der Universität Hohenheim
Fachgebiet Angewandte Genetik und Pflanzenzüchtung
Prof. Dr. A. E. Melchinger

Dissection of the genetic architecture of stalk mechanical strength and *in vivo* haploid induction in maize

Dissertation
zur Erlangung des Grades eines
Doktors der Agrarwissenschaften
vorgelegt
der Fakultät Agrarwissenschaften
der Universität Hohenheim

von
M. Sc. der Agrarwissenschaften
Haixiao Hu
aus Shandong (China)

Stuttgart–Hohenheim
2016

Die vorliegende Doktorarbeit wurde am 30. September 2015 von der Fakultät Agrarwissenschaften der Universität Hohenheim als “Dissertation zur Erlangung des Grades eines Doktors der Agrarwissenschaften (Dr. sc. agr.)” angenommen.

Tag der mündlichen Prüfung: 29. Januar 2016

1. Prodekan:	Prof. Dr. T. Streck
Berichterstatter, 1. Prüfer:	Prof. Dr. A. E. Melchinger
Mitberichterstatter, 2. Prüfer:	Prof. Dr. C. Schön
3. Prüfer	Prof. Dr. C. Zebitz

Contents

[1] Hu H, Meng Y, Wang H, Liu H, Chen S (2012) Identifying quantitative trait loci and determining closely related stalk traits for rind penetrometer resistance in a high-oil maize population. Theor Appl Genet 124: 1439–1447

[2] Hu H, Liu W, Fu Z, Homann L, Technow F, Wang H, Song C, Li S, Melchinger AE, Chen S (2013) QTL mapping of stalk bending strength in a recombinant inbred line maize population. Theor Appl Genet 126: 2257–2266

[3] Hu H, Schrag T, Peis R, Unterseer S, Schipprack W, Chen S, Lai J, Yan J, Prasanna BM, Nair SK, Chaikam V, Rotarenco V, Shatskaya OA, Zavalishina A, Scholten S, Schön CC, Melchinger AE (2016) The genetic basis of haploid induction in maize identified with a novel genome-wide association method. Genetics (Accepted Paper)

Abbreviations

CHE	conditional haplotype extension
DH	double haploid
F_{max}	maximum load exerted to breaking
GWAS	genome-wide association study
HI	haploid induction
HIR	haploid induction rate
kb	kilobase pair
LD	linkage disequilibrium
MAS	marker-assisted selection
Mb	megabase pair
M_{max}	breaking moment
QTL	quantitative trait loci
RIL	recombinant inbred line
RPR	rind penetrometer resistance
SBS	stalk bending strength
σ_{max}	critical stress

Chapter 1

General Introduction

Maize plays an important role in securing world's food supply, animal husbandry and the deep processing industry. According to statistics of the international grain council (IGC), global production of maize, at 982 million tons in 2014, has far exceeded that of wheat and rice during the past 5 years. However, stalk lodging causes yield losses in maize cultivation ranging between 5 to 20% annually worldwide (Flint-Garcia et al. 2003). Reduction of stalk lodging is required for both yield improvement and application of harvesting machines. Constantly increasing market demands urge maize geneticists and breeders not only to enhance the field performance of new hybrids but also to improve breeding process. During the last decade, advances in the double haploid (DH) technology based on *in vivo* haploid induction (HI) shifted breeding paradigms and greatly accelerate the breeding process (Melchinger et al. 2013). However, the number of currently available maize haploid inducers is far smaller in comparison to the actual demands to be utilized in various environments due to lack of a simple and efficient way for developing maize haploid inducers. In this frame, further investigation of stalk lodging and HI is compulsory to develop new germplasm resistant to lodging and obtain further progress in the utilization of DH technology.

1.1 Stalk mechanical strength and *in vivo* HI in maize

1.1.1 Stalk mechanical strength and its measurement

The most direct way to evaluate stalk lodging resistance is counting the number of lodged plants at harvest. However, it cannot always be reliably determined in field trials, because it strongly depends on the environmental conditions (Thompson 1963; Hu et al. 2012).

Many studies (Zuber et al. 1961; Colbert et al. 1984; Jia et al. 1992; Gou et al. 2007) found that stalk mechanical strength is positively correlated with stalk lodging resistance in the field. Enhancing overall mechanical strength in maize will make stalks stronger and ultimately reduce yield and grain quality losses (Ching et al. 2010). Since stalk lodging is extremely affected by the environmental conditions and, therefore, has often a low heritability, stalk mechanical strength can be considered as a reliable indicator for evaluating stalk lodging resistance.

Several methods have been developed and applied to evaluate stalk mechanical strength. Zuber and Grogan (1961) measured stalk crushing strength. Jia and Bai (1992) developed a field bending approach. Sibale et al. (1992) used rind penetrometer resistance (RPR), and Kokubo et al. (1989), Ma (2009) and Ching et al. (2010) measured stalk bending strength (SBS) with a three point bending approach (Gere and Timoshenko 1984). Among all these approaches, the RPR method is simple, rapid and most importantly does not damage the stalk during data collection (Hu et al. 2000). The SBS approach is more closely associated with stalk lodging in the field because under natural conditions, stalk lodging occurs when the stalk bending level exceeds the critical bending point (Yuan et al. 2002). In plant physiology and breeding, only the maximum load exerted to breaking (F_{max}) was used to characterize SBS (Kokubo et al. 1989; Ma 2009; Ching et al. 2010). However, in the field of mechanics of materials, bending strength is actually reflected by the breaking moment (M_{max}) and the critical stress (σ_{max}) besides F_{max} (Gere and Timoshenko 1984). Thus, to have a complete understanding of the genetic architecture of SBS in maize, we decided to use RPR and SBS (measured by F_{max}, M_{max} and σ_{max}) to characterize stalk strength in our study.

1.1.2 In vivo HI in maize

Generating DH lines from maternal haploids in maize consists of four major steps: i) *In vivo* haploid induction; ii) haploid seed identification using morphological markers; iii) chromosome doubling of putative haploids; and iv) generating D_1 seed from D_0 seedlings (Prasanna et al. 2012). *In vivo* haploid induction is achieved by crossing a specially developed

maize genetic stock called “inducer” (as male) with a source population (as female) from which homozygous DH lines are developed. The haploid seeds can be identified with various approaches: (i) morphological marker systems, such as the *R1-nj* marker system. Briefly, haploid kernels can be distinguished from diploid kernels (purple scutellum and purple aleurone) by their colorless embryo and colored endosperm (Li et al. 2009); (ii) the liguleless test. A liguleless tester is pollinated with the inducer in the field. Then, in the testing phase, diploid progenies show a ligule and auricle, while haploids are characterized by a missing ligule and auricle (Prigge et al. 2012); (iii) agronomic traits evaluation in field trials. Compared to diploids, haploid plants display shorter stature, slender weak stems, erecter and narrower leaves, and reduced growth rate (Xu et tal. 2013).

The haploid induction ability varies from inducer to inducer, which is assessed by the so-called haploid induction rate (HIR). Xu et al. (2013) proposed two methods to compute HIR. One relies on self-pollinated ears and termed S-HIR. S-HIR is calculated as the percentage of haploid plants obtained by selfing a pollinated inducer. The other, termed T-HIR, is to calculate the percentage of haploid plants obtained by crossing an inducer with source germplasm. In practice, S-HIR or T-HIR is usually estimated from putative haploid kernels identified by a morphological marker system, such as the *R1-nj* marker system. However, for specific purposes that require very accurate HIR estimates, such as QTL mapping or fine mapping of HIR, the estimated HIR must be confirmed by growing all putative haploid kernels in the field, and then calculating the percentage of haploid plants.

The study of haploids in maize started in 1929 by Randolph and Stadler (Chase 1969). However, only during the last decade, the DH technology was widely adopted and routinely used in scientific studies and practical breeding (Melchinger et al. 2013). Lacking of haploid inducers with acceptable HIR, i.e. ≥1%, can explain the extremely low development and spread of DH technology. During the long period of time, Chase and Coe played key roles in pushing forward the development of inducers with high HIR. Chase firstly recognized the prospect of

DH technology and demonstrated its practical values in maize breeding by utilizing DH lines in maize hybrid breeding (Chang and Coe 2009). Coe released the famous inducer Stock 6 in 1959, showing highest HIR (S-HIR = 3.23%, Coe 1959) at that time. Based on Stock 6, the second and the third generation of inducers with higher HIR have been developed worldwide, which essentially promoted the wide spread of the DH technology. Although dozens of maize inducers have been developed in public institutes worldwide, development of a new inducer line is still difficult due to massive work necessary for phenotypic evaluation. In addition, the number of presently available inducers is far smaller in comparison to the actual demand to be utilized in various environments. Further, a much simpler and more efficient approach for developing maize haploid inducers is needed to be built for public institutes or small-scale breeding companies to breed their own maize haploid inducers.

1.2 Methodology of dissecting the genetic architecture of quantitative traits

Many agronomically and economically important traits in crop plants, such as grain yield, quality, flowering time, resistance to diseases and stress tolerance are quantitative traits (Ren et al. 2005). In our study, both stalk mechanical strength (Flint-Garcia et al. 2003; Ching et a. 2010) and HIR (Prigge et al. 2012) are quantitative traits. In contrast to quality traits, quantitative traits (i) are controlled by multiple genes, (ii) exhibit continuous variation and (iii) are modified by environmental effects (Lander and Botstein, 1989). Therefore, it’s relatively difficult to investigate quantitative traits. For most of the period up to 1980, geneticists were limited to characterize quantitative traits by the means, variances and covariances of relatives at the phenotypic level (Kearsey and Farquhar, 1998). Moreover, research about (i) the number and genome locations of genes/quantitative trait loci (QTL) that affect a trait, (ii) the magnitude of their effects, and (iii) the relative contribution of additive, dominant, and epistatic gene/QTL effects was hindered. These three factors define the architecture of a quantitative trait (Holland 2007).

1.2.1 Mapping QTL of quantitative traits

The advent of molecular makers enables arrangement of crosses with genetic markers densely spaced throughout an entire genome (Lander and Botstein, 1989). Linkage analysis based on molecular markers, i.e. QTL mapping, has become a routine tool to dissect the genetic architecture of a quantitative traits since Paterson et al. (1988) used a complete restriction fragment length polymorphism (RFLP) linkage map to analyze QTLs controlling yield and quality related traits in tomato. Various statistical methods for QTL mapping have been developed, which range from simple interval mapping (Lander and Botstein, 1989), to more effective methods such as composite interval mapping (Zeng, 1994), mixed-model based composite interval mapping (Zhu, 2000) and inclusive composite interval mapping (Li et al. 2007). Correspondingly, software packages publicly available for QTL mapping were developed, such as MAPMAKER/QTL (Lincoln et al. 1993), QTL Cartographer (Basten et al. 1994), PLABQTL (Utz and Melchinger, 1996), QTL Network (Yang et al. 2010) and Icimapping (Wang et al. 2010).

For a successful exploitation of QTL mapping in plant breeding, reliable estimates of the phenotypic/genotypic variance explained by QTL are needed. Melchinger et al. (1998) demonstrated that QTL effects are overestimated if the same data set is used to locate QTL and estimate their effects. Utz et al. (2000) proposed a k-fold cross-validation (CV) approach to evaluate the bias and sampling error in estimation of QTL effects. For performing 5-fold CV, the dataset (DS) is evenly and randomly divided into 5 subsets, from which 4 subsets serve as estimation set (ES) for locating QTL and the fifth subset, termed test set (TS), is used for estimating the QTL effects. This procedure is repeated for a reasonable number of times (e.g. 1,000) with different partitions of the DS into ES and TS. The CV approach has been implemented in PLABQTL software, and has been widely applied in QTL mapping of various crops (e.g. Schön et al. 2004; Micic et al. 2005; Lisec et al. 2008; Würschum 2012; Hu et al. 2013).

1.2.2 Conventional fine-mapping approaches

The nature of a trait may sometimes suggest that much of the quantitative variation is controlled by a few genes with large effects (Bernardo 2008), such as the *qhir1* for HI in our study (Prigge et al. 2012). Other examples include the *fw2.2* locus for tomato fruit size (Frary et al. 2000), the *ZmWAK* locus for head smut resistance in maize (Zuo et al. 2014), the *DGAT1-2* locus for maize oil content (Zheng et al. 2008), and the *Fhb1* locus for resistance to Fusarium head blight in wheat (Anderson et al. 2008). In this situation, fine-mapping the major QTL would be of great help for crop improvement with marker-assisted selection (MAS) (Bernardo 2008; Anderson et al. 2008; Chai et al. 2012; Zhao et al. 2012).

For fine-mapping the major QTL, the quantitative trait is usually simplified to a quality trait according to whether a genotype has the target segment or not (Yang et al. 2012). Based on linkage mapping, two widely used approaches are: (i) developing a set of near-isogenic lines (NILs) and (ii) developing a set of introgression lines (ILs), and subsequently narrow down the target QTL gradually by comparison of genotypes and phenotypes of recombinant-derived progenies (Zheng et al. 2008; Tan et al. 2008). The major disadvantage of the two approaches is that many generations are needed to breed ideal NILs or ILs. Yang et al. (2012) modified the NILs-based approach by starting screening recombinants in BC_3 and testing recombinants-derived progenies in BC_4. However, this approach is still time consuming (at least four generations are needed to develop the BC_3 generation). Moreover, there still exists genetic background noise among the BC_3 recombinants-derived progenies in generation BC_4.

1.2.3 Case-control association mapping

Case-control association mapping can also be used for fine mapping the major QTL of a quantitative trait. In this approach, an individual carrying a target genomic segment is treated as case, otherwise as control. In contrast to linkage mapping based approaches like NILs or ILs, the major advantage of case-control association mapping is no need to produce a segregating

mapping population. Instead, a collection of individuals can be used to form a diversity population, and then identify common genetic variants associated with the target trait.

1.2.4 Detection of Selective sweeps

Genome scan for selective sweeps can also be used for fine-mapping major genes, because major genes constantly associate with strong selective signatures. Examples include the lactase locus in humans (Voight et al. 2006) and the malaria related locus in African populations (Sabeti et al. 2002). Three categories of approaches for detecting selective sweeps are commonly used in studies of humans and animals. One is based on site frequency spectrum (SFS) variation, represented by Tajima's D statistic (Tajima 1989) and Fay and Wu's H statistic (Fay and Wu 2000). These statistics measure departures from neutrality that are reflected by site-specific variations in natural populations. The second category attempts to detect long-range haplotypes resulting from positive selection. This approach was first proposed as extended haplotype homozygosity (EHH) by Sabeti et al. (2002). Voight et al. (2006) extended EHH for genome-wide scanning in a single population and termed it as integrated haplotype score (iHS). Tang et al. (2007) further extended the EHH approach in a pair of populations and termed this new approach as Rsb. The third category exploits the variance of population differentiation (F_{ST}). The preliminary version of F_{ST} based approach was described by Lewontin and Krakauer (1973), and is therefore named as LK test. As the original LK test did not consider complex demographic structures, Bonhomme et al. (2010) modified it by taking the population kinship matrix into account and named it FLK. Fariello et al. (2013) further extended FLK test for application to haplotypes, termed hapFLK, and demonstrated its higher power compared to FLK.

1.3 MAS and genomic selection

Currently, there are two approaches to utilize molecular markers for artificial selection in plant breeding. One is MAS (Lande and Thompson 1990), and the other is genomic selection

(Meuwissen et al. 2001). To perform MAS, a QTL/fine mapping must be performed beforehand to obtain markers that are either located inside the QTL (called functional markers) or linked to the QTL. MAS includes two cases: (i) MAS solely based on molecular markers, termed pure MAS; and (ii) MAS based on combining both phenotypic values and QTL information, termed combined MAS. The relative efficiency of MAS for both cases compared with classical phenotypic selection was described by Lande and Thompson (1990) with the following formulas:

Relative efficiency of pure MAS = $\sqrt{\frac{p}{h^2}}$,

where p refers to genetic variance explained by all QTL and h^2 refers to heritability of a given trait.

Relative efficiency of combined MAS = $\sqrt{\frac{p}{h^2} + \frac{(1-p)^2}{1-h^2 p}}$.

In contrast, genomic selection does not need a QTL identification step and enables utilizing genome-wide markers for predicting breeding values of candidates. The procedure and methods of genomic selection have been described by Meuwissen et al. (2001) and Technow (2013). Genomic selection has been successfully implemented in dairy cattle breeding (Hayes et al. 2009), and received considerable interest recently among plant geneticists and breeders (Albrecht et al. 2011; Zhao et al. 2012; Technow et al. 2013; Würschum et al. 2013; Albrecht et al. 2014).

1.4 Recent advances in investigation of the genetic architecture of stalk mechanical strength and *in vivo* HI in maize

1.4.1 QTL mapping of stalk mechanical strength

Although stalk mechanical strength is very important in maize breeding, understanding of its genetic basis is still very limited. To our knowledge, at the time our study was started, there were only two studies about the genetic architecture of this trait: one was a QTL mapping

on RPR by Flint-Garcia et al. (2003) and the other was an association mapping on SBS by Ching et al. (2010). Flint-Garcia et al. (2003) detected eight, ten, eight and nine single-effect QTL and four, two, zero, and five epistatic interactions for RPR in four $F_{2:3}$ maize populations, respectively. Ching et al. (2010) detected QTL of F_{max} on chromosomes 1, 5 and 9 with 189 non-Stiff Stalk lines. Recently, Peiffer et al. (2013) identified 18 QTL with joint linkage mapping and 141 significant markers with genome-wide association mapping for RPR in NAM and IBM populations. Li et al. (2014) detected seven QTL for RPR with two recombinant inbred line populations, among which a major QTL, *qRPR3-1*, explains 18.9% of the phenotypic variance. Moreover, these authors further narrowed down the support interval of *qRPR3-1* to a 3.1 Mb region and identified four genes involved in the biosynthesis of cell wall components as putatively candidate genes of RPR.

1.4.2 QTL mapping and fine mapping of HIR

QTL mapping of HIR was initiated by Deimling et al. (1997) using a segregating population derived from the cross of W23ig × Stock 6 with 84 RFLP markers at the University of Hohenheim. Two QTL were detected in bin 1.03-1.06 and bin 2.04-2.06, respectively. Barret et al. (2008) located a QTL for HIR between marker umc1917 and bnlg1811 in bin 1.04 using 101 SSR markers in the F_2 population of the cross DH99 × PK6. Prigge et al. (2012) identified two QTL for HIR based on four segregating populations from crosses of type non-inducer × inducer. The *qhir1* was a major QTL explaining large genetic variance (66%) and was also located in bin 1.04. Moreover, they also detected a large-effect QTL ($\hat{p}$ > 20%) in bin 9.01 influencing HIR in a segregating population of CAUHOI × UH400. To narrow down the support interval of *qhir1*, Dong et al. (2013) searched for recombinants in the *qhir1* region in a large (N=14,375) F_2 population derived from a cross between haploid inducer UH400 and non-inducer line 1680. After determining the recombination sites in the F_2 recombinants and analyzing corresponding F_3 families for segregation of HIR, the authors claimed to have narrowed down the *qhir1* locus to a 243 kb region. Considering the complex genetic

architecture of HI and genetic background noise possibly affecting the fine-mapping of *qhir1* by Dong (2013), it seems prudent to validate these results with an alternative approach before embarking on map-based gene isolation.

1.5 Objectives

The goals of this treatise were to dissect the genetic architecture of stalk mechanical traits, RPR and SBS, with a recombinant inbred line population and to investigate the major QTL for HI in a diversity panel of inducers and non-inducer lines in maize. In detail, our objectives were to:

(1) estimate genetic variances and heritability for RPR and SBS and their correlations with other stalk traits,

(2) identify QTL associated with RPR and SBS as well as their related stalk traits,

(3) evaluate the reliability of these QTL of SBS by cross-validation,

(4) compare the prospects of marker-assisted and genomic selection for SBS-related traits,

(5) analyze the genetic diversity between inducers and a worldwide collection of non-inducers,

(6) describe a novel approach for detecting selective sweeps in a population of closely related genotypes and compare its performance with other genome-wide association mapping or selective sweep approaches,

(7) detect genomic regions harboring the major gene of HI and compare the results with previous QTL and fine mapping studies.

Chapter 2

Identifying quantitative trait loci and determining closely related stalk traits for rind penetrometer resistance in a high-oil maize population

Haixiao Hu, Yujie Meng, Hongwu Wang, Hai Liu, Shaojiang Chen

Haixiao Hu, Yujie Meng and Shaojiang Chen, National Maize Improvement Center of China, China Agricultural University, 2# Yuanmingyuan West Road, 100193 Beijing, China

Hongwu Wang, Institute of Crop Science, Chinese Academy of Agricultural Sciences, 12# Zhongguancun South Street, 100081 Beijing, China

Hai Liu, Sorghum Institute, Shanxi Academy of Agricultural Sciences, 238# Yunhua West Road, Yuci District, 030600 Jinzhong, China

Theor Appl Genet 124:1439–1447

DOI 10.1007/s00122-012-1799-5

Article published: May 2012

The original publication is available at:

http://link.springer.com/article/10.1007%2Fs00122-012-1799-5

Abstract

Stalk lodging in maize causes annual yield losses between 5–20% worldwide. Many studies have indicated that maize stalk strength significantly negatively correlates with lodging observed in the field. Rind penetrometer resistance (RPR) measurements can be used to effectively evaluate maize stalk strength, but little is known about the genetic basis of this parameter. The objective of this study was to explore a genetic model and detect quantitative trait loci (QTL) of RPR and determine relationships between RPR and other stalk traits, especially cell wall chemical components. The RPR trait is quantitative in nature, and both additive and non-additive effects may be important for the improvement of RPR. Nine additive-effect QTLs covering nine chromosomes, except chromosome 5, and one pair of epistatic QTLs were detected for RPR. *CeSA11* involved in cellulose synthesis and *colorless2* involved in lignin synthesis were identified as putative candidate genes for RPR. Internode diameter (InD), fresh weight of internode (FreW), dry weight of internode (DryW), fresh weight and dry weight as well as cell wall components per unit volume significantly positively correlated with RPR. The internode water content (InW) significantly negatively correlated with RPR. Notably, for these traits significantly correlated with RPR, their QTL also co-localized with QTL of RPR. The consistent results obtained from correlation analysis and QTL mapping suggested the presence of pleiotropism or tight linkage between genes, and indicated that these different approaches may be used for cross authentication of relationships between different traits.

Chapter 3

QTL mapping of stalk bending strength in a recombinant inbred line maize population

Haixiao Hu, Wenxin Liu, Zhiyi Fu, Linda Homann, Frank Technow, Hongwu Wang, Chengliang Song, Shitu Li, Albrecht E. Melchinger, Shaojiang Chen

Haixiao Hu, Linda Homann, Frank Technow, Albrecht E. Melchinger, Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70599 Stuttgart, Germany

Wenxin Liu, Shaojiang Chen, National Maize Improvement Center of China, China Agricultural University, 2# Yuanmingyuan West Road, 100193 Beijing, China

Zhiyi Fu, Chengliang Song, Shitu Li, Applicational Mechanics Department, College of Science, China Agricultural University (East Campus), 17# Qinghua East Road, 100083 Beijing , China

Hongwu Wang, Institute of Crop Science, Chinese Academy of Agricultural Sciences, 12# Zhongguancun South Street, 100081 Beijing, China

Theor Appl Genet 126: 2257–2266

DOI10.1007/s00122-013-2132-7

Article published: September 2013

The original publication is available at:

http://link.springer.com/article/10.1007/s00122-013-2132-7

Abstract

Stalk bending strength (SBS) is a reliable indicator for evaluating stalk lodging resistance of maize plants. Based on biomechanical considerations, the maximum load exerted to breaking (F_{max}), the breaking moment (M_{max}) and critical stress (σ_{max}) are three important parameters for characterizing SBS. We investigated the genetic architecture of SBS by phenotyping F_{max}, M_{max} and σ_{max} of the fourth internode above the ground of maize plants in a population of 216 recombinant inbred lines (RILs) derived from the cross B73 × Ce03005 evaluated in four environments. Heritability of F_{max}, M_{max} and σ_{max} was 0.81, 0.79 and 0.75, respectively. F_{max} and σ_{max} positively correlated with several other related stalk traits. By using a linkage map with 129 SSR markers, we detected two, three and two quantitative trait loci (QTL) explaining 22.4%, 26.1% and 17.2% of the genotypic variance for F_{max}, M_{max} and σ_{max}, respectively. The QTL for F_{max}, M_{max} and σ_{max} located in adjacent bins 5.02 and 5.03 as well as in bin 10.04 for F_{max} were detected with high frequencies in cross validation. As our QTL mapping results suggested a complex polygenic inheritance for SBS related traits, we also evaluated the prediction accuracy of two genomic prediction methods (GBLUP and BayesB). In general, we found that both methods explained considerably higher proportions of the genetic variance than that obtained in QTL mapping with cross validation. Nevertheless, the identified QTL regions, spanning bins 5.02 and 5.03, could be used as a starting point for fine mapping and gene cloning of SBS traits.

Chapter 4

The genetic basis of haploid induction in maize identified with a novel genome-wide association method

Haixiao Hu[*], Tobias A. Schrag[*], Regina Peis, Sandra Unterseer, Wolfgang Schipprack, Shaojiang Chen, Jinsheng Lai, Jianbing Yan, Boddupalli M. Prasanna, Sudha K. Nair, Vijay Chaikam, Valeriu Rotarenco, Olga A. Shatskaya, Alexandra Zavalishina, Stefan Scholten, Chris-Carolin Schön, and Albrecht E. Melchinger

Haixiao Hu, Tobias A. Schrag, Wolfgang Schipprack, Stefan Scholten and Albrecht E. Melchinger, Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70599 Stuttgart, Germany

Regina Peis, Sandra Unterseer and Chris-Carolin Schön, Plant Breeding, TUM School of Life Sciences Weihenstephan, Technische Universität München, 85354 Freising, Germany

Shaojiang Chen and Jinsheng Lai, National Maize Improvement Center, China Agricultural University, 100193 Beijing, China

Jianbing Yan, National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, 430070 Wuhan, China

Boddupalli M. Prasanna, Sudha K. Nair, and Vijay Chaikam, International Maize and Wheat Improvement Center (CIMMYT)

Valeriu Rotarenco, Procera Genetics, Muncii 47, Fundulea, 915200 Calarasi, Romania

Olga A. Shatskaya, P.P. Luk'yanenko Krasnodar All-Russia Research and Development Institute of Agriculture, Russian Academy of Agricultural Sciences, 350012 Krasnodar-12, Russia

Alexandra Zavalishina, Department of Genetics Faculty of Biology, Saratov State University, 410012 Saratov, Russia

[*]These authors contributed equally to this work.

Genetics. Accepted Paper.

Abstract

In vivo haploid induction (HI) triggered by pollination with special intra-specific genotypes, called inducers, is unique to Zea mays L. within the plant kingdom and has revolutionized maize breeding during the last decade. However, the molecular mechanisms underlying HI in maize are still unclear. To investigate the genetic basis of HI, we developed a novel genome-wide association method, termed conditional haplotype extension (CHE) test that allows detection of selective sweeps even under almost perfect confounding between genetic ancestry and trait expression. Here, we applied this test to identify genomic regions required for HI expression and dissected the combined support interval (50.34 Mb) of the QTL *qhir1*, detected in a previous study, into two closely linked genomic segments relevant for HI expression. The first, termed *qhir11* (0.54 Mb), comprises an already fine-mapped region but was not diagnostic for differentiating inducers and non-inducers. The second segment, termed *qhir12* (3.97 Mb), had a haplotype allele common to all 53 inducer lines but not found in any of the 1,482 non-inducers. By comparing resequencing data of one inducer line CAU5 with 14 non-inducers, we detected in the *qhir12* region three candidate genes involved in DNA or amino acid binding, however none for *qhir11*. We propose that the CHE test can be utilized in introgression breeding and different fields of genetics to detect selective sweeps in heterogeneous genetic backgrounds.

Chapter 5

General Discussion

Choices of experiment design for investigation of stalk strength in maize

A central question in biology is whether the observed variation in a particular trait is due to environmental factors or biological factors (Visscher et al. 2008). Heritability is a concept which summarizes how heritable a phenotype of interest is. The divergence of narrow-sense heritability and broad-sense heritability is attributable to the phenotypic variance explained by purely additive affect or all genetic factors including additive, dominant and epistatic effects. Heritability is a key parameters used by both breeders and geneticists to determine the selection gain and statistical power of QTL mapping.

High heritability of RPR and SBS, ranging from 0.75 to 0.92, has been demonstrated by all QTL mapping studies (Flint-Garcia et al. 2003; Hu et al. 2012; Hu et al. 2013; Li et al. 2014) with the exception of Peiffer et al. (2013). The lower heritability of RPR reported by Peiffer et al. (2013) could be due to (i) relatively low number of replications in the field trials, and (ii) a wide range in the flowering time among NAM families.

Increasing the number of both genotypes and replications of each genotype can increase the power of QTL identification. A larger number of genotypes leads to more recombination events and reduced sampling error. In contrast, more replications enable to improve the precision of phenotypic evaluation and therefore increase the heritability of the target trait. Increasing the replication number of each genotype can be achieved by increasing the number of years, locations, or replications at each location in the field trial design. However, the optimum choice of the number of genotypes against the number of replications of each genotype in QTL mapping still needs further investigation. Sprague and Federer (1951) demonstrated

that to increase genetic advance, the optimum distribution of a given number of plots would be one replicate per location with an increase in number of locations and years. This design has been successfully applied for testing maize hybrids (Becker 2011). The low heritability of RPR obtained by Peiffer et al. (2013) indicates that the single replication design at each location might be not suitable for obtaining a good evaluation of stalk strength traits of inbred lines. The reason might be that inbreds are less vigorous compared to hybrids, which causes them to be easily affected by environmental or random factors. Therefore, to obtain more precise phenotypic values, it is recommended that at least two replications should be included at each location and at least five random plants should be evaluated in each replication in future studies of stalk strength.

Besides, parental lines selected for constructing a segregating population can also influence the precision of heritability estimation and QTL mapping. A large divergence of flowering time between two parental lines would lead to a large variation in the mapping population. Moreover, Gou et al. (2010) demonstrated that RPR measured before and after flowering time varies largely. Thus, phenotypic values of stalk strength would be confounded with flowering time if the population is measured at the same time. Heritability estimation and QTL mapping based on such phenotypic values would be inaccurate or biased. Alternatively, measuring RPR could be conducted separately for each genotype according to its actual flowering time. However, a phenotypic evaluation of this type would be too complex, and therefore it seems unrealistic for a large population such as NAM consisting of around 5,000 RILs. In conclusion, for QTL mapping of stalk strength, we suggest choosing parental lines similar in flowering time to reduce its influence on precision of phenotypic evaluation and make stalk strength measurements easy and simple.

Choice of growth stage to measure stalk strength

Some studies (Jia et al. 1992; Li et al. 2004) and our study showed that stalk strength is positively correlated with stalk diameter. However, some breeders observed that maize plants with larger diameter have higher lodging rate at harvest time compared to plants with smaller diameter (C. Huang, personal communication 2012).

To explain this apparent contradiction, it is worth mentioning that stalk strength is measured by researchers at flowering time or the milky stage, whereas breeders observe stalk lodging at harvest time. During flowering time or milky stage, besides the composition of the dry matter in the stalk, the water content of the stalk also plays a crucial role in stalk strength (Stojsin et al. 1991), because the turgor pressure from water in cells translates into increased stalk strength. This is supported by the highly positive correlation between RPR/SBS and fresh weight and water content of the internode in our study. However, at harvest time, water content in the stalk plays a less important role than in the milky stage, because the water content is usually dramatically reduced in the stalk. Moreover, stalks with larger diameter can lose more water than the smaller ones, resulting in more space inside the stalk and greatly decreased stalk strength (Wang et al. 1998).

This indicates that stalk strength is dynamic during the whole life of a maize plant (Gou et al. 2010). On one hand, a stalk strength study can focus on a specific growth stage chosen appropriate for the local climate. For example, in North China, storms usually happen during the milky stage. Thus, the study of stalk strength at this stage is most promising for reducing stalk lodging. On the other hand, evaluating the dynamic change of stalk strength across several growth stages could provide a better understanding of the regulations behind and be helpful for selecting breeding materials with high mechanical strength during the entire life span of the maize plant. According to breeders' observations, stalk lodging occurs most frequently during the twelve-leaf stage, the milky stage and the mature stage. Thus, these three stages are recommended to be included in future studies of stalk strength in maize.

Is there a complex gene network controlling maize stalk strength?

An important prerequisite of QTL mapping with bi-parental populations is that there is genetic segregation at the loci responsible for the target trait. There is consensus that complex traits, such as RPR and SBS in our study as well as yield, resistance, stress tolerance, are controlled by a large number of genes each having minor effect (Bernado 2008). Assuming a complex trait is controlled by 100 genes with minor effects and a pair of parent lines segregates only at 20 out of them, there would be at most 20 QTL to be detected in the derived segregating population. Another segregating population derived from two other lines might enable to detect another 20 QTL. This explains why different QTL of RPR were identified in different mapping populations (Bernardo 2002).

Among the four QTL mapping studies of RPR (Flint-Garcia et al. 2003; Hu et al. 2012; Peiffer et al. 2013; Li et al. 2014), 84 single-effect (additive and dominant) QTL and 12 epistatic interactions were detected across 33 segregating populations. All QTL had minor effects and none of them was consistently detected in all the four studies except the QTL in bin 3.06, which explained 4.87% to 18.7% of the phenotypic variance. Considering the sampling error resulting in upward estimation of QTL effect (Melchinger et al. 1998; Utz et al. 2000), cloning the gene underlying the QTL in bin 3.06 makes little sense for breeding, although it might be helpful to understand the mechanism of stalk strength.

High correlations and co-locations of QTL were observed in our study between RPR and other stalk traits, such as internode diameter, fresh weight of the internode, dry weight of the internode and some stalk chemical components. Moreover, genes involved in synthesis of lignin and cellulose were reported as putative genes affecting stalk strength (Flint-Garcia 2003; Ma et al. 2009; Li et al. 2014). This suggests that some genes for morphological traits and stalk chemical components might be involved in stalk strength.

Only one QTL with a small effect located on chromosome 10 had overlapping support intervals between RPR and SBS, which indicates that most likely different sets of genes affect

the two traits, although they might share a few genes. This is supported by their positive but medium correlations (around 0.40) observed in our study. Considering both traits contribute to stalk strength and relate to stalk lodging, the gene network behind stalk lodging seems to be much more complex.

From stalk mechanical strength to stalk lodging

Stalk lodging is of major concern for plant geneticists and breeders unless the mechanism of it is well understood and the entire gene network behind is completely deciphered. Occurrence of stalk lodging highly depends on environmental factors such as wind, rainfall, and their combination, which could vary greatly over years or locations. This means that stalk lodging cannot be constantly evaluated in multi-location field trials for two consecutive years. Moreover, there is still no effective way to artificially simulate stalk lodging with high precision. Thus, most geneticists and breeders proposed evaluating stalk lodging resistance of breeding materials indirectly by measuring their mechanical strength. Subsequently, various approaches have been developed such as stalk crushing strength (Zuber and Grogan 1961), RPR (Sibale 1992) and SBS (Kokubo et al. 1989). Despite these various attempts, neither a robust way to evaluate stalk lodging has been widely accepted nor the mechanism of stalk lodging is well understood. Drawbacks of established stalk strength measurements include: (i) they only measure a single internode of the maize stalk (which cannot reflect the lodging resistance of the whole plant), and (ii) nods of stalk are usually not considered, where stalk lodging could happen.

Procedures towards evaluating stalk lodging directly have been developed by some breeding companies. For example, DuPont Pioneer uses a mobile windy machine termed Boreas to evaluate lodging resistance of their breeding materials (Barreiro et al. 2008). However, the effectiveness of this approach is questioned. Some wind tunnel experiments conducted both in the field and laboratory indicated that wind load with very high speed from a single direction cannot break the stalk of maize hybrids studied , but their leaves have been seriously broken (Z.

Fu, personal communication 2014). However, under natural conditions in the field, what we actually observe is that the stalk is broken but the leaves are not. A further experiment (K. Chen, personal communication 2015) found that if the wind load acted on maize plants in two opposite directions in turn, the stalk was broken at a lower wind speed than in their first experiment. Therefore, to really understand the mechanism of stalk lodging, we need to shift the focus from simply measuring stalk strength to knowing more about what happens in the field in the event of stalk lodging. Based on the knowledge collected in the field, simulating stalk lodging with high precision in the laboratory could help for understanding the mechanical mechanism behind stalk lodging. A breakthrough in mechanical study of stalk lodging would improve our understanding of the genetic mechanism of stalk lodging and provide a direct method for stalk lodging evaluation in plant breeding.

The influence of genetic structure on GWAS

Association mapping aims at linking phenotypic variation to common sequence polymorphism in collections of unrelated individuals (Mezmouk et al. 2011). Given high-density genome-wide markers in a population, GWAS intends to identify a subset of markers, ranging from several to dozens, which are significantly correlated with the target trait. If a large number of markers are observed significantly in the single-marker test, it means that there is strong inflation due to genetic structure (Devlin and Roeder 1999).

Genetic structure affecting GWAS includes population structure (also called population stratification or population admixture) and familial relatedness (also called cryptic relatedness). Because of population structure, significant associations observed in the whole population might not actually exist in each subpopulation. Familial relatedness inflates frequency of alleles shared by related individuals, which possibly results in spurious associations. Five types of genetic structures (Fig. S4*A*-*E*) encountered in association mapping studies have been described by Yu and Buckler (2006), which include population structure, familial relatedness and their

combinations. All these types can be effectively controlled by corresponding statistical methods except the last one.

In our study, the segregation of HI is heavily confounded with population structure of inducers and non-inducers, and severe familial relatedness exists in the inducer group. This corresponds to a more complex situation of genetic structure, which was not described by Yu and Buckler (2006). This extreme case encountered in GWAS is termed as perfect confounding and no statistical method is currently available for performing an association analysis (D. Balding, J. Yu, and A. Price, personal communication 2013). The reason is that an association study is a statistical approach based on Mendelian segregation, and therefore it requires segregation of the target trait in each subpopulation if population stratification exists. This is instructive for future association studies regarding the collection of samples.

Application of selective sweep approaches in plant breeding populations

The severe familial relatedness among inducers is attributable to introgression breeding, which is commonly used by plant breeders. Briefly, a trait of particular interest controlled by one or a few genes with large effects (e.g. disease resistance) was found in one progenitor (a mutation or a genotype identified by screening a large number of germplasm). Subsequently, these genes were introgressed by breeders from this initial source (or a descendant derived from it, e.g. offspring, grandchild, great-grandchild) to a large number of progenies through crossing or backcrossing with other germplasm not carrying the target genes.

Given such a collection of lines derived from introgression breeding, the statistical approaches based on single marker tests such as the case-control association study or FLK test (Bonhomme 2010), were found futile in analyzing our dataset. Gautier and Naves (2011) demonstrated that haplotype-based methods, iHS and Rsb tests, successfully detected selective sweeps in cattle breeds with strong genetic structure. This suggested that these approaches might also be able to solve our problem. Nevertheless, all existing haplotype-based selective

sweep approaches are founded on the strong assumption that the individuals in the study population are unrelated (Voight et al. 2006; Tang et al. 2007; Fariello et al. 2013). The reason is that these approaches are specially designed for detecting selective signatures that have undergone a long-term evolution, such as the divergence of Europeans and Africans in humans and races of teosinte and domesticated maize. Under such a setting, researchers have enough candidates for selecting unrelated individuals and the collected individuals met the following characteristics: (i) the allele frequency in the regions unrelated to selection is randomly distributed, and (ii) the LD due to co-ancestry in these regions is maintained at a low level.

In the case of introgression breeding, the collected materials usually encompass no more than five generations, which is a rather short time scale. There might be many stretches with strong LD outside the selected genomic regions, which have nothing to do with selection of the target trait but are merely due to relatedness. Thus, we termed these stretches pseudo sweeps. The existence of pseudo sweeps in plant breeding population largely reduces the detection power of established haplotype-based selective sweep approaches and therefore invalidates them for plant breeding populations. This is demonstrated by the failure of these approaches in analyzing our dataset. Our CHE approach can eliminate the influence of pseudo sweeps based on a conditional haplotype extension. Therefore, we expect that this approach can fill the gap of detecting selective sweeps in plant breeding populations.

QTL and fine mapping of HI

When checking carefully the positions of *qhir1* detected in the four segregating populations described by Prigge et al. (2012), we found that none of them shares the same pair of flanking markers. Moreover, the intervals between flanking markers of the *qhir1* region from the F_2 population of cross 1680 × UH400 (1680-F_2) and the F_3 population of cross 1680 × UH400 (1680-F_3) have no overlap with that from the F_3 population of cross CML395 × UH400 (CML395-F_3). Generally, the QTL positions of *qhir1* identified in 1680-F_2 and 1680-F_3 shift to

an upstream region compared to those identified in CML395-F_3 and the F_3 population of cross CML495 × UH400 (CML495-F_3). Furthermore, the marker density was low at the *qhir1* and nearby regions in the four populations, which possibly resulted in a biased estimation of the position of the peak signal. Combined with a previous QTL mapping study (Barret et al. 2008) that located a QTL of HIR at a similar region, we can conclude that the likely position of the major QTL of HIR is located in bin 1.04, but its precise position cannot be assigned to an unique interval across different mapping populations. Thus, it is prudent to further confirm the QTL interval of *qhir1* before embarking on fine mapping.

The fine mapping study conducted by Dong et al. (2013) was based on the flanking markers of *qhir1* identified in populations 1680-F_2 and 1680-F_3, which means that the divergence of QTL positions was not actually considered. In their fine mapping study, Dong et al. (2013) searched for recombinants between marker umc1917 and bnlg1811, which is outside of the interval formed by the flanking markers of *qhir1* identified in population CML395-F_3. In addition, testing the phenotypic difference among progenies of F_2 recombinants in generation F_3 cannot completely rule out an influence of the genetic background. Briefly, if the *qhir1* has a strong hitchhiking effect but progeny testing employed for fine mapping focuses on only one piece of the whole hitchhiking segment, the significant difference observed in phenotypes between genotype categories of the small piece is also possibly due to other pieces belonging to the same hitchhiking segment. This problem can be overlooked by progeny testing in their fine mapping approach, because it disregards any genetic variation outside the focal region, underpinning the need to confirm the target region of HI in *qhir1* by alternative approaches.

In contrast, case-control association mapping and selective sweep approaches for fine-mapping a major QTL are less influenced by hitchhiking effects. With these approaches, we can have an overview of the interesting region and its neighbor regions across all individuals using graphical genotype software such as Flapjack (Milne et al. 2010). The stretch length and haplotype allele frequency of the interesting region due to hitchhiking and their decay due to

recombination can be clearly observed. This could help to prevent fine mapping from choosing a wrong region.

Genetic model and biological mechanism of HI

Although various attempts have been undertaken towards revealing the genetic architecture (Deimling et al. 1997; Barret et al. 2008; Prigge et al. 2012) and biological mechanism (Zhang et al. 2008; Li et al. 2009; Xu et al. 2013; Qiu et al. 2014) behind HI in maize, these issues are still not solved. Progress for revealing the genetic model of HI was made by Prigge et al. (2012). Their study corroborated that the major QTL (*qhir1*) of HI is located in bin 1.04 and they proposed a hypothesis that HI is controlled by a mandatory gene underlying *qhir1*, and other modifier genes with minor effects. In our study with a diversity panel of 51 worldwide inducers, we found that a long stretch with 4Mb, termed *qhir12*, is monomorphic across all inducers, but has no overlap with the fine mapping region, termed *qhir11*, of Dong et al. (2013). Moreover, *qhir12* is located inside the support interval of QTL *qhir1*. If we assume a mandatory gene of HI exists and is unique, further experiments are needed to confirm the genomic regions harboring the *qhir1*. Otherwise, there would be two alternative hypotheses: (i) both *qhir11* and *qhir12* are mandatory for HI; or (ii) each of them can trigger HI independently.

The mechanism of haploid occurrence is an interesting and important topic in plant species. In *Arabidopsis thaliana*, paternal haploids can be produced by crossing CENH3 mutants with wild types as pollinators. The mechanism is the CENH3-mediated genome elimination (Ravi and Chan 2010). In barley, haploid production is the result of crossing cultivated barley *Hordeum vulgare* as female and wild *Hordeum bulbosum* as male. A similar mechanism related to the CENH3 protein for haploid induction was revealed by Seymour et al. (2012). In wheat, haploids are produced by crossing wheat source germplasm with maize as pollinator (Pret'ová et al. 2006). Obviously, the mechanism is chromosome elimination due to distant hybridization. For another relative of maize, rice haploids are mainly produced by *in*

vitro anther culture and there is no inducer currently available (J. Pauk, personal communication 2015). Different from all the species above, maternal haploid induction in maize is based on pollination with pollens of the same species. Thus, the mechanism of maize HI is most likely different from other species mentioned above. Moreover, the CENH3 gene in maize is located in bin 6.05 (maizeGDB), and no QTL for HI has been detected there until now. Currently, there are two main hypotheses for explaining maternal haploid induction in maize. One is single fertilization (Barret et al. 2008) and the other is chromosome elimination (Zhang et al. 2008; Li et al. 2009; Xu et al. 2013; Qiu et al. 2014). However, these studies are only based on either phenotypic observation (Xu et al. 2013) or limited evidence from molecular markers (Li et al. 2009). Since borrowing information from other plant species might be futile, the crucial step for revealing the mechanism of HI in maize is cloning the major gene located in bin 1.04.

Selection strategies of breeding for HI and stalk strength with molecular markers

Quantitative traits can be classified into two categories: (i) those controlled by a few genes with large effects, and (ii) those controlled by many genes with small effects (Bernado 2008).

In our study, the HI trait is a typical example of the first category. The major QTL, *qhir1*, for HI explains up to 66% of generic variance (Prigge et al. 2012). Dong et al. (2013) further narrowed down the *qhir1* to a region of 243kb by screening recombinants in a large F_2 population from the cross 1680 × UH400. Subsequently, successful applications of combined MAS for developing new inducers have been reported from two independent breeding programs at China Agricultural University (CAU; Dong et al. 2014) and International Maize and Wheat Improvement Center (CIMMYT; S. Nair, personal communication 2015), respectively. In the CAU breeding program, two markers, X18 and X109, developed by their fine mapping study (Dong et al. 2013) were used for genotyping all plants in each generation. In contrast, CIMMYT

only used one marker, umc1917, identified in the QTL mapping study of Prigge et al. (2012). This indicates that application of MAS in selection of HI is feasible and effective. Moreover, the effectiveness of MAS based on preliminary QTL mapping results indicates strong LD in the genomic region harboring the major QTL *qhir1*.

In contrast, stalk strength represents a typical example of the second category of quantitative traits. No major QTL were observed in all QTL studies conducted on stalk strength traits so far. Therefore, pure MAS would be less effective than classical phenotypic selection, which is illustrated in our study of SBS (Hu et al. 2013). Combined MAS would be slightly more effective than phenotypic selection. However, it is very laborious compared with pure MAS. The third choice is genomic selection (Meuwissen et al. 2001, Albrecht et al. 2011). For SBS, a considerably higher proportion of the genetic variance was explained by genomic selection than by QTL mapping with cross validation (Hu et al. 2013). In the QTL mapping study of Peiffer et al. (2013), they also obtained high prediction accuracy between phenotypes of RPR and predicted line means. Therefore, considering the drawbacks of MAS and classical phenotypic selection, genomic selection is expected to be of great potential in selecting superior breeding materials with high stalk mechanical strength.

Chapter 6

Summary

Stalk lodging causes yield losses in maize cultivation ranging from 5 to 20% annually worldwide and stalk mechanical strength is widely accepted as an indirect indicator for its measurement. QTL mapping can reveal the genetic basis of stalk strength and provide information about markers suitable for marker-assisted selection (MAS). Constantly increasing market demands urge maize geneticists and breeders not only to enhance the field performance of new hybrids, but also to improve the breeding process. During the last decade, advances in the double haploid (DH) technology based on *in vivo* haploid induction (HI) shifted the breeding paradigm and greatly accelerated the breeding process in maize. Further spread of DH technology urgently demands a simple but efficient way for developing new inducers, which could be achieved by introducing the mandatory QTL/gene(s) of HI to advanced breeding lines. Therefore, the main goal of my thesis was to dissect the genetic architecture of stalk strength and detect the mandatory genomic region(s) of HI using genome-wide molecular markers.

Several methods have been developed and applied in the literature to evaluate stalk mechanical strength, among which the rind penetrometer resistance (RPR) is a simple, rapid and non-destructive measurement during data collection, whereas stalk bending strength (SBS) is more closely associated with stalk lodging in the field. According to common knowledge in the mechanics of materials, SBS is reflected by the maximum load exerted to breaking (F_{max}), the breaking moment (M_{max}) and the critical stress (σ_{max}). Thus, to have a complete understanding of the genetic architecture of stalk strength in maize, RPR and SBS (measured by F_{max}, M_{max} and σ_{max}) were used to characterize stalk strength in our study.

Utilizing a segregating population with 216 recombinant inbred lines, our analysis showed that stalk strength traits, RPR and SBS, have high heritability, ranging from 0.75 to 0.91. Nine QTL and one epistatic interaction between QTL were detected for RPR. Two, three and two QTL were detected for Fmax, Mmax and σmax, respectively. All QTL showed minor

effects and only one QTL on chromosome 10 had overlapping support intervals between RPR and SBS. Co-locations of QTL and high positive correlations between stalk strength traits and other stalk traits suggested presence of pleiotropism and a complex genetic architecture of stalk strength. Owing to lack of major QTL, MAS solely based on molecular markers was found to be less effective than classical phenotypic selection for stalk strength. However, for SBS we observed considerably higher proportions of genetic variance explained by a genomic selection approach than obtained in QTL mapping with cross validation. Therefore, genomic selection might be a promising tool to improve the efficiency of breeding for stalk strength.

All QTL mapping studies conducted hitherto for unraveling the genetic architecture of HI rate detected a major QTL, termed *qhir1*, in bin 1.04. Dong et al. (2013) further narrowed down this QTL to a 243 kb region. Considering the complex genetic architecture of HI and genetic background noise possibly affecting fine mapping of *qhir1*, we attempted to validate these results with an alternative approach before embarking on map-based gene isolation. Utilizing 51 maize haploid inducers and 1,482 non-inducers collected worldwide, we were able to investigate the genetic diversity between inducers and non-inducers and detect genomic regions mandatory for HI. The genetic diversity analyses indicated that the inducer group was clearly separated from other germplasm groups and had high familial relatedness. Analyzing our data by a case-control association approach failed because the segregation of HI was heavily confounded with population structure. Moreover, selective sweep approaches commonly used in the literature that are designed for capturing selective sweeps in a long-term evolutionary context failed due to high familial relatedness among inducers. To solve this problem, we developed a novel genome scan approach to detect fixed segments among inducers. With this approach, we detected a segment, termed *qhir12*, 4.0 Mb in length, within the support interval of the *qhir1*. This segment was the longest genomic segment detected by our novel approach and was entirely absent in all non-inducers analyzed. However, *qhir12* has no overlap with the fine mapping region of Dong et al. (2013), termed *qhir11*. This indicates that the genomic region harboring the mandatory gene of HI should be confirmed by further experiments to corroborate its existence and identify its location in the maize genome.

Chapter 7

Zusammenfassung

Weltweit führt Stängellager im Maisanbau zu Ertragsausfällen zwischen 5 und 20 %. Als indirekter Indikator für Lageranfälligkeit wird die mechanische Stabilität des Stängels angesehen. Die genetische Architektur des Merkmals Stängelstabilität kann durch eine QTL-Kartierung analysiert werden. Die mit dieser Methode identifizierten Marker-Merkmals-Assoziationen können anschließend für eine Marker-gestützte Selektion genutzt werden. Die stetig steigende Nachfrage nach verbesserten Maissorten drängt Genetiker und Züchter nicht nur zu Leistungssteigerungen bei neuen Hybridsorten, sondern auch zur Verbesserung des ganzen Zuchtverfahrens. Basierend auf der *in vivo* Haploideninduktion führten Fortschritte in der Doppelhaploiden(DH)-Technik im letzten Jahrzehnt zu einem Paradigmenwechsel bei der Entwicklung von Inzuchtlinien und einer beachtlichen Beschleunigung des Zuchtprozesses. Für viele Klimazonen, in denen Mais angebaut wird, stehen jedoch keine geeigneten Induktor-Genotypen zur Verfügung, was einer weiteren Verbreitung der DH-Technologie im Wege steht. Deshalb sind viele öffentliche Einrichtungen sowie kleinere Züchtungsunternehmen an einem einfachen und effizienten Verfahren zur Entwicklung eigener Induktorlinien interessiert. Eine Identifizierung von Major-QTL für Haploideninduktion und das Auffinden damit gekoppelter Marker wäre eine vielversprechende Lösung für dieses Problem. Die Hauptziele der vorliegenden Arbeit bestanden daher in der Analyse der genetischen Architektur des Merkmals Stängelstabilität, sowie dem Auffinden von Genomabschnitten, welche die Haploideninduktionsrate bestimmen.

In der Literatur wurden diverse Methoden zur Bestimmung der Stängelstabilität entwickelt und angewandt, wobei der Penetrationswiderstand der Rinde (RPR = rind

penetrometer resistance) sehr einfach, schnell und nicht-destruktiv erfassbar ist, wobei die Messung der Stängelbiegestärke (SBS = stalk bending strength) mit dem Stängellager unter Feldbedingungen als enger assoziiert gilt. Nach dem allgemeinen Kenntnistand der Werkstoffmechanik wird die SBS durch die Höchstbelastung beim Bruch (F_{max}), dem Bruchmoment (M_{max}) und dem kritischen Stress (σ_{max}) erfasst. Um ein möglichst vollständiges Bild der genetischen Architektur der Stängelstabilität bei Mais zu erlangen, wurden in dieser Studie sowohl RPR als auch SBS (gemessen als F_{max}, M_{max} und σ_{max}) zur Charakterisierung der Stängelstabilität herangezogen.

Die Untersuchung einer spaltenden Population aus 216 rekombinanten Inzuchtlinien zeigte, dass die Stängelstabilitätsmerkmale RPR und SBS eine hohe Heritabilität mit Werten zwischen 0,75 und 0,91 besitzen. Beim RPR konnten neun QTL und ein Paar epistatischer QTL gefunden werden, bei SBS zwei QTL für F_{max}, drei für M_{max} und zwei für σ_{max}. Alle QTL zeigten nur kleine Effekte und nur ein QTL auf Chromosom 10 wies eine überlappende Region zwischen RPR und SBS auf. Die Lokalisation der QTL und die hohen positiven Korrelationen zwischen den erfassten Merkmalen weisen auf eine pleiotrope Genwirkungsweise und eine komplexe genetische Architektur der Stängelstabilität hin. Aufgrund des Fehlens von Haupt-QTL wäre eine Marker-gestützte Selektion alleine deutlich weniger effizient als die klassische phänotypische Selektion. Bei den Komponenten von SBS erklärte die Methode der genomischen Selektion deutlich höhere Anteile an der genetischen Varianz als die in der QTL-Kartierung gefundenen Genomabschnitte. Daher scheint die genomische Selektion ein sehr aussichtsreicher Ansatz zur Verbesserung der Stängelstabilität zu sein.

Alle bisherigen QTL-Kartierungsstudien für das Merkmal Haploideninduktionsrate detektierten einen mit *qhir1* bezeichneten Haupt-QTL auf Chromosom 1. Dong et al. (2013) konnten diesen auf eine Region von 243 kb eingrenzen. In Anbetracht der komplexen genetischen Architektur der Haploideninduktionsrate und der durch den genetischen Hintergrund bedingten Unschärfe bei der Feinkartierung von *qhir1* scheint es vor einer

kartengestützten Klonierung deshalb ratsam, die Wirkung und Lokalisation dieses QTL mit einer anderen Methode zu bestätigen. Hierzu wurde ein Satz von 51 Induktorlinien und 1.482 Nicht-Induktoren aus einer weltweiten Kollektion von Maislinien verwendet. Analysen zur genetischen Diversität wiesen darauf hin, dass sich die Induktorlinien deutlich von den anderen Inzuchtlinien aus verschiedenen Formenkreisen unterscheiden und untereinander eng verwandt sind. Eine „Case-control“ Assoziationsstudie scheiterte bei dem vorliegenden Datensatz, da die Ausprägung der Haploideninduktion nahezu komplett mit der Populationsstruktur vermengt war. Aufgrund der engen Verwandtschaft unter den Induktorlinien waren auch sämtliche Ansätze für eine „selective sweep“ Analyse erfolglos, da dieses Verfahren speziell darauf ausgelegt ist, Langzeitänderungen im Genom unter Selektion zu erfassen. Im Gegensatz dazu reicht der Stammbaum von Zuchtstämmen, in die ein Gen aus einem gemeinsamen Vorfahren eingekreuzt wurde, gewöhnlich nicht mehr als fünf Generationen zurück. Dies bedeutet, dass neben den durch Selektion beeinflussten Genombereichen auch unbeeinflusste Genomsegmente, sogenannte „pseudo sweeps“, zu erwarten sind. Zur Lösung dieses Problems entwickelten wir ein neues Genome-Scan Verfahren, das fixierte Segmente innerhalb der Induktorlinien erfasst und dabei „pseudo sweeps“ mit großer Wahrscheinlichkeit ausschließt. Mit dieser Methode konnte innerhalb von *qhir1* ein Segment der Länge 4.0Mb gefunden werden. Dieses als *qhir12* bezeichnete Segment trat in allen Induktorlinien auf, wurde jedoch in keinem Nicht-Induktor gefunden, und wies auch keine Überlappung mit der von Dong et al. (2013) identifizierten Feinkartierungsregion *qhir11* auf. Dies unterstreicht, dass die Haploideninduktion auslösende Genomregion durch weitere Experimente bestätigt werden sollte, um ihre Existenz und Lage im Maisgenom zweifelsfrei zu belegen.

Chapter 8

References

Albrecht T, Wimmer V, Auinger H-J, Erbe M, Knaak C, Ouzunova M, Simianer H, Schön CC (2011) Genome-based prediction of testcross values in maize. Theor Appl Genet 123: 339–350

Albrecht T, Auinger HJ, Wimmer V, Ogutu JO, Knaak C, Ouzunova M, Piepho HP, Schön CC (2014) Genome-based prediction of maize hybrid performance across genetic groups, testers, locations, and years. Theor Appl Genet 127: 1375–1386

Anderson J, Chao S, Liu S (2007) Molecular breeding using a major QTL for *Fusarium* head blight resistance in wheat. Crop Sci 47: 112–119

Barreiro R, Carrigan L, Ghaffarzadeh M, Goldman DM, Hartman ME, Johnson DL, Steenhoek L(2008) Device and method for screening a plant population for wind damage resistance traits. United States Patent Publication. No. US 7412880 B2

Barret P, Brinkmann M, Beckert M (2008) A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize. Theor Appl Genet 117: 581–594

Basten CJ, Weir BS, and Zeng Z (1994) QTL Cartographer. Dep. of Statistics, North Carolina State Univ., Raleigh, NC, USA

Becker H (2011) Pflanzenzüchtung, 2nd edn. Ulmer, Stuttgart, pp146–165

Bernardo R (2002) Breeding for quantitative traits in plants. Stemma Press, Woodbury, MN

Bernardo R (2008) Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci 48: 1649–1664

Bonhomme M, Chevalet C, Servin B, Boitard S, Abdallah J et al (2010) Detecting selection in population trees: the Lewontin and Krakauer test extended. Genetics 186: 241–262

Chai Y, Hao X, Yang X, Allen WB, Li J, Yan J, Shen B, Li J (2012) Validation of DGAT1-2 polymorphisms associated with oil content and development of functional markers for molecular breeding of high-oil maize. Mol Breed 29: 939–949

Chang M, Coe EH (2009) Doubled haploids. In: Kriz AL, Larkins BA (eds) Molecular genetic approaches to maize improvement. Biotechnology in agriculture and forestry, Vol 63. Springer, Berlin

Chase SS (1969) Monoploids and monoploid derivatives of maize (Zeamays L.). Bot Rev 35: 117–167

Ching A, Rafalski JA, Luck S, Butruilie MG (2010) Genetic loci associated with mechanical stalk strength in maize. United States Patent Publication. No. US2010/0015623A1

Coe EH (1959) A line of maize with high haploid frequency. Am Nat 93: 381–382

Colbert TR, Darrah LL, Zuber MS (1984) Effect of recurrent selection for stalk crushing strength of agronomic characteristics and soluble stalk solids in maize. Crop Sci 24: 473–478

Deimling S, Röber FK, Geiger HH (1997) Methodology and genetics of in vivo haploid induction in maize (in German). Vortr Pflanzenzüchtg 38: 203–224

Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 55: 997−1004

Dong X, Xu X, Miao J, Li L, Zhang D, Mi X, Liu C, Tian X, Melchinger AE, Chen S (2013) Fine mapping of qhir1 influencing in vivo haploid induction in maize. Theor Appl Genet 126: 1713–1720

Dong X, Xu X, Li L, Liu C, Tian X, Li W, Chen S (2014) Marker-assisted selection and evaluation of high oil in vivo haploid inducers in maize. Mol Breeding 34: 1147–1158

Fariello MI, Boitard S, Naya H, San Cristobal M, Servin B (2013) Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. Genetics 193: 929–941

Fay JC, Wu I (2000) Hitchhiking under positive Darwinian selection. Genetics 155: 1405–1413

Flint-Garcia SA, Jampatong C, Darrah LL, McMullen MD (2003) Quantitative trait locus analysis of stalk strength in four maize populations. Crop Sci 43: 13–22

Frary A, Nesbitt TC, Frary A, Grandillo S, van der Knaap E, Cong B, Lui J, Meller J, Elber R, Alpert KB, Tanksley SD (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science 289: 85–88

Gautier M, Naves M (2011) Footprints of selection in the ancestral admixture of a New World Creole cattle breed. Mol Ecol 15: 3128–3143

Geiger HH (2009) Doubled haploids. In Handbook of Maize. Vol. II. Genetics and Genomics (Bennetzen JJ and Hake S, eds). Springer, Berlin, pp 641–657

Gere JM, Timoshenko SP (1984) Mechanics of Materials. Van Nostrand Reinhold Company Ltd, New York

Gou L, Huang J, Zhang B, Li T, Sun Y, Zhao M (2007) Effect of population density on stalk lodging resistant mechanism and agronomic characteristics of maize (In Chinese with English abstract). Acta Agronomica Sinia 33:1688–1695

Gou L, Huang JJ, Sun R, Ding Z, Dong Z, Zhao M (2010) Variation characteristic of stalk penetration strength of maize with different density-tolerance varieties (In Chinese with English abstract). Trans CSAE 26: 156–162

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.

Holland JB (2007) Genetic architecture of complex traits in plants. Curr Opin Plant Biol 10: 156–161

Hu H, Meng Y, Wang H, Liu H, Chen S (2012) Identifying quantitative trait loci and determining closely related stalk traits for rind penetrometer resistance in a high-oil maize population. Theor Appl Genet 124: 1439–1447

Hu H, Liu W, Fu Z, Homann L, Technow F, Wang H, Song C, Li S, Melchinger AE, Chen S (2013) QTL mapping of stalk bending strength in a recombinant inbred line maize population. Theor Appl Genet 126: 2257–2266

Hu J, Bao Y, Luo F et al (2000) Research on technology for measuring crop rind penetrometer resistance (In Chinese with English abstract). J Henan Agric Univ 34: 77–80

Jia Z, Bai Y (1992) Study on identification of lodging in maize inbred line (In Chinese with English abstract). China seeds 3: 30–32

Kearsey MJ, Farquhar AGL (1998) QTL analysis; where are we now? Heredity 80: 137−142

Kokubo A, Kuraishi S, Sakurai N (1989) Culm strength of barley correlation among maximum bending stress, cell wall dimensions, and cellulose content. Plant Physiol 91: 876–882

Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124: 743−756

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185–199.

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74: 175−195

Li D, Li H, Zhou L (2004) Selection and estimation of indices of stem-lodging resistance in corn (In Chinese with English abstract). J Northwest Science-Technology Univ Agric For 32: 53–55

Li H, Ye G, Wang J (2007) A modifiend algorithm for the improvement of composite interval mapping. Genetics 175: 361–374

Li L, Xu X, Jin W, Chen S (2009) Morphological and molecular evidences for DNA introgression in haploid induction via a high oil inducer CAUHOI in maize. Planta 230: 367–376

Li K, Yan JB, Li JS, Yang XH (2014) Genetic architecture of rind penetrometer resistance in two maize recombinant inbred line populations. BMC Plant Biology 14:152

Lincoln SE, Daly MJ, and Lander ES (1993) Mapping genes controlling quantitative traits using MAPMAKET/QTL version 1.1: A tutorial and reference manual. Whitehead Inst., Cambridge, MA

Lisec J, Meyer RC, Steinfath M, Redestig H, Becher M, Witucka-Wall H, Fiehn O, Torjek O, Selbig J, Altmann T et al (2008) Identification of metabolic and biomass QTL in Arabidopsis thaliana in a parallel analysis of RIL and IL populations. Plant J 53: 960–972

Ma Q (2009) The expression of caffeic acid 3-O-methyltransferase in two wheat genotypes differing in lodging resistance. J Exp Botany 60: 2763–2771

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

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

Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829

Mezmouk S, Dubreuil P, Bosio M, Décousset L, Charcosset A, Praud S, Mangin B (2011) Effect of population structure corrections on the results of association mapping tests in complex maize diversity panels. Theor Appl Genet 122:1149–1160

Micic Z, Hahn V, Bauer E, Schön CC, Melchinger AE (2005) QTL mapping of resistance to Sclerotinia midstalk rot in RIL of sunflower population NDBLOSsel × CM625. Theor Appl Genet 110: 1490–498

Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas TB, Flavell AJ Marshall D (2010) Flapjack – Graphical Genotype Visualization. Bioinformatics 26: 401–402

Nanda DK, Chase SS (1966) An embryo marker for detecting monoploids of maize (Zea mays L.). Crop Sci 6: 213–215

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335: 721–726

Peiffer JA, Flint-Garcia SA, De Leon N, McMullen MD, Kaeppler SM, Buckler ES (2013) The genetic architecture of maize stalk strength. PLoS One 8: e67066

Prasanna BM, Chaikam V, Mahuku G (2012) Doubled haploid technology in maize breeding: theory and practice. CIMMYT, Mexico

Pret'ová A, Obert B, Bartošová Z (2006) Haploid formation in maize, barley, flax, and potato. Protoplasma 228: 107–114

Prigge V, Xu X, Li L, Babu R, Chen S, Atlin GN, Melchinger AE (2012) New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. Genetics 190: 781–793

Qiu F, Liang Y, Li Y, Liu Y, Wang L, Zheng Y (2014) Morphological, cellular and molecular evidences of chromosome random elimination in vivo upon haploid induction in maize. Current Plant Biology 1: 83–90

Ravi M, Chan SW (2010) Haploid plants produced by centromere-mediated genome elimination. Nature 464:615–618

Ren Z, Gao J, Li L, Cai X, Huang W, Chao D, Zhu M, Wang Z, Luan S, Lin H (2005) A rice quantitative trait locus for salt tolerance encodes a sodium transporter. Nat Genet 37: 1141–1146

Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ et al (2002). Detecting recent positive selection in the human genome from haplotype structure. Nature 419: 832–837

Schön CC, Utz HF, Groh S, Truberg B, Openshaw S, Melchinger AE (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

Seymour DK, Filiault DL, Henry IH, Monson-Miller J, Ravi M et al (2012) Arabidopsis doubled haploids-rapid homozygous lines for quantitative trait locus mapping. Proc Natl Acad Sci USA 109: 4227–4232

Sibale EM, Darrah LL, Zuber MS (1992) Comparison of two rind penetrometers for measurement of stalk strength in maize. Maydica 37: 111–114

Sprague GF, Federer WT (1951) A comparison of variance components in corn yield trials: II. Error, year x variety, location x variety, and variety components. Agron J 43: 535–541

Stojsin R, Ivanovic M, Kojic L, Stojsin D (1991) Inheritance of grain yield and several stalk characteristics significant in resistance to stalk lodging maize (Zea mays L.). Maydica 36: 75–81

Tan L, Li X, Liu F, Sun X, Li C, Zhu Z et al (2008) Control of a key transition from prostrate to erect growth in rice domestication. Nat Genet 40: 1360–1364

Tang K, Thornton KR, Stoneking M(2007). A new approach for using genome scans to detect recent positive selection in the human genome. PLoS Biol 5: e171

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595

Technow F (2013) Extensions of genomic prediction methods and approaches for plant breeding. Diss. Univ. of Hohenheim, Stuttgart, Germany

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 3:197–203

Thompson DL (1963) Stalk strength of corn as measured by crushing strength and rind thickness. Crop Sci 3: 323–329

Utz HF, Melchinger AE (1996) PLABQTL: A program for composite interval mapping of QTL. Inst. Plant Breed. Seed Sci. Pop. Genet., Univ. of Hohenheim, Stuttgart, Germany

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

Visscher PM, Hill WG, Wray NR. 2008. Heritability in the genomics era – concepts and misconceptions. Nature Rev Genet 9:255–266

Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006). A map of recent positive selection in the human genome. PLoS Biol 4: e72

Wang J, Li H, Zhang L et al (2010) Users' manual of QTL IciMapping V3.0, Chinese Academy of Agricultural Sciences, Beijing, China

Wang Y, Li Q, Li C, Li A (1998) Studies on the culm quality and anatomy of wheat varieties (in Chinese with English abstract). Acta Agronomica Sinia 24: 452−458

Würschum T (2012) Mapping QTL for agronomic traits in breeding populations. Theor Appl Genet 125: 201–210

Würschum T, Reif JC, Kraft T, Janssen G, Zhao Y (2013) Genomic selection in sugar beet breeding populations. BMC Genet 14: 85

Xu X, Li L, Dong X, Jin W, Melchinger AE, Chen S (2013) Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. J Exp Bot 64: 1083–1096

Yang J, Hu CC, Ye XZ, Zhu J (2010) QTL Network 2.2 user manual: software for mapping QTL with epistatic and QE interaction effects. Zhejiang University, Hanzhou, China

Yang Q, Zhang D, Xu M (2012) A sequential QTL fine-mapping strategy using recombinant-derived progeny. J Integr Plant Biol 54: 228–237

Yu J, Buckler ES (2006) Genetic association mapping and genome organization of maize. Curr Opin Biotechnol 17: 155–160

Yuan Z, Feng B, Zhao A, Liang A (2002) Dynamic analysis and comprehensive evaluation of crop-stem lodging resistance (in Chinese with English abstract). Trans CSAE 18: 30–31

Zhang Z, Qiu F, Liu Y, Ma K, Li Z, Xu S (2008) Chromsome elimination and in vivo haploid induction by stock 6-derived inducer line in maize (*Zea mays* L.). Plant Cell Rep 27: 1851–1860

Zhao X, Tan G, Xing Y, Wei L, Chao Q, Zuo W, Lubberstedt T, Xu M (2012) Marker-assisted introgression of qHSR1 to improve maize resistance to head smut. Mol Breeding 30: 1077–1088

Zhao Y, Gowda M, Liu W, Würschum T, Maurer HP, Longin FH, Ranc N, Reif JC (2012) Accuracy of genomic selection in European maize elite breeding populations. Theor Appl Genet 124: 769–776

Zeng Z (1994) Precision mapping of quantitative trait loci. Genetics 136: 1457–1468

Zheng P, Allen WB, Roesler K, Williams ME, Zhang SR, Li JM, Glassman K, Ranch J, Nubel D, Solawetz W, Bhattramakki D, Llaca V, Deschamps S, Zhong G, Tarczynski MC, Shen B (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat Genet 40: 367–372

Zhu J (2000) Mixed linear model approaches for analyzing genetic models of complex quantitative traits. J Zhejiang Univ (Sci) 1: 78–90

Zuber MS, Grogan CO (1961) A new technique for measuring stalk strength in corn. Crop Sci 1: 378–380

Zuo W, Chao Q, Zhang N, Ye J, Tan G, Li B, Xing Y, Zhang B, Liu H, Fengler KA, Zhao J, Zhao X, Chen Y, Lai J, Yan J, Xu M (2015) A maize wall-associated kinase confers quantitative resistance to head smut. Nat Genet 47: 151–157

Acknowledgements

I am indebted to many individuals who have enabled the completion of this dissertation. First of all, I am grateful to my academic supervisor Prof. Dr. A. E. Melchinger, who gave me the opportunity to do my PhD at University of Hohenheim. Prof. Melchinger works very hard in scientific research and in supervising/ training his PhD students. I benefit a lot from his face-to-face meeting every two/three weeks during my PhD study. His hard working attitude will motivate me forever in my future career.

I am also grateful to Dr. T. Schrag, who gave me great support in data analysis, especially, at the beginning of my PhD. Dr. Schrag is very skilled in database management, R programming, use of office software, etc. I learned a lot from him in structuring the database, preparing meeting protocols, etc. All of these made my work more efficient.

I thank Prof. C. Schön, Prof. C. Zebitz and Prof. T. Streck for their volunteer work for my Ph.D. oral examination. I would like to give special thanks to Prof. Schön who also gave me great support in my maize haploid induction study.

Many thanks to Prof. Dr. Shaojiang Chen, Prof. Dr. Zhiyi Fu, Dr. Regina Peis, Dr. Wenxin Liu, Dr. Hongwu Wang, Dr. Frank Technow, Ms. Sandra Unterseer and Ms. Linda Homann as well as other co-authors and colleagues who contributed to all my publications. Without their strong support, my research would not have been published quickly.

I also thank Mrs. Helga Kösling, Dr. Xuefei Mi, Sen Han, José Marulanda, Markus Winter. Mrs. Kösling does excellent secretary work for our group. She is always patient to help me any time I asked her for help. Dr. Mi helped me a lot in adapting the life in Germany at the beginning of my PhD. Sen and José helped me a lot both in wiring this dissertation and in life. Markus helped me translate the summary from English to German.

I would like to express my appreciation to all the colleagues in the Institute of Plant Breeding, Seed Science and Population Genetics, especially to the PhD students who work in the same time period as me. I miss those days we together playing football, having Christmas parties, discussing questions, etc.

Finally, I owe my deepest gratitude to my family and especially my wife, Yujie Meng. She always encourages me any time either I felt stressful or got stuck in my research.

Curriculum Vitae

NAME	Haixiao Hu
BIRTH	31 October 1985 in Shandong, China

SCHOOL EDUCATION

1992 – 1997	Yinzhuang Primary School, Jinan, China
1997 – 2000	Kongcun Junior Middle School, Jinan, China
2000 – 2004	Pingyin No.3 Senior Middle School, Jinan, China

UNIVERSITY EDUCATION

2004 – 2008	B.Sc. in Agronomy, Shandong Agricultural University, Tai'an, China
2008 – 2011	M.Sc. in Plant Genetics and Breeding, China Agricultural University, Beijing, China
2012 – 2016	Doctoral student in Applied Genetics and Plant Breeding (Prof. Dr. A. E. Melchinger), University of Hohenheim, Stuttgart, Germany

WORK EXPERIENCE

2011 – 2012	Research assistant in China Agricultural University, Beijing, China

Haixiao Hu

Hohenheim, 18.02.2016