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

Genome-wide association mapping of molecular and physiological component traits in maize

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

von
Master of Science
Christian Riedelsheimer
aus Affing-Mühlhausen

Stuttgart-Hohenheim
2013

Die vorliegende Arbeit wurde am 31.10.2013 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: 16.12.2013

1. Prodekan:	Prof. Dr. M. Rodehutscord
Berichterstattter, 1. Prüfer:	Prof. Dr. A.E. Melchinger
Mitberichterstatter, 2. Prüfer:	Prof. Dr. L. Willmitzer
3. Prüfer:	Prof. Dr. J. Bennewitz

Contents

[1] Riedelsheimer, C., J. Lisec, A. Czedik-Eysenberg, R. Sulpice, A. Flis, C. Grieder, T. Altmann, M. Stitt, L. Willmitzer, and A. E. Melchinger, 2012. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl. Acad. Sci. USA* 109:8872-8877.

[2] Riedelsheimer, C., Y. Brotman, M. Méret, A. E. Melchinger, and L. Willmitzer , 2013. The maize leaf lipidome shows multilevel genetic control and high predictive value for agronomic traits. *Sci. Rep.* 3:2479.

[3] Strigens, A., N. M. Freitag, X. Gilbert, C. Grieder, C. Riedelsheimer, T. A. Schrag, R. Messmer, and A. E. Melchinger, 2013. Association mapping for chilling tolerance in elite flint and dent maize inbred lines evaluated in growth chamber and field experiments. *Plant Cell Environ.* 36:1871-1887.

Abbreviations

GS	Genomic selection
GWA	Genome-wide association
GWAS	Genome-wide association study
LD	Linkage disequilibrium
LM	Linkage mapping
MS	Mass spectroscopy
NAM	Nested association mapping
QTL	Quantitative trait loci

1 General Introduction

Maize has a long and fruitful history of genetics studies aiming at understanding the inheritance of quantitative traits. During the last twenty years, catalogues of specific results for quantitative trait loci (QTL) have been assembled by conducting numerous linkage mapping (LM) studies with biparental populations that carry genetic mosaics of two contrasting parental inbred lines (Mauricio, 2001). Despite the high power of LM for detecting QTL specific to the parental lines, however, most detected QTL were not subsequently fine mapped for identifying the underlying causal genetic variants. Thus, most LM studies have contributed only little in deepening our functional understanding of how complex traits are regulated at the genetic level. Most detected QTL were also not used in marker-assisted breeding (MAS) (Bernardo, 2008), mainly because the amount of explained genetic variance of the detected QTL turned out to be only weakly transferable to different populations and backgrounds (Melchinger *et al.*, 1998; Xu and Crouch, 2008), especially if no major QTL are present.

GWAS as an alternative to LM

The development of inexpensive high-throughput genotyping platforms such as the Illumina MaizeSNP50 Beadchip (Ganal *et al.*, 2011) or genotyping-by-sequencing (Elshire *et al.*, 2011) has generated new hope to overcome the limitations of LM. Genotyping costs have fallen so much that they are now frequently among the least-expensive parts of an experiment (Wallace *et al.*, 2014). With such an abundant marker density, it became possible to conduct genome-wide association studies (GWAS) by exploiting ancestral linkage disequilibrium (LD) in a diverse population capturing a much broader diversity than the biparental populations used in LM (Fig. 1.1). Depending on the level of LD in the population, the resolution can be up to the single nucleotide level.

GWAS was first developed as a necessity for large-scale human studies and has been extremely popular in this field. As of May 30$^{\text{th}}$, 2013, 1,613 human GWA studies have been published and their specific results combined in a public database (`http://www.genome.gov/gwastudies/`). In contrast to human genetics, the application of GWA mapping to plant populations and especially elite breeding material is hampered by high levels of population structure and cryptic relatedness which can lead to spurious associations (Astle and Balding, 2009). However, powerful techniques became available for decoupling genetic associations with confounding factors (Sillanpää, 2011; Yu *et al.*, 2006). Their application with current genotyping platforms was however only possible after the development of variance component based computational

algorithms which allow an extremely efficient testing of thousands to millions of SNPs while correcting for spurious associations in a mixed model framework (Svishcheva and Axenovich, 2012; Zhang *et al.*, 2010).

Meanwhile, several GWA studies in maize revealed that most traits are highly polygenic, *e.g.* controlled by a large number of small effect QTL. The largest QTL detected for most maize traits typically explained $< 5\%$ of the phenotypic variance (Wallace *et al.*, 2014). Interestingly, this has been even found for traits like flowering time (Buckler *et al.*, 2009), disease resistances (Kump *et al.*, 2011; Poland *et al.*, 2011), kernel starch, protein, and oil content (Cook *et al.*, 2011), or morphological traits such as leaf architecture (Tian *et al.*, 2011) which were initially expected to be genetically less complex than *e.g.* grain or biomass yield. The total summed up variance explained by all QTL detected in these studies is typically well below 30 %, *e. g.* 25.9 % for upper leaf angle, 23.2 % for leaf length, 23.3 % for leaf width, and 21.6 % for southern leaf blight (calculated from supporting information of Tian *et al.* (2011) and Kump *et al.* (2011)).

It is important to note that these results were not the result of a very limited population size as several of theses studies analysed the nested association mapping (NAM) panel comprising $\approx 5,000$ genotypes, a number which is barely manageable even for large public organizations. Thus, the majority of genetic factors underlying these traits are still unknown and unravelling the chain from the genes to the phenotype is still largely unresolved for most quantitative traits in maize. As this is scientifically dissatisfying, the problem has been coined 'bridging the genotype-phenotype gap' and to do so remains a big challenge. An obvious approach would be to further extend the population size to increase the power for detecting QTL with even smaller effect sizes. In fact, GWA studies in maize with more than 30,000 genotypes are currently conducted (Ed Buckler, Cornell, personal communication).

This thesis research aims to present an alternative route by mapping not the polygenic trait of primary interest itself, but genetically correlated molecular and physiological component traits. As such components represent biological sub-processes underlying the trait of interest, they are supposed to be genetically less complex and thus, more suitable for genetic mapping. This approach is demonstrated with (i) biomass yield by using metabolites and lipids as molecular component traits and (ii) chilling sensitivity by using physiological component traits such as photosynthesis parameters derived from chlorophyll fluorescence measurements.

Metabolites and lipids as molecular component traits of biomass yield

Metabolomics refers to the mass spectrometry-based quantitative measurement of hundreds of different biochemical compounds from a wide range of chemical classes within a single sample (Fig. 1.2) (Saito and Matsuda, 2010). Whereas metabolomics is already a rather mature field in medicine (Suhre and Gieger, 2012), it has received larger attention in plants only in recent

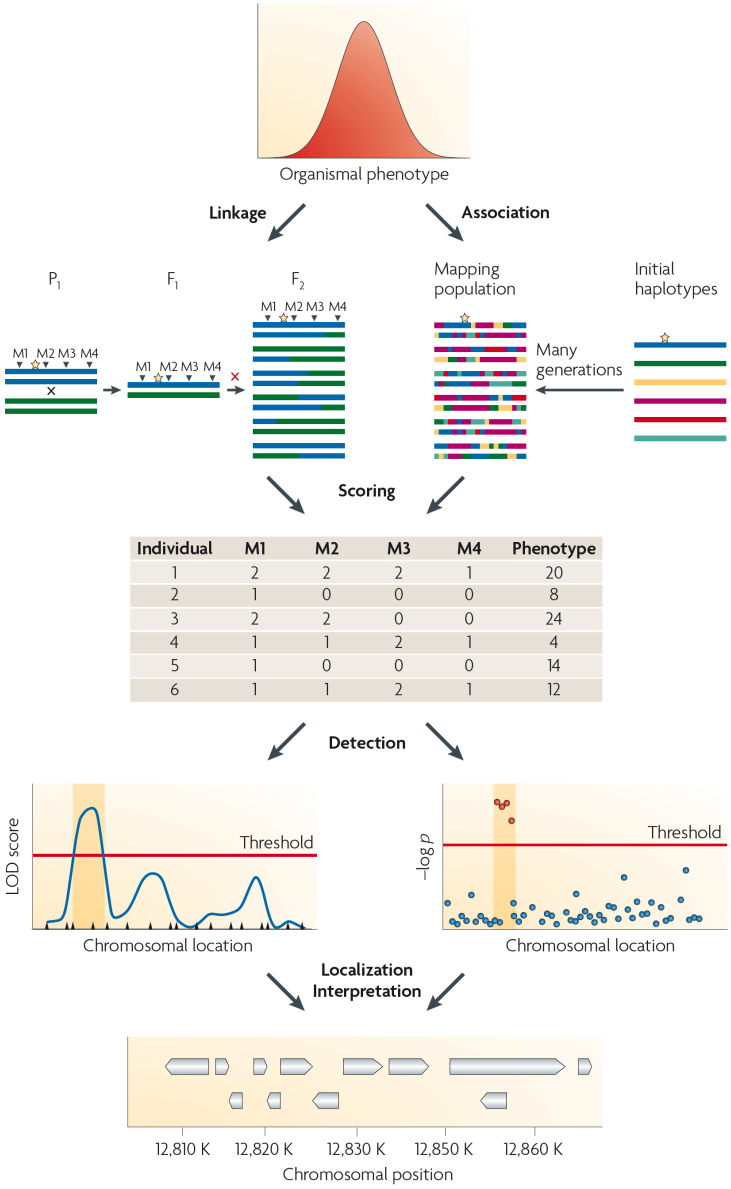

Figure 1.1: Conceptual overlappings and differences between linkage mapping (left) and genome-wide association mapping (right) (Mackay *et al.*, 2009).

years after analytical methods for blood and human tissues could be adapted to plant tissues (De Vos *et al.*, 2007; Lisec *et al.*, 2006).

An important finding of some early metabolomics studies in *Arabidopsis thaliana* was that an array of metabolites can be linked to biomass accumulation (Meyer *et al.*, 2007; Sulpice *et al.*, 2009). In turn, from a physiological perspective, biomass accumulation can be seen as the plants' ultimate outcome of its metabolic performance (Stitt *et al.*, 2010). The tight connection between metabolism and growth can arise from different reasons: either a high supply of metabolites triggers growth, or growth drains metabolites to a minimum tolerable level. Alternatively, metabolites may exert control on growth not only by acting as substrates for the synthesis of cellular components but also by acting as signals that are sensed leading to subsequent changes in growth (Meyer *et al.*, 2007; Stitt *et al.*, 2010).

Because the correct functioning of metabolic networks is crucial for proper development, the regulation of individual metabolites is tightly controlled (Kooke and Keurentjes, 2012). Thus, their genetic architecture is probably less complex than for most agronomic traits. As a consequence, instead of trying to map the genetic factors of highly polygenic traits, it has been suggested to study the inheritance of genetically more simple metabolites representing distinct physiological sub-processes underlying the trait of interest (Keurentjes, 2009)

Previous genetic analysis of metabolites are rather limited and restricted to either single compounds like carotenoids (Wong *et al.*, 2004) or model plants like *Arabidopsis* in which metabolic QTL could be successfully mapped with LM of recombinant inbred lines (Lisec *et al.*, 2008). In this thesis, our objective was to significantly expand this "genetical metabolomics" coined approach (Keurentjes, 2009) by moving (i) from the model plant *Arabidopsis* to the staple crop maize, (ii) from well-controlled greenhouse conditions to field conditions, and (iii) from LM to GWAS. In contrast to *Arabidopsis*, where hundreds of plants can be easily cultivated under controlled conditions in the greenhouse without a sophisticated randomization, the application of metabolomics in the field requires more consideration of the randomization layout. Thus, one important objective in the first place was to develop a sampling and processing scheme which allows to integrate the field randomization with the processing layout for measuring metabolites in the lab.

Recently, the lipidome has emerged as an especially rich subgroup of the metabolome (Mutch *et al.*, 2006). The lipidome contains hundreds to thousands of individual lipids species showing an enormous chemical diversity due to the high plasticity of the underlying biosynthetic machinery (Broun, 1998). Commonly only known for their role as a storage compound, lipids are also involved in many other processes including cell integrity, membran formation and scaffolding for membrane proteins, energy storage, and cell signaling (Brown and Murphy, 2009). In this study, we seeked for the first time to (i) genetically characterize the diversity of the lipidome by GWAS and (ii) explored its connection to complex traits in maize measured at the testcross level in multi-environment field trials.

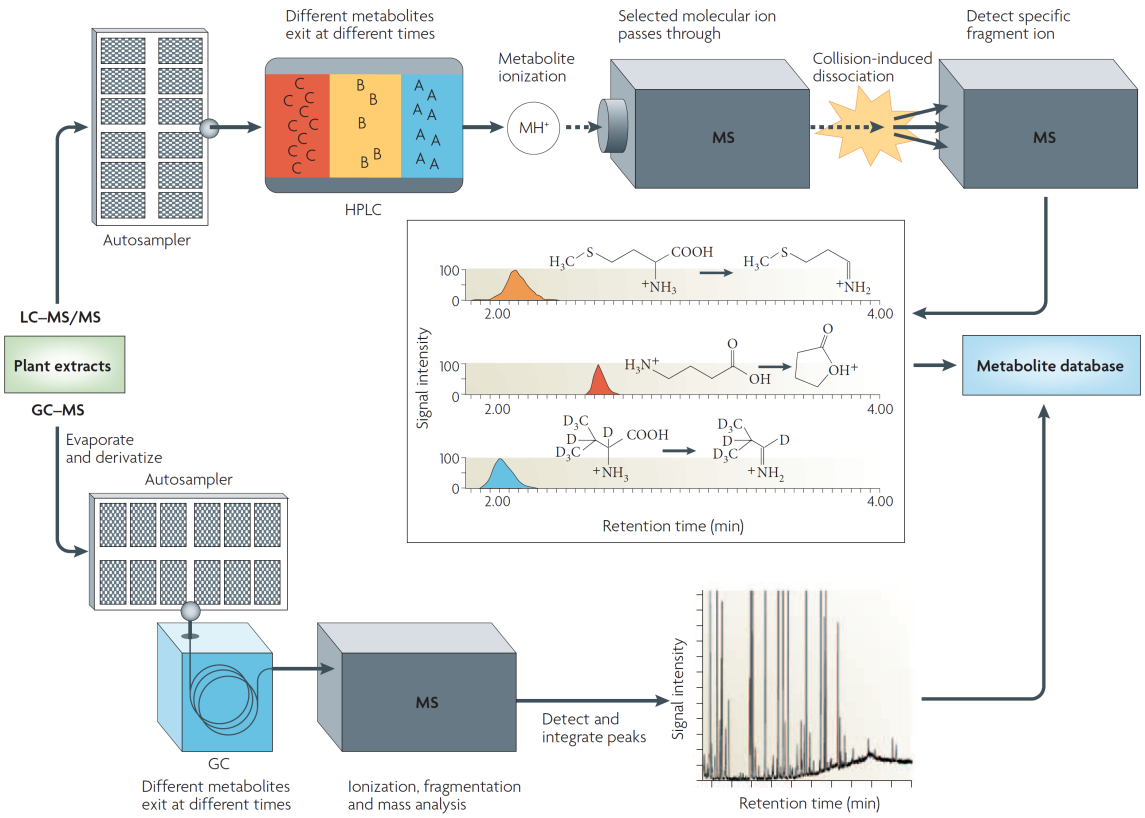

Figure 1.2: Untargeted metabolic profiling of a single plant sample by mass spectrometry with chromatographic separation (Last *et al.*, 2007).

Physiological component traits of chilling sensitivity

Chilling sensitivity is of increasing relevance for expanding the cultivation of maize into cooler regions (Frei, 2000). However, fluctuating climatic conditions and high genotype × environment interactions hamper its evaluation in the field (Presterl *et al.*, 2007) and thus aggravate the identification of cold tolerant genotypes by phenotypic selection. Hence, marker-based approaches would be a promising alternative to increase the selection efficiency for chilling sensitivity.

Unfortunately, previous LM studies for chilling tolerance yielded only erratic results, probably because of the large genetic complexity of the trait. In a large biparental population of 720 doubled haploid lines, the seven and ten QTL found for line *per se* and testcross performance explained without cross-validation on average only 11.3 % and 4.3 % of the phenotypic variance, respectively (Presterl *et al.*, 2007). In the intermated B73 × Mo17 (IBM) population, the detected QTL explained only 3.7 % of the genetic variance after cross-validation (Rodríguez *et al.*, 2008). In this case, marker-assisted selection could therefore not be recommended. The fact that chilling sensitivity is no simple trait is reflected in the deep and major changes with which plants react after perception of cold temperatures. These involve changes in gene and osmotic regulation, hormone and energy balance as well as major modifications of membranes and cell walls (Xin, 2000) and reconfigurations of metabolic networks (Guy *et al.*, 2008).

Whereas changes at these levels are difficult to measure in populations of hundreds of genotypes, many physiological reactions in response of cold temperatures are easier detectable, at least in growth chambers with a controlled temperature regime (Table 1.1). Similar to metabolites for biomass yield, genetic analysis of these physiological parameters would allow to genetically dissect physiological sub-processes of chilling tolerance, which are probably genetically less complex, easier to interpret, and probably more reliable than barely significant minor QTL found for chilling sensitivity itself. Building on previous LM results for these parameters, we aimed to perform GWAS with high precision to dissect the genetic causes underlying the observed phenotypic differences in chilling sensitivity in a parallel evaluation in both field environments and growth chambers.

Objectives

The goal of this thesis research was to examine the feasibility to dissect the genetically complex traits biomass yield and chilling sensitivity in maize by performing GWAS on genetically simpler molecular and physiological component traits correlated with the target trait itself. In particular, the objectives were to

1. develop a sampling and randomization procedure for the application of metabolomics and lipidomics in large-scale field trials of maize inbred lines;

2. explore repeatabilities and the correlation patterns in the leaf metabolome and lipidome in a diversity panel of maize inbred lines grown under field conditions;

3. perform GWAS of the leaf metabolome and lipidome using 56k SNPs;

4. perform GWAS of physiological component traits of chilling sensitivity in two diversity panels of maize inbred lines grown in parallel under field conditions and growth chambers;

5. identify plausible candidate genes underlying QTL mapped for metabolites, lipids, and physiological traits;

6. explore whether and how the genotype-phenotype gap of complex traits in field-grown maize can be narrowed by genetically characterized component traits.

Table 1.1: Examples of physiological processes / traits and associated measurable parameters which have been associated with chilling sensitivity in maize.

Physiological process / trait	Parameter	Source
Plant morphology	Specific leaf area	Hund *et al.* (2005); Verheul *et al.* (1996)
Biomass partitioning	Root/shoot ratio	Tollenaar (1989)
Photosynthetic performance	Chlorophyll content	Lee *et al.* (2002)
	Quantum efficiency of PSII (Ψ_{PSII})	Fracheboud (2002); Hund *et al.* (2005) Lee *et al.* (2002)
Photoinhibition	Ratio of variable to maximum fluorescence (F_v/F_m)	Fracheboud (2002); Ortiz-Lopez *et al.* (1990)
Leaf gas exchange	Carbon exchange rate	Lee *et al.* (2002)
	Stomatal conductance	Lee *et al.* (2002)

2 Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize

Christian Riedelsheimer[1a], Jan Lisec[2a], Angelika Czedik-Eysenberg[3], Ronan Sulpice[3], Anna Flis[3], Christoph Grieder[1], Thomas Altmann[4], Mark Stitt[3], Lothar Willmitzer[2,5], and Albrecht E. Melchinger[1]

[1]Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany

Departments of [2]Molecular Physiology and [3]Metabolic Networks, Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam, Germany

[4]Department Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06446 Gatersleben, Germany

[5]King Abdulaziz University, Jeddah 21589, Saudi Arabia

[a]These authors contributed equally

Proc. Natl. Acad. Sci. USA 109:8872-8877 (2012).

The original publication is available at
`http://dx.doi.org/10.1073/pnas.1120813109`

Abstract

The diversity of metabolites found in plants is by far greater than in most other organisms. Metabolic profiling techniques, which measure many of these compounds simultaneously, enabled investigating the regulation of metabolic networks and proved to be useful for predicting

important agronomic traits. However, little is known about the genetic basis of metabolites in crops such as maize. Here, a set of 289 diverse maize inbred lines was genotyped with 56,110 SNPs and assayed for 118 biochemical compounds in the leaves of young plants, as well as for agronomic traits of mature plants in field trials. Metabolite concentrations had on average a repeatability of 0.73 and showed a correlation pattern that largely reflected their functional grouping. Genome-wide association mapping with correction for population structure and cryptic relatedness identified for 26 distinct metabolites strong associations with SNPs, explaining up to 32.0% of the observed genetic variance. On nine chromosomes, we detected 15 distinct SNP-metabolite associations, each of which explained more then 15% of the genetic variance. For lignin precursors, including p-coumaric acid and caffeic acid, we found strong associations (P-values 2.7×10^{-10} to 3.9×10^{-18}) with a region on chromosome 9 harboring cinnamoyl-CoA reductase, a key enzyme in monolignol synthesis and a target for improving the quality of lignocellulosic biomass by genetic engineering approaches. Moreover, lignin precursors correlated significantly with lignin content, plant height, and dry matter yield, suggesting that metabolites represent promising connecting links for narrowing the genotype-phenotype gap of complex agronomic traits.

3 The maize leaf lipidome shows multilevel genetic control and high predictive value for agronomic traits

Christian Riedelsheimer[1a], Yariv Brotman[2a], Michaël Méret[2a], Albrecht E. Melchinger[1], and Lothar Willmitzer[2]

[1]Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany

[2]Department Molecular Physiology, Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam, Germany

[a]These authors contributed equally

Sci. Rep. 3:2479 (2013)

The original publication is available at
`http://dx.doi.org/10.1038/srep02479`

Abstract

Although the plant lipidome shows an enormous level of structural and functional diversity, our knowledge about its genetic control and its connection to whole-plant phenotypes is very limited. Here, we profiled 563 lipid species with UPLC-FT-MS in 289 field-grown inbred lines genotyped with 56,110 SNPs, Genome-wide association study identified 174 associations for 76 lipids explaining up to 31.4% of the genetic variance (P-value 8.4×10^{-18}). Candidate genes were found for lipid synthesis, breakdown, transfer, and protection against peroxidation. The detected SNP-lipid associations could be grouped into associations with 1) individual lipids, 2) lipids from one biochemical class, and 3) lipids from several classes, suggesting a multilevel genetic control architecture. We further found a strong connection between the lipidome and agronomic traits in field-evaluated hybrid progeny. A cross-validated prediction model yielded

correlations of up to 0.78 suggesting that the lipidome accurately predicts agronomic traits relevant in hybrid maize breeding.

4 Association mapping for chilling tolerance in elite flint and dent maize inbred lines evaluated in growth chambers and field experiments

Alexander Strigens[1], Niclas M. Freitag[2], Xavier Gilbert[2], Christoph Grieder[1], Christian Riedelsheimer[1], Tobias A. Schrag[1], Rainer Messmer[2], and Albrecht E. Melchinger[1]

[1]Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany

[2]Group of Crop Science, Swiss Federal Institute of Technology Zurich, Zurich, Switzerland

Plant Cell Environ. 36:1871-1887 (2013)

The original publication is available at
http://dx.doi.org/10.1111/pce.12096

Abstract

Chilling sensitivity of maize is a strong limitation for its cultivation in the cooler areas of the northern and southern hemisphere because reduced growth in early stages impairs on later biomass accumulation. Efficient breeding for chilling tolerance is hampered by both the complex physiological response of maize to chilling temperatures and the difficulty to accurately measure chilling tolerance in the field under fluctuating climatic conditions. For this research, we used genome-wide association (GWA) mapping to identify genes underlying chilling tolerance under both controlled and field conditions in a broad germplasm collection of 375 maize inbred lines genotyped with 56,110 single nucleotide polymorphism (SNP). We identified 19 highly significant association signals explaining between 5.7 and 52.5% of the phenotypic variance observed for early growth and chlorophyll fluorescence parameters. The allelic effect of several SNPs identified for early growth was associated with temperature and

incident radiation. Candidate genes involved in ethylene signalling, brassinolide, and lignin biosynthesis were found in their vicinity. The frequent involvement of candidate genes into signalling or gene expression regulation underlines the complex response of photosynthetic performance and early growth to climatic conditions, and supports pleiotropism as a major cause of co-locations of quantitative trait loci for these highly polygenic traits.

5 General Discussion

Genome-wide association studies (GWAS) link genetic variants to complex traits at high precision by exploiting ancestral linkage disequilibrium (LD) between genetic markers and causal variants in diverse population (Rafalski, 2010; Stich and Melchinger, 2010). To cope with the polygenic architecture of most agronomic traits, our objective was to investigate the feasibility to apply GWAS to molecular and physiological component traits (Fig. 5.1) representing genetically often simpler controlled sub-processes of the target trait itself.

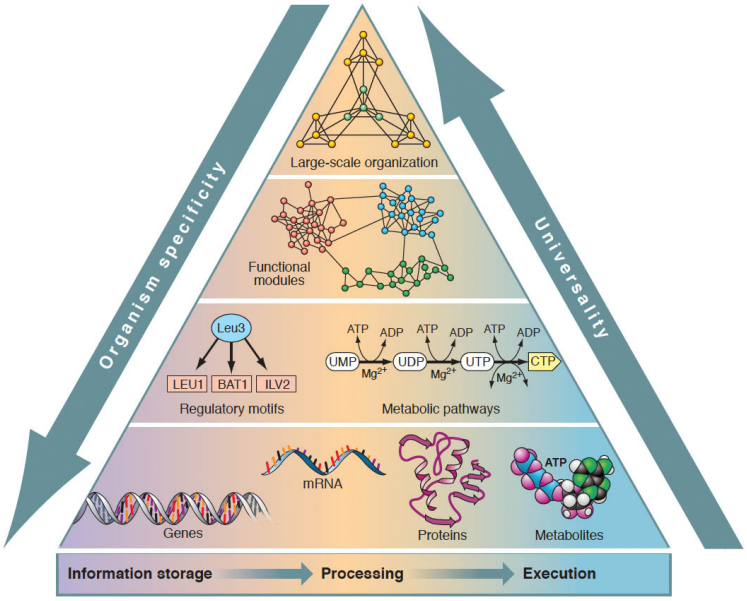

Figure 5.1: Life's complexity pyramid illustrating the different 'omics' layers as well as their integration to build up the large-scale organization of an organism and its phenotypic expression in a given environment (Oltvai and Barabási, 2002).

Challenges for phenotyping and data integration of molecular component traits

A first necessity was to ensure the high quality phenotyping of metabolites and lipids. Prior to this thesis research, research on metabolic profiles has mainly been performed under controlled

conditions and not with large-scale field trials. Thus, a sampling procedure had to be developed to obtain highly repeatable metabolic profiles in the field. The main challenges were to:

i shock freeze samples of 600 field plots (6,000 individual plants pooled into 600 sample bags) in a short period of time to minimize metabolic changes over time;

ii correct for potential time trends in the metabolic profile;

iii correct for environmental differences within the field due to *e.g.* soil differences;

iv randomize the measurement batches to catch potential systematic lab effects due to *e.g.* differences in intensity of extraction, derivatization, chromatographic separation;

v minimize loss in precision if one batch defrosts or gets destroyed;

vi obtain genotypic means in a single-step procedure, *e.g.* correct for all systematic effects simultaneously.

The solution we developed fulfils all these requirements (Fig. 5.2). The high mean reapeatabilities of 0.73 for metabolites and 0.66 for lipids show that highly repeatable metabolic and lipid profiles can be generated in large-scale trials of maize inbred lines. The obtained repeatabilites were significantly higher than those reported for recombinant inbred lines of *Arabidopsis* grown under controlled conditions (Chan *et al.*, 2010; Keurentjes *et al.*, 2006).

The stability of metabolites is however still an open research question. Because leaf samples were collected in one year and one location only, the repeatabilities might not necessarily reflect the heritabilities over a series of years and locations. In addition, it remains unclear how sensitive metabolic profiles are regarding the climatic conditions during the day of sampling and the days before.

Since metabolic profiling is still an expensive and resource-demanding task, the question arises how metabolic profiling data from different experiments can be compared and integrated. If possible, this would allow to perform GWAS meta-analyses as routinely done in human genetics. By expanding the population size in the combined analysis, power to detect QTL could be substantially increased.

A crucial first criteria to meet is a sufficiently high repeatability. If one dataset with high repeatability is combined with one having a low repeatability, the power to detect QTL might be decreased by diluting a high quality signal with random noise. As spatial metabolic analysis showed a clear separation of different organs on the metabolic level (Hanhineva *et al.*, 2008), an important decision concerns the organ from which the sample is taken and its developmental stage. Ensuring comparability in the environmental conditions appears to to be more difficult. Although quantification of the relative contributions of the genotype $\times$ environment interactions to the phenotypic variance is still missing, metabolomics studies with a case-control

set-up found clearly differentiable metabolic pattern when changing different growth conditions such as light (Lubbe *et al.*, 2012) or fungicide treatment (Hanhineva *et al.*, 2008). In maize, it has recently been shown that the common fungus *Ustilago maydis* which often infect plants in the field, produces a secretory protein which severely alters the metabolic status of the host plant (Djamei *et al.*, 2011). As it is clear that many more factors might influence the obtained metabolic profile, information about their relative importance seems to be crucial for assessing whether distinct data sets can be combined.

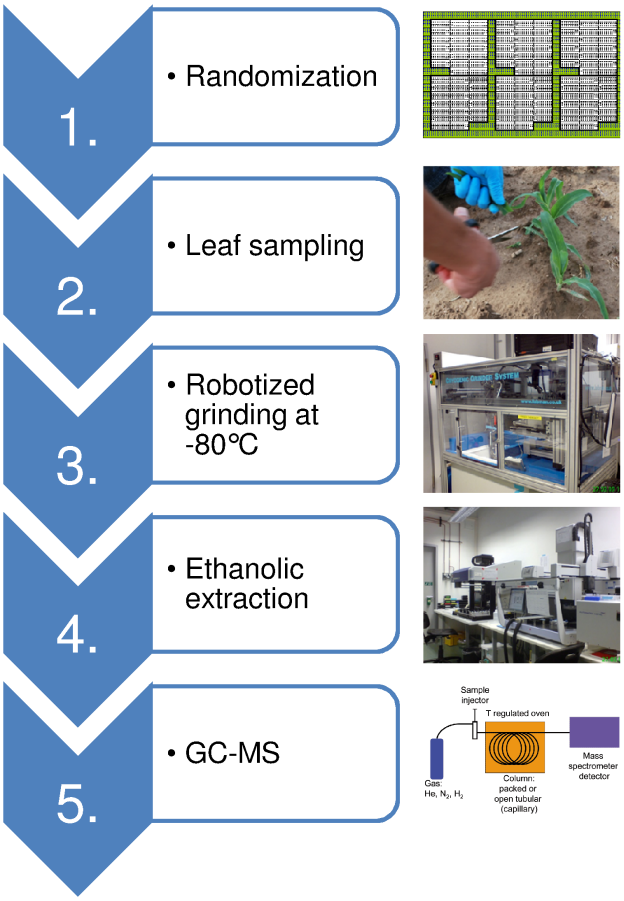

Figure 5.2: Steps involved until plant extracts for MS-based metabolic profiling were obtained from the field experiments. After randomizing the genotypes in three adjacent α-lattice designs according to their three maturity groups, ten plants per plot were pooled and immediately frozen using dry ice. Grinding of plant material was done with a robotized platform which keeps the plant samples at -80°C to prevent metabolic activities during grinding. Ethanolic extraction with pipetting robots was done in batches of 50 samples allocated on 96-well plates. To account for batch effects in the subsequent mixed model normalization, batches contained samples from randomly chosen field blocks of one field replication within one maturity group only. Randomization of batches was therefore nested in between complete field replications and field blocks which allowed a combined correction of both field and batch effects in one analysis step.

The genetic architecture of the maize leaf metabolome and lipidome

Using a false discovery rate (Storey and Tibshirani, 2003) of 2.5 %, GWAS detected significant associations for 26 metabolites and 76 lipids. For both metabolites and lipids, associations explaining more than 30 % of the genetic variance were detected. Thus, our results demonstrate that at least some molecular component traits are under a relatively simple genetic control with large-scale allelic effects detectable by GWAS. Metabolic QTL with allelic effects of this size have not been detected in previous studies in *Arabidopsis* using either LM (Lisec *et al.*, 2008) or GWAS (Chan *et al.*, 2010; Keurentjes *et al.*, 2006). However, studies in human urine (Suhre *et al.*, 2011) or blood (Illig *et al.*, 2010) suggest that associations of this size exist in nature. Thus, the inability to detect QTL with large effects in *Arabidopsis* despite large number of metabolites and SNPs is more likely attributable to insufficient populations size or too low repeatabilities.

Although several studies reported the agglomeration of SNP-metabolite associations (Fu *et al.*, 2009; Keurentjes *et al.*, 2006; Lisec *et al.*, 2008), we could not find any. Besides the small number of total associations found, an explanation for such hotspots might be the occurrence of biochemically connected or otherwise highly correlated metabolites leading to indirect associations with the same genetic region.

An important finding of the GWA mapping of the lipidome was the ability to establish a conceptual framework for assigning hierarchical levels at which the underlying candidate genes control the lipidome. The detected SNP-lipid associations could be grouped into associations with 1) individual lipids, 2) lipids from one biochemical class, and 3) lipids from several classes. This suggested a multilevel genetic control architecture similar as employed in engineering of complex controlled mechanical systems. In many cases, the assigned control levels matched with *a priori* knownledge about the specificity of candidate genes found in the vicinity of the QTL region. Whereas *e.g.* specific peroxidases were uniquely associated with single lipids, prominent lipid signaling genes such as sphingosine kinase were associated with large sets of lipids from multiple chemical classes.

Linking genetic variants for component traits back to agronomic traits traits

Our results show that GWAS results from individual molecular components, *e.g.*, monolignols can yield valuable information on the genetic control of correlated biomass-related agronomic traits in mature plants. Such information can be a valuable source for both (i) enriching the functional understanding how complex traits are controlled and (ii) identifying targets for knowledge-based breeding approaches aiming at improving a specific physiological process.

The identification of genetic loci underlying molecular component traits are however not likely to yield markers suitable for directly selecting the trait of interest. For example, the detected QTL on chr. 9 for caffeic acid explains $> 30\%$ of the genetic variance but less than 2 % of the correlated agronomic traits. This was however not unexpected from the high polygenic architecture of most agronomic traits.

The different signs of the correlations between caffeic acid and biochemical related metabolites with lignin content also indicate that the relationships between pathway intermediates and the final product is not simple. Directly modeling complex traits with individual component traits may therefore require consideration of feedback loops and other interdependencies. Structural equation modeling might be a promising technique for this task. This statistical method originally developed by the geneticist Sewall Wright (1921) tests causal relations using both experimental data and qualitative causal assumptions, the latter of which could be derived from known plant biochemistry. This approach has already been applied for deciphering complex interrelationships in various settings such as ecological modeling (Arhonditsis *et al.*, 2006) or in the pathophysiology of human diseases (Stevenson *et al.*, 2012). Knowledge about interdependencies of metabolites such as the lignin pathway intermediates are crucial for successful genetic engineering projects aiming to *e.g.* alter lignin composition without detrimental side effects on biotic or abiotic stress resistance (Weng *et al.*, 2008).

If the goal is merely to predict the phenotypic value of a genotype across a given set of environments without deciphering any causative gene-phenotype relationships, genomic selection (GS) seems to be the current tool of choice. The approach uses a training population to create a statistical model by assigning effects to all markers instead of only the significant ones detected in GWAS or LM studies (Lorenz *et al.*, 2011). Initially developed in the field of animal breeding (Meuwissen *et al.*, 2001), it is currently replacing traditional marker-assisted selection procedures in plant breeding (Heffner *et al.*, 2009). GS is anticipated to be highly successful for increasing selection gain in breeding practice (Riedelsheimer *et al.*, 2012a) and is already implemented in the private sector. GS is designed for and work best with a highly polygenic genetic architecture by sourcing its predictive information from a very precise estimation of genetic relatedness, *i.e.*, the deviation from the expected relationship due to Mendelian sampling (Hill and Weir, 2011).

Our results suggest that molecular component traits can be used in the same way as SNPs for GS by estimating effects for all predictor variables irrespective of their individual impact on the target trait. For both metabolites (Riedelsheimer *et al.*, 2012a) and lipids (Riedelsheimer *et al.*, 2013), prediction accuracies for general combining abilities evaluated in multi-environment field trials were found to be close to those obtained with SNPs (Riedelsheimer *et al.*, 2012a).

It is still an open question though, what extra value molecular component traits may carry on top of SNPs as their combination did not yield improved prediction accuracies. In contrast to SNPs, metabolic or lipid profiles are expected to carry information about the environment in which they were measured. Thus, if the goal is to predict the performance of genotypes in environment A but phenotypic data is only available for environment B (which shows a genotypic correlation < 1 with environment A), we speculate that a GS prediction model with

both SNPs and molecular profile data from the target environment A can be superior to a model with SNPs alone. Further research is, however, necessary to investigate this hypothesis.

GWAS conceptually complements GS in that it delivers insights about the genetic architecture of the trait, commonly described as the number of genes and the distribution of their effects affecting the trait as well as their interactions (Hill, 2012). Both approaches therefore deliver answers to different questions. Figure 5.3 illustrates the two different strategies for using metabolites either for dissecting or predicting complex traits. Prediction models with molecular component traits might also be further enriched with other sources of information like structural variation (copy number variants, CNVs) or knowledge from crop models or climate data about the conditions at a specific target environment.

Recently, Bayesian GS models were developed which precisely capture the individual QTL effects instead of merely modeling genome-wide relationship structure at the SNP effect level (de Los Campos *et al.*, 2012). It was suggested that such models could integrate GWAS and GS as one would be able to both identify the genetic regions controlling the trait and predicting it at the same time. Recently, however, Gianola (Gianola, 2013) demonstrated that SNP effects from such Bayesian models should be treated with caution if claims about genetic architecture are to be made. It was shown that in a situation with many more predictor variables (*i.e.* SNPs) than genotypes, the obtained predictor effects are always influenced by the chosen prior distribution because its parameters are never likelihood-identified. Alternatively, one might compare prediction accuracies obtained with models having different assumptions about the distribution of the underlying genetic effects. The model which yields the highest accuracies is then assumed to be the one whose assumptions match best. However, differences between GS models were in general found to be only very small (Heslot *et al.*, 2012) and also prone to a large sampling variance. Riedelsheimer *et al.* (2012b) found models assuming a non-normal distribution of genetic effects slightly superior for predicting metabolic traits with strong GWAS signals explaining > 30 % of genetic variances. However the differences between the models were only small which limits the use of such model comparisons for drawing inferences about the genetic architecture of the trait of interest. GS models therefore provide only limited biological insights into how the genetic make-up results in the variation of a trait.

Linking genetic variants for physiological traits back to chilling sensitivity

As outlined in the general introduction, chilling sensitivity is an economically important trait for which only little is known about its underlying genetic control. Because exposure to chilling temperatures is expressed in an array of physiological responses in the plant, we aimed to map the genetic factors underlying physiological component traits correlated with chilling sensitivity. The detected QTL explained up to 12 % of the phenotypic variance for the measured parameters, thus, demonstrating the feasibility of the approach.

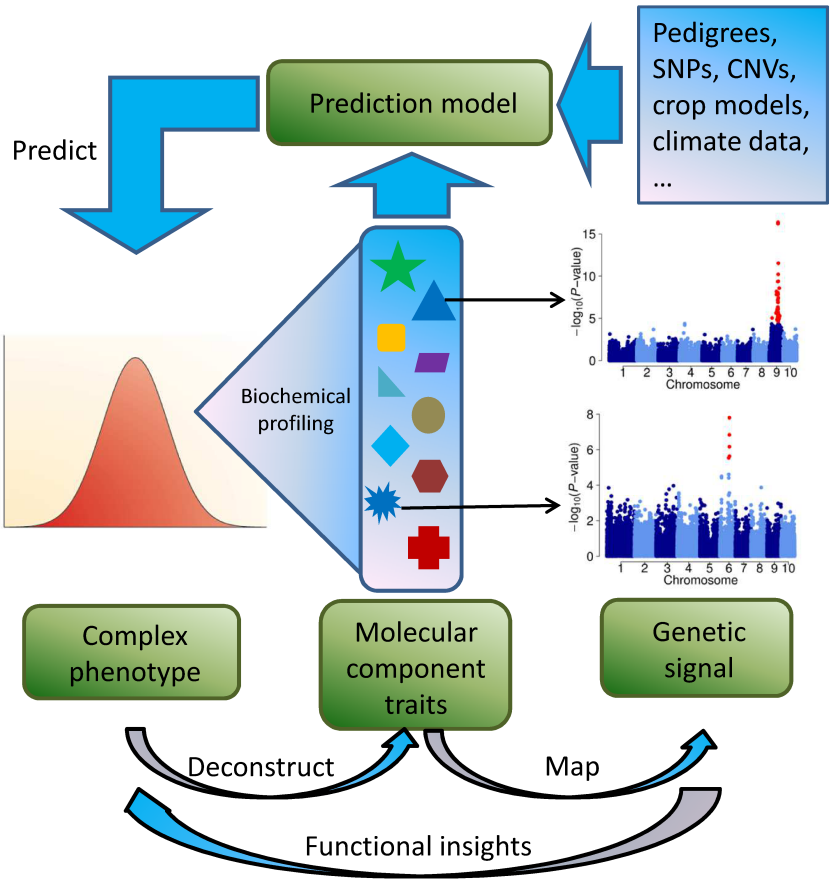

Figure 5.3: Strategies for using molecular component traits for dissecting and predicting complex traits. Complex traits are deconstructed with *e.g.* mass-spectroscopy based profiling techniques into genetically simpler regulated molecular component traits like metabolites or lipids. Genetic analysis of them can reveal genetic signals for shedding light on the genetic regulation of the (correlated) complex trait. As a complementary approach, the full amount of molecular profile data might be used in a prediction model for predicting the complex trait. The statistical prediction model might be adopted from fields such as machine learning or genomic selection and might further integrate several layers of information including pedigrees, SNPs, copy number variants (CNVs) or knowledge from crop models or climate data about the target environment.

However, relative growth rate in the field were in many cases not or only for specific heterotic groups significantly correlated with the physiological parameters measured under controlled conditions in growth chambers. Thus, drawing inferences from the genetic variants for those parameters to early growth in the field remains highly speculative. On the other hand, early growth in the field showed as expected a high level of genotype × environment interactions resulting in QTL which are highly specific for a certain environment only and, thus, are not suitable for marker assisted selection either.

For improving selection gain, we speculate that GS is the way to go here, too. For drawing

inferences about biological processes underlying chilling sensitivity, the space of inferences from the results of GWAS is limited to a single physiological process (such as photosynthetic performance) in a single environment unless genotype × environment interactions can be successfully controlled and stable relationships to chilling sensitivity in the field can be established.

General limitations of GWAS of component traits

Confounding with structure

The confounding of associations with population structure and cryptic relatedness is a well known problem in GWAS (Astle and Balding, 2009). Yet, it appeared to be a smaller problem for molecular component traits than for *e.g.* flowering time, a prominent example trait which correlates strongly with population structure (Van Inghelandt *et al.*, 2012). Meanwhile, it has even been shown that the prominent flowering time locus *dwarf8* which was published in *Nature Genetics* in 2001 (Thornsberry *et al.*, 2001) is most likely an artefact due to insufficient control of population structure. This has become evident after reanalysis of the results from Thornsberry *et al.* (2001) by the same authors of the original study (Larsson *et al.*, 2013).

To overcome the problem of confounding with population structure in maize, the maize community has created the nested association mapping (NAM) population by crossing 25 diverse lines to a common parent (B73) and developing ≈ 200 recombinant inbred lines from each cross (McMullen *et al.*, 2009). The NAM panel allows to perform joint-linkage association mapping. This approach combines the advantages of LM (high power) with GWAS (high precision) while breaking confounding with population structure by having a common parent for all crosses. Meanwhile, the NAM population has been very effective in dissecting complex agronomic traits although no major QTL were found for different traits ranging from kernel composition (Cook *et al.*, 2011) and leaf architecture (Tian *et al.*, 2011) to various disease resistances (Kump *et al.*, 2011; Poland *et al.*, 2011).

Epistasis

Classical GWAS with inbred lines as performed in this thesis research provides only estimates of additive effects. Unfortunately, power to detect moderate sized two-way interactions is low and extremely low for higher-order interactions. With increasing tests to be made, also the significance threshold for controling the experiment-wise error becomes more stringent. In addition, the number of tests to be made for all all two-way interactions rise for the 56k SNP chip from 5.6×10^4 to 3.1×10^9 if all two-way interactions should be tested. This produces enormous practical challenges requiring the development of new algorithms for efficient testing.

Given the natural complexity of biological pathways and gene regulatory networks, epistasis

seems to be prevalent in nature. Yet, the relative importance of epistatic interaction effects play a role is a question which still has to be resolved. For maize, the literature gives contradicting answers to this question. For example, whereas Durand *et. al.* (2012) found a strong contribution of epistasis in a major locus for flowering time, Mihaljevic *et al.* (2005) did not find evidence for epistasis using generation mean and LM analysis for grain yield and grain moisture.

In natural or diverse populations, empirical evidence points towards mainly additive genetic variance (Hill *et al.*, 2008). For example, Riedelsheimer *et al.* (2012b) found that fitting all 56k SNPs simultaneously in an additive regression model accounts for most of the heritability of dry matter yield, plant height, and lignin content in a diversity panel of maize inbred lines. However, Hill (2012) recently pointed out that such findings do *not* imply that gene action is additive. Under the assumption of mutation-drift balance, the frequency density of alleles is proportional to $1/[p(1-p)]$, *i.e.*, U-shaped where p is the allele frequency. Thus, if at most loci one genotype is likely to be very infrequent in the diversity panel subject to GWAS, epistasis can contribute only little to the total phenotypic variance. This imbalance in allele frequencies is no issue in biparental populations. Therefore, it has been suggested to favor classical LM as the more appropriate tool for mapping epistatic interactions (Rafalski, 2010).

Rare variants

Empirical studies suggest that most alleles are rare in maize. For example, 30 % of the polymorphisms between 27 diverse maize inbred lines were found to be unique to a single line (Myles *et al.*, 2009). Unfortunately, detecting rare variants in GWAS is difficult unless their effect is very large (Rafalski, 2010).

For a given number of individuals and level of environmental variation, power to detect a QTL with additive effect a on the trait is proportional to r^2a^2 where r^2 is the squared LD correlation between QTL and marker alleles (Hill, 2012). If a QTL is rare ($p < 5$ %) any SNP within this QTL has been discarded because only markers with a minor allele frequency (MAF) > 5 % were used for GWA mapping in our studies to decrease the number of false positives. In addition, the used marker panel was designed with the idea in mind that most practitioners should be able to work with a high proportion of polymorphic SNPs. It is likely that most rare SNPs were not represented in the used SNP panel. Thus, low frequent QTL would have been only found if in LD with a nearby SNP showing a MAF > 5 %. However, only for 48.1 % of the SNPs, there exists at least one SNP in strong LD ($r^2 > 0.8$) in the vicinity. In addition, the higher the MAF of an SNP, the lower the maximum LD with a rare variant in the vicinity.

Recently, Mackay *et al.* (2012) found an inverse relationship between allele frequency and allelic effects detected with GWA mapping in *Drosophila*, *i.e.* less frequent variants showed stronger allelic effects making them easier to detect even with lower r^2 with genotyped SNPs. However, Marjoram *et al.* (2014) pointed out that this finding is probably due to the fact that the expected overestimation of effect sizes increases with decreasing allele frequency (Lynch

and Walsh, 1998).

Increasing r^2 between genotyped SNPs and rare variants could be achieved by more dense genotyping platforms with a wider allele frequency distribution, especially around the MAF cutoff used in GWAS. A currently available option to accomplish this in maize would be genotyping-by-sequencing (Elshire *et al.*, 2011). However, even if r^2 would be high, the small number of genotypes with the causal rare variant limit the power for testing the significance of phenotypic differences between the allele classes using a standard t-test. In order to increase the power of detecting rare variants in humans, it has been suggested to not only increase the population size but also to use 'collapsing strategies' which test the combined effect of multiple rare variants (Bansal *et al.*, 2010). Further research is necessary to investigate whether this approach is also feasible in maize.

In conclusion, as pointed out by Rafalski (2010), classical LM with biparental populations derived from contrasting inbred lines remains the easier method for mapping rare alleles such as disease resistances or alleles introgressed from exotic germplasm. In the aforementioned NAM population, variants unique to one of the founder lines can also be mapped within the individual segregating population created by crossing the line to the common parent line B73.

Validation of QTL and candidate genes

Because GWAS is a *data-driven* approach which yields only statistical *i.e.* indirect evidence for the association of a genomic region with the target trait, validation of the detected QTL are necessary. Such validations can be (i) confirmative in different germplasm, or (ii) functional. Failure of confirmative validation of detected marker-trait are often observed. A prominent example is the aforementioned *dwarf8* locus for flowering time (Andersen *et al.*, 2005; Larsson *et al.*, 2013). Because associations are called significant if below a certain arbitrary significance threshold, lack of QTL confirmation might be due to spurious associations because of a two weakly chosen P-value threshold. Another reason might be insufficient control of population structure or LD between unlinked loci. In addition, failure of confirmative validation might be because of insufficient power in the validation population, QTL $\times$ environment interactions, or QTL $\times$ genetic background interactions. The latter might only be alleviated by testing the QTL allele effect in nearly isogenic backgrounds by marker-assisted introgression or genetic engineering approaches.

In companion to confirmative validations, functional validations need to be conducted if the biological basis underlying a QTL is of interest. With the availability of searchable genome databases, candidate genes are searched with a preconceived idea about the underlying mechanism which might, however, not be true. Thus, GWAS takes biological knowledge into account only *a posteriori* (Marjoram *et al.*, 2014). In addition, the SNPs on the used MaizeSNP50 BeadChip were selected to lie within gene-rich genomic regions (Ganal *et al.*, 2011). As a consequence, tens to hundreds of coded genes exist (or in most cases: are predicted to exist using bioinformatic algorithms) within the region around the most significant SNP in which

LD is sufficiently high.

Possible approaches for a functional validation include antisense methods or the production of knock-out mutants for inducing loss-of-function point mutations in the candidate genes. For metabolic QTL, a first approach would be measure the transcription level of the candidate genes. Because the levels of many metabolites are assumed to be transcriptionally regulated, it is likely that the transcription rate of the causative gene shows the same association as the metabolite. This would be an additional line of empirical evidence pointing to the same candidate gene.

Conclusions

The results of this thesis research demonstrate that the genotype-phenotype gap of highly polygenic biomass-related traits can be successfully narrowed by the genetic analysis of genetically simpler component traits from metabolic and lipid profiles. High levels of genotype × environment interactions for chilling sensitivity were however found to limit the ability to use physiological component traits to draw inferences about its genetic control. Our specific conclusions are:

i In contrast to previous experiments in *Arabidopsis*, the developed sampling and randomization procedure allows to generate highly repeatable metabolic and lipid profiles from a diversity panel of maize inbred lines grown under field conditions.

ii The maize leaf metabolome and lipidome is intensively structured with a correlation pattern reflecting their functional grouping.

iii GWAS with 56k SNPs is able to unravel the genetic architecture of the maize leaf metabolome and lipidome. At least some metabolites and lipids have a very simple genetic architecture with individual SNPs explaining more than 30 % of genetic variance in GWAS.

iv The lipidome shows a multilevel genetic control architecture similar as used in engineering complex regulated mechanical systems.

v Functional, biological connections between genetic variants of molecular component traits and agronomic traits are possible in some instances. Examples are enzymes in the monolignol pathways detected for lignin precursors, which correlated with dry matter yield, plant height, and dry matter yield. Interrelations and feedback loops seems to be present in many instances.

vi Whereas GWAS of molecular component traits provides information about functional relationships, it does not provide allelic effects of SNPs usable in marker-assisted selection

programs for improving the polygenic agronomic traits. If the goal is merely to predict the phenotypic value, the application of black-box genomic selection methodology with either SNPs, molecular profiles, or both combined are promising tools to achieve this goal.

vii QTL for physiological component traits of chilling sensitivity under controlled conditions can be detected and plausible candidate genes assigned. However, connecting genetic variants underlying these physiological components with early growth in the field is hampered by insignificant correlations and large amount of genotype × environment interactions. The complex and environment-dependent response of plants after exposure to chilling temperatures challenges the genetic dissection and modeling of chilling sensitivity with physiological component traits.

6 Summary

Genome-wide association (GWA) mapping emerged as a powerful tool to dissect complex traits in maize. Yet, most agronomic traits were found to be highly polygenic and the detected associations explained together only a small portion of the total genetic variance. Hence, the majority of genetic factors underlying many agronomically important traits are still unknown. New approaches are needed for unravelling the chain from the genes to the phenotype which is still largely unresolved for most quantitative traits in maize.

Instead of further enlarging the mapping population to increase the power to detect even smaller QTL, this thesis research aims to present an alternative route by mapping not the polygenic trait of primary interest itself, but genetically correlated molecular and physiological component traits. As such components represent biological sub-processes underlying the trait of interest, they are supposed to be genetically less complex and thus, more suitable for genetic mapping. Using large diversity panels of maize inbred lines, this approach is demonstrated with (i) biomass yield by using metabolites and lipids as molecular component traits and with (ii) chilling sensitivity by using physiological component traits such as photosynthesis parameters derived from chlorophyll fluorescence measurements.

In a first step, we developed a sampling and randomization scheme which allowed us to obtain metabolic and lipid profiles from large-scale field trials. Both profiles were found to be intensively structured reflecting their functional grouping. They also showed repeatabilities higher than in comparable profiles obtained in previous studies with the model plant *Arabidopsis* under controlled conditions.

By applying GWAS with 56,110 SNPs to metabolites and lipids, large-scale genetic associations explaining more than 30 % of the genetic variance were detected. Confounding with structure was found to be a problem of less extent for molecular components than for agronomic traits like flowering time. The lipidome was also found to show a multilevel control architecture similar as employed in controlling complex mechanical systems. In several instances, direct links between candidate genes underlying the detected associations and agronomic traits could be established. An example is cinnamoyl-CoA reductase, a key enzyme in the lingin biosynthesis pathway. It was found to be a candidate gene underlying a major QTL found for several intermediates in the lignin biosynthesis pathways. These intemediates were in turn found to be correlated with plant height, lignin content, and dry matter yield at the end of the vegetation period. The different signs of these correlations indicated that the relationships between pathway intermediates and the final product is not simple. Directly modeling complex traits with individual component traits may therefore require consideration of feedback loops and other interdependencies.

Such connections were however found difficult to be established with physiological components underlying chilling sensitivity. The main reasons for this were the weak correlations between physiological components under controlled conditions and chilling sensitivity in the field as well as high levels of genotype × environment interactions caused by the complex and environment-dependent responses of maize after perception of chilling temperatures.

The approach explored in this thesis research uses component traits to gain biological insights about the genetic control of biomass yield and chilling sensitivity evaluated in diverse populations of still manageable sizes. We showed that GWAS with 56k SNPs can identify large additive effects for component traits correlated with these traits. For mapping epistatic interactions and rare variants, classical linkage mapping with biparental populations will be a reasonable complementary approach. However, controlling and modeling genotype × environment interactions remains an important issue for understanding the genetic basis of especially chilling sensitivity. If the goal is merely to predict the phenotypic value in a given set of environments, black-box genomic selection methods with either SNPs, molecular profiles, or a combination of both, are very promising strategies to achieve this goal.

7 Zusammenfassung

Die genomweite Assoziationskartierung (GWA) hat sich als hilfreiches Werkzeug zur genetischen Analyse komplexer Merkmale in Mais erwiesen. Die meisten Merkmale von agronomischer Bedeutung haben sich allerdings als hochgradig polygen herausgestellt, mit sehr vielen genetischen Loci, die in der Summe nur einen Bruchteil der gesamten genetischen Varianz erklären. Folglich sind die meisten genetischen Faktoren, welche wichtige agronomische Merkmale kontrollieren, immer noch unbekannt. Es besteht daher ein großes Interesse, Alternativen für die Aufklärung der kausalen Kette zwischen den Genen einerseits und der phänotypischen Ausprägung komplexer Merkmale andererseits zu finden.

In dieser Arbeit wird ein solcher alternativer Ansatz untersucht. Anstatt die komplexen agronomischen Merkmale direkt zu kartieren, verfolgt diese Arbeit das Ziel, deren korrelierte molekulare und physiologische Komponenten zu untersuchen. Da solche Komponenten meist klar interpretierbare biologische Unterprozesse repräsentieren, sind sie höchstwahrscheinlich genetisch einfacher kontrolliert und daher für GWAS besser als die meisten agronomischen Merkmale geeignet. In dieser Arbeit wird dieser Ansatz mit Hilfer großer diverser Maispopulationen an zwei Beispielen untersucht: (i) Biomasseertrag mit Metabolit- und Lipidprofilen als molekulare Komponentenmerkmale, und (ii) Kältetoleranz mit physiologischen Komponentenmerkmalen wie z.B. photosynthetische Leistungsparametern, die von Chlorophyllfluoreszenz-Messungen abgeleitet wurden.

Als erster Schritt wurde eine Probenahme- und Randomisationsstruktur entworfen, die es ermöglichte, Metabolit- und Lipidprofile von mehreren Hundert Maislinien unter Feldbedinungen zu generieren. Die Wiederholbarkeiten der molekularen Komponenten überstiegen dabei deutlich jene, welche für vergleichbarer Profile in *Arabidopsis* unter kontrollierten Bedingungen ermittelt wurden. Die Profile zeigten eine innere Korrelatiosstruktur, welche die funktionelle Gruppierung der Komponenten widerspiegelt.

Bei den genetischen Analysen der Metabolit- und Lipidprofile mittles GWAS mit 56.110 SNPs wurden Assozationen mit Genorten gefunden, welche mehr als 30 % der gesamten genetischen Varianz erklären. Der Einfluss der Populationsstruktur war dabei geringer als bei vielen agronomischen Merkmalen, wie beispielsweise der Blühzeitpunkt. Für das Lipidom wurde außerdem eine genetische Mehrebenen-Kontrollarchitetur gefunden, welche Ähnlichkeit mit Steuereinrichtungen für komplexe mechanische Systeme besitzt. Direkte Verknüpfungen zwischen Kandidatengenen der detektierten Assoziationen und den agronomischen Merkmalen konnten in mehreren Fällen festgestellet werden. Bespielsweise wurde für mehrere Lignin-vorstufen die gleiche genetische Assoziation mit Cinnamoyl-CoA Reductase gefunden – einem

Schlüsselenzym der Ligninbiosynthese. Im Umkehrschluss korrelierten diese Ligninvorstufen signifikant mit Pflanzenhöhe, Lingingehalt und Trockenmasseertrag am Ende der Vegetationperiode. Die unterschiedlichen Vorzeichen dieser Korrelationen wiesen allerdings darauf hin, dass die Verbindung zwischen Stoffwechselprodukten und Endprodukt komplex ist. Die Modellierung von komplexen Merkmalen mit molekularen Komponenten erfordert daher die Einbeziehung von gegenseitgen Abhängigkeiten und Rückkopplungsschleifen.

Solche Brücken zwischen einzelnen Genen und phänotypischen Merkmalen konnten allerdings nur bedingt für Kältetoleranz etabliert werden. Die Gründe lagen hierfür sowohl in den schwachen Korrelationen zwischen physiologischen Komponten unter kontrollierten Bedingungen und Kältetoleranz im Feld, als auch in den bedeutenden Genotyp $\times$ Umwelt Interaktionen, welche durch die komplexen und stark umweltabhängigen Reaktionen von Mais nach Einwirkung kühler Temperaturen hervorgerufen wurden.

Der in dieser Arbeit verfolgte Ansatz benützt Komponentenmerkmale, um Erkenntnisse über die genetische Regulierung komplexer Merkmale in diversen Maispopulationen von handhabbarer Größe zu erlangen. Unsere Ergebnisse zeigen, dass mittles 56 Tausend SNPs für die Komponentenmerkmale Assoziationen mit starken additiven Effekten gefunden werden können. Zur Kartierung von epistatischen Interaktionen oder niedrig-frequenten (rare) Varianten kann die klassische Kopplungsanalyse mit biparentalen Populationen ein hilfreicher komplementärer Ansatz sein. Ist das Ziel allerdings aussschließlich die Vorhersage der phänotypischen Leistung unter gegebenen Umweltbedingungen, bieten Methoden der genomischen Selektion mit SNPs, molekularen Profilen, oder beiden Biomarkern kombinert, einen erfolgsversprechenden komplementären Ansatz.

References

Andersen, J. R., T. Schrag, A. E. Melchinger, I. Zein, and T. Lübberstedt, 2005. Validation of Dwarf8 polymorphisms associated with flowering time in elite European inbred lines of maize (Zea mays L.). *Theor. Appl. Genet.* 111:206–217.

Arhonditsis, G., C. Stow, L. Steinberg, M. Kenney, R. Lathrop, S. McBride, and K. Reckhow, 2006. Exploring ecological patterns with structural equation modeling and Bayesian analysis. *Ecol. Model.* 192:385–409.

Astle, W. and D. J. Balding, 2009. Population Structure and Cryptic Relatedness in Genetic Association Studies. *Stat. Sci.* 24:451–471.

Bansal, V., O. Libiger, A. Torkamani, and N. J. Schork, 2010. Statistical analysis strategies for association studies involving rare variants. *Nat. Rev. Genet.* 11(11):773–85.

Bernardo, R., 2008. Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years. *Crop Sci.* 48:1649–1664.

Broun, P., 1998. Catalytic Plasticity of Fatty Acid Modification Enzymes Underlying Chemical Diversity of Plant Lipids. *Science* 282:1315–1317.

Brown, H. A. and R. C. Murphy, 2009. Working towards an exegesis for lipids in biology. *Nat. Chem. Biol.* 5:602–606.

Buckler, E., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J. C. Glaubitz, M. M. Goodman, C. Harjes, K. Guill, D. E. Kroon, S. Larsson, N. K. Lepak, H. Li, S. E. Mitchell, G. Pressoir, J. a. Peiffer, M. O. Rosas, T. R. Rocheford, M. C. Romay, S. Romero, S. Salvo, H. Sanchez Villeda, H. S. da Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich, and M. D. McMullen, 2009. The genetic architecture of maize flowering time. *Science* 325:714–718.

Chan, E. K. F., H. C. Rowe, B. G. Hansen, and D. J. Kliebenstein, 2010. The Complex Genetic Architecture of the Metabolome. *PLoS Genet.* 6:e1001198.

Cook, J. P., M. D. McMullen, J. B. Holland, F. Tian, P. Bradbury, J. Ross-Ibarra, E. S. Buckler, and S. a. Flint-Garcia, 2011. Genetic Architecture of Maize Kernel Composition in the Nested Association Mapping and Inbred Association Panels. *Plant Physiol.* 158:824–834.

de Los Campos, G., J. M. Hickey, R. Pong-Wong, H. D. Daetwyler, and M. P. L. Calus, 2012. Whole Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. *Genetics* 193:327–345.

De Vos, R. C. H., S. Moco, A. Lommen, J. J. B. Keurentjes, R. J. Bino, and R. D. Hall, 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Prot.* 2:778–791.

Djamei, A., K. Schipper, F. Rabe, A. Ghosh, V. Vincon, J. Kahnt, S. Osorio, T. Tohge, A. R. Fernie, I. Feussner, K. Feussner, P. Meinicke, Y.-D. Stierhof, H. Schwarz, B. Macek, M. Mann, and R. Kahmann, 2011. Metabolic priming by a secreted fungal effector. *Nature* 478:395–398.

Durand, E., S. Bouchet, P. Bertin, A. Ressayre, P. Jamin, A. Charcosset, C. Dillmann, and M. I. Tenaillon, 2012. Flowering time in Maize: Linkage and Epistasis at aMajor Effect Locus. *Genetics* 190:1547–1562.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. a. Poland, K. Kawamoto, E. Buckler, and S. E. Mitchell, 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* 6:e19379.

Fracheboud, Y., 2002. Identification of quantitative trait loci for cold-tolerance of photosynthesis in maize (Zea mays L.). *J. Exp. Bot.* 53:1967–1977.

Frei, O. M., 2000. Changes in yield physiology of corn as a result of breeding in northern Europe. *Maydica* 45:173–183.

Fu, J., J. J. B. Keurentjes, H. Bouwmeester, T. America, F. W. a. Verstappen, J. L. Ward, M. H. Beale, R. C. H. de Vos, M. Dijkstra, R. a. Scheltema, F. Johannes, M. Koornneef, D. Vreugdenhil, R. Breitling, and R. C. Jansen, 2009. System-wide molecular evidence for phenotypic buffering in Arabidopsis. *Nat. Genet.* 41:166–167.

Ganal, M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler, A. Charcosset, J. D. Clarke, E.-M. Graner, M. Hansen, J. Joets, M.-C. Le Paslier, M. D. McMullen, P. Montalent, M. Rose, C.-C. Schön, Q. Sun, H. Walter, O. C. Martin, and M. Falque, 2011. A Large Maize (Zea mays L.) SNP Genotyping Array: Development and Germplasm Genotyping, and Genetic Mapping to Compare with the B73 Reference Genome. *PLoS ONE* 6:e28334.

Gianola, D., 2013. Priors in whole-genome regression: the Bayesian alphabet returns. *Genetics* 194:573–596.

Guy, C., F. Kaplan, J. Kopka, J. Selbig, and D. K. Hincha, 2008. Metabolomics of temperature stress. *Physiol. Plant.* 132:220–235.

Hanhineva, K., I. Rogachev, H. Kokko, S. Mintz-Oron, I. Venger, S. Kärenlampi, and A. Aharoni, 2008. Non-targeted analysis of spatial metabolite composition in strawberry (Fragaria-axananassa) flowers. *Phytochemistry* 69:2463–2481.

Heffner, E., M. Sorrells, and J. L. Jannink, 2009. Genomic Selection for Crop Improvement. *Crop Sci.* 49:1–12.

Heslot, N., H.-P. Yang, M. E. Sorrells, and J.-L. Jannink, 2012. Genomic Selection in Plant Breeding: A Comparison of Models. *Crop Sci.* 52:146–160.

Hill, W. G., 2012. Quantitative genetics in the genomics era. *Curr. Genomics* 13:196–206.

Hill, W. G., M. E. Goddard, and P. M. Visscher, 2008. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* 4:e1000008.

Hill, W. G. and B. S. Weir, 2011. Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet. Res. (Camb.)* 93:47–64.

Hund, A., E. Frascaroli, J. Leipner, C. Jompuk, P. Stamp, and Y. Fracheboud, 2005. Cold Tolerance of the Photosynthetic Apparatus: Pleiotropic Relationship between Photosynthetic Performance and Specific Leaf Area of Maize Seedlings. *Mol. Breeding* 16:321–331.

Illig, T., C. Gieger, G. Zhai, W. Römisch-Margl, R. Wang-Sattler, C. Prehn, E. Altmaier, G. Kastenmüller, B. S. Kato, H.-W. Mewes, T. Meitinger, M. H. de Angelis, F. Kronenberg, N. Soranzo, H.-E. Wichmann, T. D. Spector, J. Adamski, and K. Suhre, 2010. A genome-wide perspective of genetic variation in human metabolism. *Nat. Genet.* 42:137–141.

Keurentjes, J. J. B., 2009. Genetical metabolomics: closing in on phenotypes. *Curr. Opin. Plant Biol.* 12:223–230.

Keurentjes, J. J. B., J. Fu, C. H. R. D. Vos, A. Lommen, R. D. Hall, R. J. Bino, L. H. W. V. D. Plas, R. C. Jansen, D. Vreugdenhil, and M. Koornneef, 2006. The genetics of plant metabolism. *Nat. Genet.* 38:842–849.

Kooke, R. and J. J. B. Keurentjes, 2012. Multi-dimensional regulation of metabolic networks shaping plant development and performance. *J. Exp. Bot.* 63:3353–3365.

Kump, K., P. Bradbury, R. Wisser, E. Buckler, A. Belcher, M. Oropeza-Rosas, J. Zwonitzer, S. Kresovich, M. McMullen, D. Ware, and Others, 2011. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* 43:163–168.

Larsson, S. J., A. E. Lipka, and E. S. Buckler, 2013. Lessons from Dwarf8 on the strengths and weaknesses of structured association mapping. *PLoS Genet.* 9:e1003246.

Last, R. L., a. D. Jones, and Y. Shachar-Hill, 2007. Towards the plant metabolome and beyond. *Nat. Rev. Mol. Cell Bio.* 8:167–174.

Lee, E., M. Staebler, and M. Tollenaar, 2002. Genetic variation in physiological discriminators for cold tolerance—early autotrophic phase of maize development. *Crop Sci.* 42:1919–1929.

Lisec, J., R. C. Meyer, M. Steinfath, H. Redestig, M. Becher, H. Witucka-Wall, O. Fiehn, O. Törjék, J. Selbig, T. Altmann, and L. Willmitzer, 2008. Identification of metabolic and biomass QTL in Arabidopsis thaliana in a parallel analysis of RIL and IL populations. *Plant J.* 53:960–972.

Lisec, J., N. Schauer, J. Kopka, L. Willmitzer, and A. R. Fernie, 2006. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Prot.* 1:387–396.

Lorenz, A. J., S. Chao, F. G. Asoro, E. L. Heffner, T. Hayashi, H. Iwata, K. P. Smith, M. E. Sorrells, and J.-l. Jannink, 2011. Genomic Selection in Plant Breeding: Knowledge and Prospects. *In* Advances in Agronomy, vol. 110, Elsevier Inc., 1st edn., 77–123.

Lubbe, A., R. Verpoorte, and Y. H. Choi, 2012. Effects of fungicides on galanthamine and metabolite profiles in Narcissus bulbs. *Plant Physiol. Biochem.* 58:116–123.

Lynch, M. and B. Walsh, 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.

Mackay, T., E. Stone, and J. Ayroles, 2009. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10:565–577.

Mackay, T. F. C., S. Richards, E. a. Stone, A. Barbadilla, J. F. Ayroles, D. Zhu, S. Casillas, Y. Han, M. M. Magwire, J. M. Cridland, M. F. Richardson, R. R. H. Anholt, M. Barrón, C. Bess, K. P. Blankenburg, M. A. Carbone, D. Castellano, L. Chaboub, L. Duncan, Z. Harris, M. Javaid, J. C. Jayaseelan, S. N. Jhangiani, K. W. Jordan, F. Lara, F. Lawrence, S. L. Lee, P. Librado, R. S. Linheiro, R. F. Lyman, A. J. Mackey, M. Munidasa, D. M. Muzny, L. Nazareth, I. Newsham, L. Perales, L.-L. Pu, C. Qu, M. Ràmia, J. G. Reid, S. M. Rollmann, J. Rozas, N. Saada, L. Turlapati, K. C. Worley, Y.-Q. Wu, A. Yamamoto, Y. Zhu, C. M. Bergman, K. R. Thornton, D. Mittelman, and R. a. Gibbs, 2012. The Drosophila melanogaster Genetic Reference Panel. *Nature* 482:173–178.

Marjoram, P., A. Zubair, and S. V. Nuzhdin, 2014. Post-GWAS: where next? More samples, more SNPs or more biology? *Heredity* 112:79–88.

Mauricio, R., 2001. Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nat. Rev. Genet.* 2:370–381.

McMullen, M., S. Kresovich, H. Villeda, P. Bradbury, H. Li, Q. Sun, S. Flint-Garcia, J. Thornsberry, C. Acharya, C. Bottoms, and Others, 2009. Genetic Properties of the Maize Nested Association Mapping Population. *Science* 325:737.

Melchinger, A., H. F. Utz, and C. C. Schön, 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.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard, 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829.

Meyer, R., M. Steinfath, J. Lisec, M. Becher, H. Witucka-Wall, O. Törjék, O. Fiehn, b. Eckardt, L. Willmitzer, J. Selbig, and Others, 2007. The metabolic signature related to high plant growth rate in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U.S.A.* 104:4759–4764.

Mihaljevic, R., 2005. No Evidence for Epistasis in Hybrid and Per Se Performance of Elite European Flint Maize Inbreds from Generation Means and QTL Analyses. *Crop Sci.* 45:2605–2613.

Mutch, D. M., L. Fauconnot, M. Grigorov, and L. B. Fay, 2006. Putting the 'Ome' in lipid metabolism. *Biotechnol. Annu. Rev.* 12:67–84.

Myles, S., J. Peiffer, P. J. Brown, E. S. Ersoz, Z. Zhang, D. E. Costich, and E. S. Buckler, 2009. Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21:2194–2202.

Oltvai, Z. and A.-l. Barabási, 2002. Life's complexity pyramid. *Science* 298:763–764.

Ortiz-Lopez, A., G. Nie, D. Ort, and N. Baker, 1990. The involvement of the photoinhibition of photosystem II and impaired membrane energization in the reduced quantum yield of carbon assimilation in chilled maize. *Planta* 181:78–84.

Poland, J. A., P. J. Bradbury, E. Buckler, and R. J. Nelson, 2011. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc. Natl. Acad. Sci. U.S.A.* 108:6893–6898.

Presterl, T., M. Ouzunova, W. Schmidt, E. M. Möller, F. K. Röber, C. Knaak, K. Ernst, P. Westhoff, and H. H. Geiger, 2007. Quantitative trait loci for early plant vigour of maize grown in chilly environments. *Theor. Appl. Genet.* 114:1059–1070.

Rafalski, J. A., 2010. Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13:174–180.

Riedelsheimer, C., Y. Brotman, M. Méret, A. E. Melchinger, and L. Willmitzer, 2013. The maize leaf lipidome shows multilevel genetic control and high predictive value for agronomic traits. *Sci. Rep.* 3:2479.

Riedelsheimer, C., A. Czedik-Eysenberg, C. Grieder, J. Lisec, F. Technow, R. Sulpice, T. Altmann, M. Stitt, L. Willmitzer, and A. E. Melchinger, 2012a. Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nat. Genet.* 44:217–220.

Riedelsheimer, C., F. Technow, and A. Melchinger, 2012b. Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. *BMC Genomics* 13:452.

Rodríguez, V. M., A. Butrón, R. a. Malvar, A. Ordás, and P. Revilla, 2008. Quantitative Trait Loci for Cold Tolerance in the Maize IBM Population. *Int. J. Plant Sci.* 169:551–556.

Saito, K. and F. Matsuda, 2010. Metabolomics for functional genomics, systems biology, and biotechnology. *Annu. Rev. Plant Biol.* 61:463–489.

Sillanpää, M. J., 2011. Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. *Heredity* 106:511–519.

Stevenson, J. E., B. R. Wright, and A. S. Boydstun, 2012. The metabolic syndrome and coronary artery disease: a structural equation modeling approach suggestive of a common underlying pathophysiology. *Metabolism* 61:1582–1588.

Stich, B. and A. Melchinger, 2010. An introduction to association mapping in plants. *CAB Reviews* 5:1–9.

Stitt, M., R. Sulpice, J. Keurentjes, M. Planck, M. Plant, and G. M. S, 2010. Metabolic Networks: How to Identify Key Components in the Regulation of Metabolism and Growth. *Plant Physiol.* 152:428–444.

Storey, J. D. and R. Tibshirani, 2003. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100:9440–9445.

Suhre, K. and C. Gieger, 2012. Genetic variation in metabolic phenotypes: study designs and applications. *Nat. Rev. Genet.* 13:759–769.

Suhre, K., H. Wallaschofski, J. Raffler, N. Friedrich, R. Haring, K. Michael, C. Wasner, A. Krebs, F. Kronenberg, D. Chang, C. Meisinger, H.-E. Wichmann, W. Hoffmann, H. Völzke, U. Völker, A. Teumer, R. Biffar, T. Kocher, S. B. Felix, T. Illig, H. K. Kroemer, C. Gieger, W. Römisch-Margl, and M. Nauck, 2011. A genome-wide association study of metabolic traits in human urine. *Nat. Genet.* 43:565–569.

Sulpice, R., E. Pyl, H. Ishihara, S. Trenkamp, M. Steinfath, H. Witucka-Wall, Y. Gibon, B. Usadel, F. Poree, M. Piques, and Others, 2009. Starch as a major integrator in the regulation of plant growth. *Proc. Natl. Acad. Sci. U.S.A.* 106:10348–10353.

Svishcheva, G. and T. Axenovich, 2012. Rapid variance components–based method for whole-genome association analysis. *Nat. Genet.* 44:1166–1170.

Thornsberry, J. M., M. M. Goodman, J. Doebley, S. Kresovich, D. Nielsen, and E. S. Buckler, 2001. Dwarf8 polymorphisms associate with variation in flowering time. *Nat. Genet.* 28:286–289.

Tian, F., P. J. Bradbury, P. J. Brown, H. Hung, Q. Sun, S. Flint-Garcia, T. R. Rocheford, M. D. McMullen, J. B. Holland, and E. Buckler, 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* 43:6–11.

Tollenaar, M., 1989. Response of dry matter accumulation in maize to temperature: I. Dry matter partitioning. *Crop Sci.* 29:1239–1246.

Van Inghelandt, D., A. E. Melchinger, J.-P. Martinant, and B. Stich, 2012. Genome-wide association mapping of flowering time and northern corn leaf blight (Setosphaeria turcica) resistance in a vast commercial maize germplasm set. *BMC Plant Biol.* 12:56.

Verheul, M. J., C. Picatto, and P. Stamp, 1996. Growth and development of maize (Zea ways L.) seedlings under chilling conditions in the field. *Eur. J. Agron.* 5:31–43.

Wallace, J. G., S. J. Larsson, and E. S. Buckler, 2014. Entering the second century of maize quantitative genetics. *Heredity* 112:30–38.

Weng, J.-K., X. Li, N. D. Bonawitz, and C. Chapple, 2008. Emerging strategies of lignin engineering and degradation for cellulosic biofuel production. *Curr. Opin. Plant Biol.* 19:166–172.

Wong, J. C., R. J. Lambert, E. T. Wurtzel, and T. R. Rocheford, 2004. QTL and candidate genes phytoene synthase and zeta-carotene desaturase associated with the accumulation of carotenoids in maize. *Theor. Appl. Genet.* 108:349–359.

Wright, S., 1921. Correlation and causation. *J. Agr. Res.* 20:557–585.

Xin, Z., 2000. Cold comfort farm: the acclimation of plants to freezing temperatures. *Plant Cell Environ.* 23:893–902.

Xu, Y. and J. H. Crouch, 2008. Marker-Assisted Selection in Plant Breeding: From Publications to Practice. *Crop Sci.* 48:391–407.

Yu, J., G. Pressoir, W. H. Briggs, I. Vroh Bi, M. Yamasaki, J. F. Doebley, M. D. McMullen, B. S. Gaut, D. M. Nielsen, J. B. Holland, S. Kresovich, and E. Buckler, 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38:203–208.

Zhang, Z., E. Ersoz, C.-Q. Lai, R. J. Todhunter, H. K. Tiwari, M. a. Gore, P. J. Bradbury, J. Yu, D. K. Arnett, J. M. Ordovas, and E. Buckler, 2010. Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42:355–360.

Acknowledgements

First and foremost, I would like to thank my academic supervisor and mentor Prof. Dr. Albrecht E. Melchinger for his advice, suggestions, and continuous support during this thesis work.

Many thanks to Prof. Dr. Mark Stitt and Prof. Dr. Lothar Willmitzer for giving me the opportunity to work at the Max-Planck-Insitute of Molecular Plant Physiology. For mastering the metabolic and lipid profiling and all issues with 'Frosty, the robot' (sometimes until late at night), I would like to thank Dr. Angelika Czedik-Eysenberg, Dr. Yariv Brotman, Dr. Jan Lisec, and Dr. Ronan Sulpice. I highly appreciated the laboratory assistance of Bea Encke, Nicole Dittrich, and Manuela Günther.

Many thanks to Jean-Luc Jannink for the opportunity to join his working group at Cornell University. I am grateful to Jeffrey B. Endelman, Dr. Julie Dawson, Dr. Martha Hamblin, Delphine Ly, Dr. Deniz Akdemir, and Nicolas Heslot for their support during my stay in Ithaca.

I am grateful to Jochen Jesse, Franz-Josef Mauch, Nicole Friedl, Birgit Schilling, Beate Devezi, and many student workers for their help in mastering the leaf sampling in the field. Without your help, this thesis would have not been possible!

Sincere thanks to the teams of the research stations of the University of Hohenheim for their skilled design, realization, and data collection of the field trials.

Many thanks to Mrs. Kösling, Mrs. Lieb, and Mrs. Meyer for their help regarding organizational matters.

I would like to also thank Prof. Dr. Hans-Peter Piepho for suggestions regarding the statistical analysis.

Special thanks goes to Frank Technow for all the discussions about plant breeding, statistics, and science in general.

Many thanks to Prof. Dr. H.F. Utz, Dr. Tobias Schrag, Willmar Leiser, Michael Stange, Vanessa Windhausen, Christoph Grieder, Frank Technow, Alexander Strigens, Delphine Van Inghelandt, Dorcus Gemenet, Peter Muth, PD Dr. Bettina Haussmann, Elfadil Adam, Dr.

Xuefei Mi, Haixiao Hu, Han Sen, and all the other members of the Institute of Plant Breeding, Seed Science, and Population Genetics for creating a pleasant work environment.

Funding for this research came from the German Federal Ministry of Education and Research (BMBF) within the project GABI-Energy (grant ID 0315045) and the AgroClustEr "Synbreed — Synergistic Plant and Animal Breeding" (grant 0315528D). I greatly acknowledge the synbreed scholarship for my stay at Cornell.

Curriculum vitae

Name	Christian Riedelsheimer	
Birth	March 19th 1984 in Augsburg, Germany	
Current position	Since 07/2013	Research Scientist, DuPont Pioneer
University education	07/2010 – 06/2013	Ph.D. student, Institute of Plant Breeding, Seed Science and Population Genetics, Prof. Dr. Albrecht E. Melchinger, University of Hohenheim
	08/2012 – 12/2012	Visiting Scientist, Cornell University, IA, USA
	10/2008 – 06/2010	M.Sc. in Agricultural Sciences, University of Hohenheim
	10/2005 – 09/2008	B.Sc. in Agricultural Sciences, University of Hohenheim
Work experience	03/2009 – 10/2009	Lab work for M.Sc. thesis, Max-Planck Institute of Molecular Plant Physiology
	08/2008 – 10/2008	Internship in inbred line development, DuPont Pioneer, Dallas Center, IA, USA
	08/2001 – 07/2003	Vocational training in farming, cereal production and dairy cattle
School education	09/2003 – 07/2005	High school, Berufsoberschule Landshut-Schönbrunn
	09/1996 – 07/2000	Secondary school, Konradin-Realschule Friedberg

Erklärung

Hiermit erkläre ich an Eides statt, dass die vorliegende Arbeit von mir selbst verfasst und lediglich unter Zuhilfenahme der angegebenen Quellen und Hilfsmittel angefertigt wurde. Wörtlich oder inhaltlich übernommene Stellen wurden als solche gekennzeichnet.

Die vorliegende Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Institution oder Prüfungsbehörde vorgelegt.

Insbesondere erkläre ich, dass ich nicht früher oder gleichzeitig einen Antrag auf Eröffnung eines Promotionsverfahrens unter Vorlage der hier eingereichten Dissertation gestellt habe.

Stuttgart-Hohenheim, Juni 2013

Christian Riedelsheimer