Lehrstuhl für Genetik
der Technischen Universität München

Regulation of Benzoxazinoid Biosynthesis in *Zea mays*: Genomic requirement for the *Bx1* gene expression

2014

Linlin Zheng

Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt
Lehrstuhl für Genetik

Regulation of Benzoxazinoid Biosynthesis in *Zea mays*: Genomic requirement for the *Bx1* gene expression

Linlin Zheng

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende(r): Univ.-Prof. Dr. R. Hückelhoven
Prüfer der Dissertation:
1. Univ.-Prof. Dr. A. Gierl
2. Univ.-Prof. Dr. Ch-C. Schön.

Die Dissertation wurde am 18.02.2014 bei der Technischen Universität München eingereicht und durch die Fakultät Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 14.05.2014 angenommen.

Abbreviations

ANOVA	analysis of variance
A.thaliana(At)	*Arabidopsis thaliana*
BAC	Bacteria Artificial Chromosomes
bp	basepairs
BSA	Bovine serum albumin
cDNA	Complementary DNA
CHORI	Children's Hospital Oakland Research Institute
CIM	Composite interval mapping
cM	centimorgen
CTAB	$C_{16}H_{33}N(CH_3)_3Br$, cetyltrimethylammonium bromide
dai	days after imbibition
dCTP	Deoxycytidine triphosphate
DIBOA	2, 4-dihydroxy -1, 4-benzoxazin-3 (4*H*)-one
DIMBOA	2, 4-dihydroxy -7-methoxy-2*H*-1, 4-benzoxazin-3 (4*H*)-one
DMSO	Dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	Deoxynucleotide
EDTA	Ethylenediaminetetraacetic Acid
EST	Expressed Sequence Tags
EuNAM	European Nested Association Mapping population
g	gram
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
kb	kilobase
h	hours
HDIMBOA-Glc	2-hydroxy-4,7-dimethoxy-1,4-benzoxazin-3-one glucoside
Hd-Zip	Homeodomain-leucin zipper
HPLC	High-performance liquid chromatography
IBM population	intermated B73 × Mo17 population
IGL	indole-3-glycerol phosphate lyase
IL	Introgression Lines
l	liter
IPTG	isopropyl β-D-1-thiogalactopyranoside
LOD value	log of the odds value
m (min)	minutes
ml	mililitre
mM	milimolar
Mb	megabase
MIM	Multiple interval mapping
NAM population	Nested Association Mapping population
mRNA	messenger RNA
NILs	near-isogenic lines
nt	nucleotides

N	not analysis
PCR	polymerase chain reaction
pg	picogramm
QTL	quantitative trait locus
qPCR	quantitative polymerase chain reaction
RIL	Recombinant Inbred Lines
RNA	ribonucleic acid
RT	room temperature
RT-PCR	reverse transcriptase PCR
s	seconds
SDS	sodium dodecyl sulfate
SIM	simple interval mapping
SM	single marker analysis
SNP	single nucleotide polymorphism
SSPE	Saline sodium phosphate EDTA stock solution
Taq DNA polymerase	*Thermus Aquaticus* DNA polymerase
TE	Tris-EDTA buffer
tb1	*teosinte branched 1*
Tris	Tris-hydroxymethyl-aminomethane
TSA	tryptophan synthase
Tukey HSD test	Tukey honestly significant difference test
µg	microgram
µl	microliter
Vgt1	*Vegetative to generative transition 1*
Z.mays	*Zea mays*

1. Introduction

1.1 Secondary metabolites and chemical defense

The history of secondary product research started with the isolation of a natural product (principium somniferum, i.e. morphine) by Friedrich Wilhelm Sertürner in the early 1800s. Primary metabolism in a plant comprises all metabolic pathways that are essential to the plant's survival and development. However, secondary metabolism is essential because it comprises all interactions of plants with their biotic and abiotic environment (Hartmann 2007). Secondary metabolites function as defense against pathogens, herbivores or competing plants, and signal compounds to attract pollinating or seed dispersing animals. They have a wide range of chemical structures and biological activities. Secondary metabolites derived from unique biosynthetic pathways using primary metabolites and intermediates as a basis. The high genetic plasticity and diversity of secondary metabolism guarantees flexible adaptations of plants to the demands of their continuously changing environment. Preformed defence chemicals like the well-known alkaloids and glucosinolates are constitutively accumulated substances against pathogens. Phytoalexins are specifically produced upon pathogen attack. Allelochemicals mediate intercellular and interspecies communication (Hartmann 2007).

The metabolic flow from primary metabolism into secondary biosynthetic pathways is often generated by duplication and neofunctionalization of genes functioning in primary metabolism. Specific metabolites are the result of modification by oxygenases (e.g. cytochrome P450 enzymes and 2-oxoglutarate dependent dioxygenases) and transferases, (e.g. methyltransferases and glycosyltransferases). Generally, secondary metabolites interact not specifically with unique targets but interfere with common structures like lipids, amino acids, or DNA. Plants have developed strategies to circumvent self-intoxication. The compound might be stored in specialized organs or cells (Turner and Croteau, 2004), synthesized only upon challenge by the pathogen or pest, or stored as a stabilized less toxic derivate and activated upon the attack of pathogens and herbivores. Effective concentrations of phytochemicals are high (e.g. 0.1-1 mg per gram fresh weight for benzoxazinoids).

1.2 Benzoxazinoid biosynthesis in maize

Benzoxazinoids, identified in the early 1960s, are secondary plant metabolites characteristically found within the grasses, including wheat, rye and maize (Sicker *et al.*, 2000, Niemeyer, 1988; Niemeyer, 2009). Benzoxazinoids are also found sporadically in some species of two independent orders of the dicots, Ranunculales (*Consolida orientalis*) and Lamiales (*Acanthus species*, *Aphelandra squarrosa, Lamium galeobdolon;* Sicker *et al.*, 2000, Alipieva *et al.* 2003). DIMBOA (2, 4-dihydroxy -7-methoxy-2*H*-1, 4-benzoxazin-3 (4*H*)-one) is the prevalent benzoxazinoid in maize and wheat, while it is DIBOA (2, 4 -dihydroxy-2*H*-1, 4 -benzoxazin-3 (4*H*)-one) in rye (Sicker *et al.*, 2000, Niemeyer, 1988; Niemeyer, 2009). Benzoxazinoids serve as defenses against a variety of herbivores in the affected plants. In maize it has been shown that benzoxazinoids function in the control of

insects and pathogens, e.g. the European corn borer (*Ostrinia nubilalis*) and opposed oat aphid (*Rhopalosiphum padi*), leaf blight (*Helminthosporium turcicum*), and the stem rot (*Stenocarpella maydis*) (Niemeyer, 1988). Generally dry seeds are free of benzoxazinoids. In the grasses benzoxazinoid biosynthesis starts with germination and the level increases until 10 days after germination and decreases subsequently with age. Around 20 days after germination DIMBOA-Glc levels drop even further. It is believed that from this point on, the benzoxazinoids concentration is decreased not only by dilution caused by increased tissue volume and reduced biosynthesis (see below) but also by release and degradation (Cambier *et al.*, 2000). In adult plants DIMBOA-levels are too low for effective pest control, e.g. while the growth of first brood of the European corn borer is reduced in high DIMBOA lines there is no effect on the second brood that feeds on older plants (Klun *et al.*, 1970; Grombacher *et al.*, 1989).

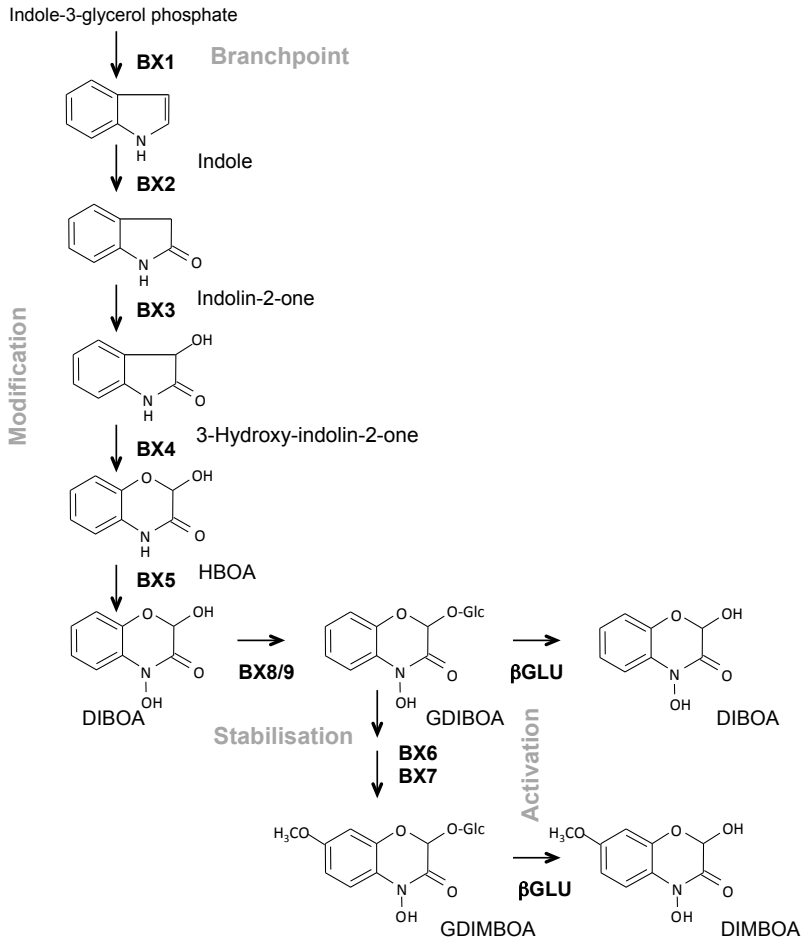


Figure 1: Benzoxazinoid biosynthesis pathway in maize. The branchpoint enzyme *Bx1* is homologous to the alpha-subunit of tryptophan synthase (TSA) of the primary metabolism. The chemical modification of the branchpoint intermediate indole involves four related cytochrome P450 enzymes (BX2-BX5) located in microsomes, the UDP-glucosyltransferases BX8 and BX9 stabilize the reactive intermediate DIBOA which is further modified by the cytosolic 2-oxoglutarate dependent dioxygenase BX6 and the *O*-methyltransferase BX7. The glucoside is stored in the vacuole.

Benzoxazinoid biosynthesis in maize, wheat and wild barley has been elucidated (Frey *et al.*, 1997, 2003; von Rad *et al.*, 2001; Jonczyk *et al.*, 2008; Grün *et al.*, 2005). In the grasses, a series of five genes (*Bx1, Bx2, Bx3, Bx4* and *Bx5*) is sufficient to encode the enzymes to synthesize DIBOA. The UDP-glucosyltransferase BX8 or BX9 stabilizes DIBOA via glucosylation. In maize, further modifications by BX6 and BX7 lead to the generation of DIMBOA-glucoside (Figure 1). The glucosides might be stored in the vacuole. The glucosides are hydrolyzed by plastidic β-glucosidases (Babcock and Esen, 1994) upon tissue disruption caused by herbivore damage or pathogens. The result is the release of biocidal aglucon.

In maize, most of the biosynthetic genes are located on the short arm of chromosome 4, except that *Bx9* is located on the long arm of chromosome 1 (Figure 2). The core gene cluster, *Bx1 -Bx5* and *Bx8*, is located within 264 kb; *Bx6* is located upstream of *Bx3*, and was mapped about 6 cM to *Bx1*, the position in kb is not yet known. *Bx7* is downstream to *Bx2*, but not in immediate proximity. Two specific β-glucosidases map to chromosome 10 (Sue *et al.*, 2011).

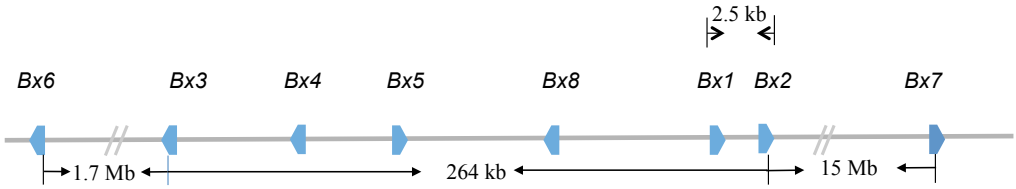


Figure 2: DIMBOA biosynthesis cluster on maize chromosome 4. Coordinates of the *Bx* genes on chromosome 4 are given for the B73 reference genome, v2 (www.maizesequence. org). *Bx1*: 3.256.291-3.258.246; *Bx2*: 3.260.685-3.262.840; *Bx3*: 2.998.897-3.000.920; *Bx4*: 3.046.687-3.049.508; *Bx5*: 3.108.397-3.110.973; *Bx7*: 18.370.242-18.371.515; *Bx8*: 3.210.119-3.212.030; *Bx6* is mapped upstream of *Bx3*.

In wheat and rye, benzoxazinoid pathway genes are dispersed into different chromosomes. In hexaploid wheat, *TaBx1* and *TaBx2* orthologs were identified in the same chromosomal bin on homologous group 4 chromosomes (4A, 4B, and 4D), while *TaBx3* to *TaBx5* homologs are in the same chromosomal bin on group 5 chromosomes (5A, 5B, and5D) (Nomura *et al.*, 2003, 2009). Like in maize, the pathway-specific glucosidases are not linked to the biosynthetic genes. Four *GT* loci (*TaGTa–TaGTd*) were mapped on chromosomes 7A, 7B (two loci), and 7D. Four glu1 loci (*Taglu1a–Taglu1d*) were on chromosomes 2A, 2B (two loci), and 2D (Sue *et al.*, 2011). The chromosome organization in rye is analogous to wheat, the *ScBx1* and *ScBx2* are clustered and located on chromosomes 7R; *ScBx3, ScBx4* and *ScBx5* genes are located to chromosomes 5R; *ScGT* and *Scglu* separately on chromosomes 4R and 2R, respectively (Nomura *et al.*, 2003; Sue *et al.*, 2011).

All *Bx* genes are highly expressed in the seedling (Frey *et al.*, 1997, v. Rad *et al.*, 2001, Jonczyk *et al.*, 2008). *Bx8* and *Bx9*, show a longer and more general expression compared to the other *Bx* genes (von Rad *et al.*, 2001). Detailed expression data for gene models containing annotated “classical genes”, syntenic orthologs, and phytozome annotations (Schnable JC *et al.*, 2011, 2012; Schnable PS *et al.*, 2009; Figure 3) include the core pathway

genes *Bx1-Bx5*. These tissue specific expression data confirm that *Bx8* is the only one with equal expression among all the tissues (Sekhon *et al.*, 2011; Winter *et al.*, 2007). Although *Bx8* is localized between *Bx3, Bx4, Bx5* and *Bx1, Bx2,* its regulation seems to be independent from the other *Bx* gene in the cluster. Interestingly *Bx1* and *Bx2* are identical in expression. A set of regulatory sequences might confer seedling specific expression, and additional elements might be responsible for fine-tuning. In hexaploid wheat, transcription of *TaBx1* to *TaBx5* was identified to be coordinated in the seedling, while the three genomes express the *Bx* gene to different levels (Nomura *et al.*, 2005).

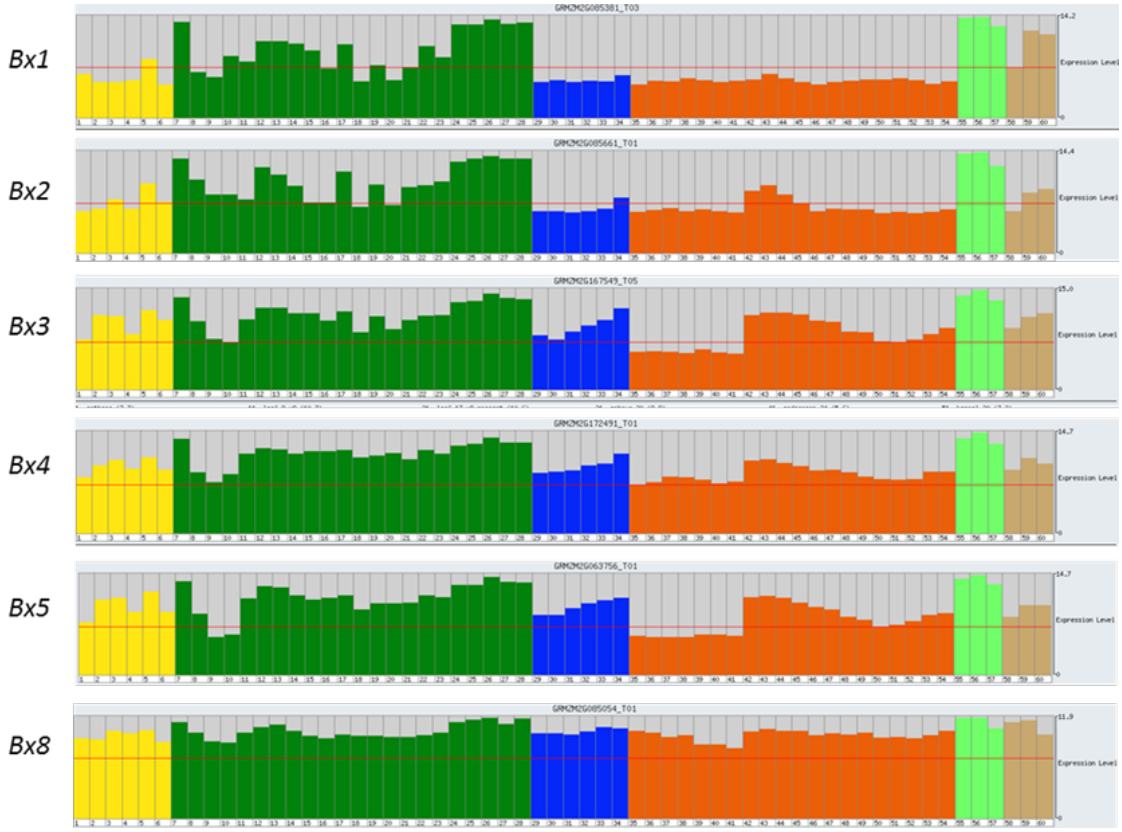


Figure 3: Tissue specific expression of the *Bx* gene cluster on chromosome 4. Columns 1-6 male and female flower (yellow), columns 7-28 leaves, increasing developmental stage (green); columns 29-34 embryo and seed (blue); columns 35-54 endosperm and kernel (orange); columns 55-57 stem (light green); columns 58-60 root (brown). Data from Sekhon *et al.*, 2011; Winter *et al.*, 2007. The red line gives background expression level.

In addition to the expression in the seedling stage, it became obvious that benzoxazinoid biosynthesis is induced by microbial pathogens and herbivores. Several investigations have shown that benzoxazinoids are formed after herbivore damage (Gutierrez, 1988; Niemeyer, 2009), and after treatment with methyl jasmonate or infection of microbes (Wang, 2007; Song, 2011). Recent investigation have shown that concordantly *Bx* gene transcription can be induced by pathogens and herbivores (Huffaker *et al.*, 2011; Dafoe *et al.*, 2013)

1.3 Regulation of plant secondary metabolic clusters

While gene clusters containing non-homologous functionally related genes of secondary metabolic pathways are common in bacterial and fungi genomes (Zheng *et al.*, 2002), the benzoxazinoids pathway of maize was the first example of such a cluster in plants (Frey *et al.*, 1997). However, with the rapid development of genome sequencing projects, huge amounts of sequence data are available nowadays, and more plant secondary metabolite gene clusters are emerging (Osbourn 2010a). By now, at least ten plant secondary metabolite gene clusters have been reported in addition to the *Bx* gene cluster. Clusters for the synthesis of triterpenes in oat (*Avena sativa*) and Arabidopsis (the avenacin, marneral and thalianol clusters; Papadopoulou *et al.*, 1999; Haralampidis *et al.*, 2001; Qi *et al.*, 2004, 2006; Field *et al.*, 2008; Mylona *et al.*, 2008; Mugford *et al.*, 2009; Osbourn *et al.*, 2010a), diterpenes in rice (*Oryza sativa*; the momilactone and phytocassane clusters, Sakamoto *et al.*, 2004; Wilderman *et al.*, 2004; 2007; Swaminathan *et al.*, 2009), anticancer alkaloid noscapine in Opium poppy (*Papaver somniferum*, Winzer *et al.*, 2012) and steroidal glycoalkaloids in potato (Itkin *et al.*, 2013). These clusters are diverse in organization and function and all appear to have evolved independently (Field *et al.*, 2008; Frey *et al.*, 2009; Swaminathan *et al.*, 2009; Osbourn *et al*., 2010a, b).

It is speculated that gene clusters have a selective advantage since superior allelic combinations are inherited preferentially once established in coupling phase (Osbourn *et al*., 2010a). Alternatively or additionally clustering might facilitate coordinate gene expression. However, the knowledge about regulation is limited. A unique example for gene cluster regulation in plants was given by Okada *et al.* (2009). A chitin oligosaccharide elicitor-inducible basic leucine zipper transcription factor, OsTGAP1 was identified to regulate the transcription of momilactone and phytocassene gene cluster in rice. Little is known about *cis*-elements or *trans*-factors, which may play a role in the regulation of benzoxazinoid biosynthesis on the gene level (Martin, 2003). According to Martin two homeodomain- leucin zipper (Hd-Zip) class I proteins (ZmHDZip 1 and ZmHDZip 2) bind as homo-and heterodimers to a sequence motif present upstream of *Bx2* and significantly increase its transcription in transient assays. Analyses of ZmHDZip 1/ZmHD-Zip 2 overexpressing transgenic maize lines and lines with mutations in these genes, however, showed no change in the *Bx* gene expression patterns. Thus, an essential function of the two transcription factors is not possible.

Standard methods like promoter reporter gene fusions in our lab did not result in the definition of promoter elements in *Bx1.* A reason might be that the major regulator is distal from *Bx1.* Long distance influence of *cis*-elements have been detected in maize: for *Vgt1*, *tb1* and *b1* regulatory sequences were located up to 100 kb upstream of the genes (Salvi *et al.*, 2007; Clark *et al.*, 2006; Stam *et al.*, 2002).

From fungus to human, different mechanisms have been shown to contribute to the co-expression of gene clusters (Splinter *et al.*, 2006; Soshnikova *et al.*, 2009; Meneghini *et al.*, 2003; Strauss *et al.*, 2011; Bok *et al.*, 2006). Mechanistically regulation of the gene cluster expression can be by long distance DNA looping (Tolhuis *et al.*, 2002). Chromatin modifications have been shown to have major impacts on cluster expression (Soshnikova *et*

al., 2009; Meneghini *et al.* 2003; Strauss *et al.* 2011; Bok *et al.* 2006). Recently Wegel *et al.* (2009) investigated the chromatin structure of the avenacin gene cluster in oat. Using DNA fluorescence *in situ* hybridization experiments, cell type-specific changes in chromatin conformation were revealed to be correlated with cluster expression. Global histone modifications in *A. thaliana* indicated that both the thalianol and the marneral gene clusters have strong histone H3 lysine 27 trimethylation markers, whereas the flanking genes have not (Zhang X *et al.*, 2007). Both findings underline the potential impact of chromatin conformation and histone modification on plant gene clusters. A final proof of an essential role of a specific chromatin conformation for gene activity in gene clusters is missing. Likewise, master regulators involved in expression of clustered plant genes have not yet been defined.

1.4 Application and methods of QTL mapping

Quantitative trait loci (QTLs) are genomic regions that contribute to the variation of quantitative traits. Most of the important agronomic and economic traits such as yield, protein content, fat content, secondary metabolite content, are quantitative traits, and jointly controlled by multiple genes or genomic loci. Quantitative traits are vulnerable to environmental impact (Johanssen, 1909).

QTL mapping exploits genetically heritable differences that can statistically be associated with a specific part of the genome. To identify a QTL, genomic markers are tested for their significant association with the variation of the trait values. The probability is assessed by the LOD score. To link genomic markers to a phenotype, several genotypes that differ at the tested marker positions need to be analyzed for their trait variation. High-density genetic linkage maps, statistical methods and the corresponding experimental population are all important for a comprehensive QTL mapping. In powerful populations, for example the Nested Association Mapping (NAM) population, genetic models can account for 80% of the genetic variance (Buckler *et al.*, 2009; Kump *et al.*, 2011; Tian *et al.*, 2011, Wallace *et al.*, 2013). Maize has a much more complicated genetic architecture than other model plants, such as rice or Arabidopsis. Usually only a few QTL exert the majority of control on traits in these models. However, in maize, traits can be controlled by dozens of QTLs. For example, over 30 small-effect QTLs were detected to influence flowering time in maize (Buckler *et al.*, 2009). In Arabidopsis and rice, only small numbers of QTLs with larger effects were found to control the flowering time variation (Izawa *et al.* 2003; EI-Lithy *et al.*, 2006).

Different methods of QTL mapping with increasing complexity have been developed. The basic level, single marker analysis (SM), involves tests of association between trait values and the genotypes at each marker locus (Edwards *et al.* 1987). Since these tests take each marker locus on a chromosome separately, the effects are underestimated and the position cannot be determined. Lander and Botstein (1989) proposed the simple interval mapping (SIM) as a systematical way to scan the whole genome for the evidence of QTLs. This method is an extension of the marker analysis by using two flanking markers to construct an interval for searching a putative QTL within the interval. However, if there is more than one QTL on a

chromosome, interference among the markers will influence QTL calculation. Furthermore effects of QTL identified by this method are likely to be biased. Zeng *et al.* (1994) introduced a method called composite interval mapping (CIM), in which markers outside of the test interval serve as covariates. CIM combines interval mapping with multiple regression analysis. The potential QTLs are selected by forward stepwise regression. CIM results contain often sharper, higher, and more peaks in the LOD curves than SIM, which can be recognized very clearly in this data set. Multiple interval mapping (MIM) is a multiple QTL oriented method combining QTL mapping analysis with the analysis of genetic architecture of quantitative traits through an algorithm to search for number, positions, effects and interaction of significant QTL simultaneously (Kao *et al.*, 1997).

Various computer programs are available to perform the sophisticated algorithms, such as MQTL (Tinker and Mather, 1995a and 1995b, SIM and CIM), PLABQTL (Utz and Melchinger, 1996, SIM and CIM), R/QTL (Broman *et al.*, 2003, SM, SIM, CIM and MIM) and QTL Cartographer (Wang *et al.* 2006, SM, SIM, CIM and MIM). In this project we took DIMBOA-content as quantitative trait and made the CIM analysis with QTL Cartographer.

1.5 Resources for gene mapping in maize

Genetic resolution is dependent on presence of recombination events. Additional opportunities for recombination in a population are provided if multiple generations of intermating are performed, thus the resolution in the resulting mapping population can be greatly enhanced (Liu *et al.*, 1996). Populations based on Mo17/B73 hybrids have a long history, since they serve as models for hybrid vigor. Hybrid vigor is an essential factor for yield. Lee *et al.* (2002) developed the intermated B73 × Mo17 (IBM) population by randomly intermating a F2 population derived from the single cross of the inbreds B73 × Mo17 for four generations. The Maize Mapping Project (http://www.maizegdb.org/ documentation/maizemap/overview.php, Schaeffer *et al.* 2011) and the development of illumina MaizeSNP50 SNPs (Ganal *et al.* 2011; Bauer *et al.* 2013) have essentially contributed to construct a densely marked high-resolution map for the IBM 302 population.

Numerous maize quantitative trait mapping studies were executed based on the IBM 302 population, for instance, resistance to southern leaf blight (Balint-Kurti *et al.*, 2007), carbon and nitrogen metabolism enzyme activities and seedling biomass (Zhang N *et al.*, 2010), density-related traits (Gonzalo *et al.*, 2010), and iron biofortification (Lung'aho *et al.*, 2011). Recently, Bommert *et al.* (2013) performed the fine mapping of a QTL detected for kernel row number in maize based on IBM 302 population. They found that CLAVATA receptor-like protein FASCIATED EAR2 controls the inflorescence meristem size and kernel row number. In recent years, a genetically diverse panel of 25 maize founder lines (NAM panel) was used to generate a nested association mapping (NAM) population of 5000 recombinant inbreds using B73 as constant parent (Flint-Garcia *et al.*, 2005; Yu *et al.*, 2008; McMullen *et al.*, 2009a). The NAM panel comprises about 80% of the variability found in the maize population. Hence for most genes, “extreme” alleles will be present in the panel. The NAM population has been used to map numerous maize quantitative traits, for example, DIMBOA-contents (Butrón *et al.*, 2010), maize leaf aphid (*Rhopalosiphum maidis*)

susceptibility of maize (Meihls *et al.*, 2013), resistance to northern leaf blight (Poland *et al.* 2011), and southern leaf blight (Kump *et al.*, 2011). Genome-wide association mapping revealed numerous QTLs with relatively small additive effect in all cases.

1.6 Research Objectives

As described above, it is well documented that benzoxazinoids are important defense compounds in maize. In addition to direct correlations between DIMBOA-content and defense (reviewed in Niemeyer 2009) many QTLs of insect and disease resistance in maize have been detected in close proximity to the *Bx* genes (McMullen and Simcox, 1995; Jampatong *et al.*, 2002; Brooks *et al.*, 2005, 2007; Cardinal *et al.*, 2006; Zwonitzer *et al.*, 2010; Kump *et al.*, 2011; Poland *et al.*, 2011) and might be ascribed to benzoxazinoids. However, benzoxazinoid biosynthesis generally is a juvenile trait, all *Bx* genes have expression maxima in the first week of germination, and high protective DIMBOA levels are only present in seedlings and young plants. The aim of the study was to screen the germplasm for prolonged maintenance of protective DIMBOA concentrations, to elucidate the underlying molecular mechanisms and eventually to provide tools to improve the chemical defense in breeding strategies. To this end the natural variation presented by the NAM core population was investigated and QTL mapping and molecular methods of analysis were applied.

2. Materials and Methods

2.1 Material

2.1.1 Chemicals and reagents

The chemicals used in this work (analytical grade) were provided by the company Bio-Rad (München), Roche (Mannheim), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) and Sigma-Aldrich Fluka (Hamburg, Seelze, Steinheim).

DNA restriction enzymes and DNA modifying enzymes were purchased from Roche (Mannheim), New England Biolabs (Frankfurt), Promega (Mannheim) and Qiagen (Hilden). Oligonucleotides were synthesized by both Biomers (Ulm) and Eurofins MWG Operon (Ebersberg).

2.1.2 Plasmids

For cloning the *E. coli* standard vectors pBlueskript KS+ (Agilent-Stratagene, Böblingen) and pGEM Easy (Promega) were used.

2.1.3 Bacteria strains

For preparing competent cells, the *E.coli* strains XL1-Blue (Bullock, *et al.*, 1978) and BL21 (Studier and Mofat, 1986) were used.

2.1.4 BACs

AC213878 (B73) is from Children's Hospital Oakland Research Institute (CHORI), and b0506A16 (Mo17) is from Bailin Li, Dupont Pioneer, USA.

2.1.5 DNA and RNA standards

1 kb Plus DNA Ladder TM, Invitrogen, Life technologies, Darmstadt
Gene Ruler TM 1 kb, Thermo Scientific, Life technologies, Darmstadt

2.1.6 Oligonucleotides

All products listed were synthesized by Eurofins MWG Operon (Ebersberg), Microsynth (Lindau) and Biomers (Ulm) and supplied freeze-dried. The primer were dissolved in TE buffer in a stock concentration of 100 mM and stored at -20 ° C. Working solutions were diluted in water to a concentration of 10 mM.

Table 1: Primers for qPCR with B73, Mo17 and hybrids, and PCR conditions.

	Primers	Sequences	Annealing ℃	Extensions	Acquisition ℃	Additive
Bx1	Bx1QF2	CGCAGCTTGGCCGAGATGAAA	50	27	80	5% DMSO
	Bx1QREV2	TWCTTTGTCATGGACTCATGGC				
Bx2	Bx2QF	CCGGGAGCTCACCGACATCAA	66	25	84	-
	Bx2QR	CTCCTGCCCCGCCGGCACGTT				
Bx3	Bx3FW4	CGGAACAGGCTGTTCAGCGAG	65	25	84	-
	Bx3REV4	TTCTTGCCGTCCGGCGAGC				
Bx4	Bx4QF	GCTCTCCGGCGGCAAGCAG	64	27	84	Q solution
	Bx4QR	CCGTACATGTCGACCTCGGC				
Bx5	Bx5QF	TGGAGATGAGGAAGCTGTCC	64	27	84	Q solution
	Bx5QR	CCGCTGTAGCTGTTGGACTT				
Bx8	GT2qua2f	TCGTCACGGCGCTCAACCCCGC	68	30	72	5% DMSO
	GT2LC1RUTA	GACTGCGTCGTCCTTGCGCTC				
Bx9	GT1F	TCGTCACCACGCTGAACGCCAG	68	30	72	5% DMSO
	GT1LC1RUTA	GGATCCTCCTTGCGCTCCTCTTTC				
GAPDH	GAPCF	GCTAGCTGCACCACAAACTGCCT	65	27	72	5% DMSO
	GAPCR	TAGCCCCACTCGTTGTCGTACCA				
Bx1 B73 specific	B73BX1F:	ACATCACCGCGGGCGACCCC	72	10	72	5% DMSO
	B73BX1R	AGGGGTCCGAGCAGGGCA				
Bx1 Mo17 specific	Mo17Bx1F	ACATCACCGCCGGCGACCCG	72	10	72	5% DMSO
	Mo17Bx1R	AGGGGTCCGAGCAGGGTACC				

Table 2: General primers for the NAM lines.

primers	sequences		Annealing (℃)	Extension (s)	Acquisition (℃	additive
Bx1	Bx1QgeneralFW	ATGGCTTTCGCGCCCAAAACG	60	16	84	5% DMSO
	Bx1QgeneralREV	GGCTCCTCCTCTCGGCGGGT				
Bx4	Bx4QgeneralFW	CCGGGAGCTCACCGACATCAA	62	16	85	5% DMSO
	Bx4QgeneralREV	GTGGCCGTACTTGGCGTGGA				

Table 3: Primers for fine mapping.

	primers	sequences	Annealing(°C)	Extension (s)	Cycles	additive
Bx1-gene	*Bx1* B73 specific primers and *Bx1* Mo17 specific primers in Table 1					
2kb	up1p2FW	GAGCAATGTCAACCTTTGGC	66	25	40	1M Betaine
	up1p2REV	ATGGTGGCTGCAGAAGGGAT				
3kb	up2p5KBFW	CTGCCATAGGAGCAGGGTAA	65	25	40	1M Betaine
	up2p5KBREV	CCCCTTCTCTCCCCTCCTTT				
5kb	MarkerA5kbB73F	CTCGTGATTCTTGTCTACTG	54	30	33	-
	MarkerA5kbB73R	AGGTTTGTATGGATCGACCA				
	MarkerA5kbMo17F	CTCGTGATTCCTGTCTACTA	51	30	35	1M Betaine
	MarkerA5kbMo17R	GTTTGTATATGGATAGACCT				
10kb	10kbB73Fw	ATGAAATGGTCAAGTTCA	58	20	35	1M Betaine
	10kbB73REV	GTTTGTGATTTTGTGTACAT				
	10kbMo17Fw	ATGAAATGATCAAGTTCG	58	20	35	1M Betaine
	10kbMo17REV	ATTTGTGATTTTGCGTACAG				
32kb	Bx8123kbFw	TCGAGAGGGACGGACTTAAC	68	30	40	1M Betaine
	21kbdownbx8REV2	TGCCAATCATGGCGGTTTCA				
43kb	Bx8up3F2	AGCCCTAGATCGCCAGGGA	68	30	40	1M Betaine
	Bx8up3R2	CTAGCACCTTGGGTGCGCT				
83kb	up80kbBACRev	TGGCAAGTCAAGAACAGAACC	65	27	40	1M Betaine
	up80kbBACFW	GGAGAGGAGGATGCTGCTTA				
137kb	3119224F3	TGTTTGGCTAGCTGTCGGAT	57	50	40	1M Betaine
	3119224R3	GGAGGGAGTACCTTGTGTAT				
αβ -marker	AlphaF	ACCATGCATGCTGAGAGAGA	57	50	40	1M Betaine
	AlphaR	AGTTGCACCGGAGCACATAT				
143kb	3113F14F2	GAGGAGCTCGACCGGTCAG	57	20	40	1M Betaine
	3113F14R2	TTTGTAGGGTTTGGGTGGGT				
148kb	marker2FW1	ATCCCTCCAATTCCCTTTGT	56	50	40	1M Betaine
	marker2REV1	GAGAATGTAGCTGTGGGGAA				
210kb	MarkerBM1B73F	CAAAAAATCTAGAATCTCAA	50	40	33	-
	MarkerBM1B73R	GGGATGACTAAGTCATGTCA				
	MarkerBM1Mo17F	TCAAACATCTAGAATCCCAT	50	40	33	-
	MarkerBM1Mo17R	GGGATGACTAAGCAATGTTG				

Table 4: Primers for determination of the genomic sequence of the αβ-region in Mo17.

	primers	Sequences	Annealing ℃	Extension	additive
1	MetFW3	CACAGTATTATTTGCTCCGT	56	1m40s	1M Betaine
	CHIPR1	GACAGGGTTGTTGTATATGCA			
2	Mo17D1D2CHIPF1	GGGCGCCATTGGCATTGTTG	50	50s	1M Betaine
	gMo17R1	GTCGTGTTCACACTCAACAT			
3	151AMo17F1	GGTGCATGTGCGAGAAAACC	56	2ms	1M Betaine
	3116KF9	GCACACTCTCAACACTACAA			
4	gMo17F2	GTTAGAACTTGGTAGCGAGT	52	50s	1M Betaine
	MetREV3	TGCCATCGATATCAAGGCAG			
5	151AB73F1	GGTGCATGTGCGACAACCTG	52	2m	1M Betaine
	gMo17R2	CAGTGACATATGTTCTCAGT			
6	gMo17F1	CCACATATTTCCAAAATCGA	52	2m	1M Betaine
	MREV3	TGCCATCGATATCAAGGCAG			
7	gMo17F4	GTTGAGTGTGAACACGACAT	53	1m30s	1M Betaine
	MREV3	TGCCATCGATATCAAGGCAG			
8	gMo17F4	GTTGAGTGTGAACACGACAT	53	1m30s	1M Betaine
	MREV4	TCACCCCATCCTAAATAAGG			

2.1.7 Plant material

The maize inbreds B73, Mo17, NAM founder lines, IBM population, and NIL lines were provided by the Maize Genetic Stock Center, Dr. Mike D. McMullen, USDA and Univ. Missouri, Columbia, USA, and Dr. Nathan Springer (University of Minnesota). The European collection of inbred lines for nested association mapping (EuNAM) lines were provided by Prof. C. -C. Schön and Dr. Peter Westermeier, Lehrstuhl für Pflanzenzüchtung, Technische Universität München. The recombinant inbred line MO038 was crossed to B73, fine mapping was made with F2 progenies. Two different cobs were used. The markers' distribution of B73 and Mo17 allele proved the integrity of the materials.

2.1.8 Plant growth conditions

The seeds were sterilized by shaking in 1.3% sodium hypochlorite for 1 min washed with tap water and incubated in 28°C with germination paper for 2 to 5 days in the dark before transfer to soil (Standard soil ED73 mixed with 10% w/w sand). 20 plants were in a tray. Incubation was in the plant growth chamber (Climate chamber HPS 2000 Heraeus VÖTSCH Industrietechnik GmbH) at 22°C for 16h in the light and 18°C for 8 h without light and constant moisture of 80%. To comparison, the plants were randomly placed in the trays, and were circulated every 24 hours.

2.1.9 Harvest of leaf material

The leaf blade of the 3rd leaf was cut at the border to the leaf sheath with the scissors. 3-4 harvested leaves of a line were "pooled", in 50 ml Falcon tubes weighed, immediately frozen in liquid nitrogen and subsequently stored until further processing at -70 ° C. The frozen plant material was ground into powder in a mortar using liquid nitrogen before performing

benzoxazinoids extraction and DNA and RNA analysis. Harvest time was between 10:00 to 16:00 h.

2.1.10 Kits and enzymes

Enzymes were purchased from Roche Applied Science, Mannheim, New England Biolabs (NEB) GmbH, Frankfurt and Promega GmbH, Mannheim.

The pGem ®-T Easy Vector system I kit was used for the ligation of inserts in the kit pGem ®-T Easy Vector.

The GE Healthcare Limited illustraTM GFXTM PCR DNA and Gel Band Purification Kit was used for purification of PCR fragments, the Macherey-Nagel (Düren) NucleoSpin ® RNA II kit was used to isolate RNA.

First strand cDNA biosynthesis was with **Applied Biosystems TaqMan ® Reverse Transcription Reagents kit** for cDNA synthesis.

qRT-PCR analysis was with **Bioline SensiFASTTM 2x SYBR No-ROX Kit.**

2.1.11 Databases

http://www.ncbi.nlm.nih.gov/
http://gbrowse.maizegdb.org/gb2/gbrowse/maize_v2
http://www.maizesequence.org/ index.html

2.2 Molecular biology methods and databases

2.2.1 Molecular biology methods

Isolation of DNA, RNA and generation of cDNA, ligation, transformation, restriction analysis, and Southern analysis was accomplished as described in Grün *et al.*, (2005) and according to the manuals of manufacturers. Quality and quantity of isolated RNA was checked by denaturating agarose electrophoresis as described by Schullehner *et al.*, (2008).

DNA isolation for fine mapping: Silica gel dried 7 days maize leaves were used for DNA-isolation in the fine-mapping approach. Isolation was in the 96-well plate format with the Qiagen TissueLyser II (Cat-No 19560), using the CTAB-method as described by Saghai-Maroof *et al.*, 1984. Briefly, 2-3 metal beads were added in each well to the dried plant material and 0.45 ml of CTAB β-mercaptoethanol mixture (150 mM Tris, pH 8.0, 1.0 M NaCl, 15 mM EDTA, 1.5 % cetyltrimethyl ammonium bromide, 1% β-mercaptoethanol) was added. A short vortexing, step was followed by 1 hour incubation at 60 °C, the mixture was inverted every 10 minutes. The mixture was chilled on ice for 10 min and extracted with one volume of chloroform/iso-amyl alcohol (24/1). RNA was digested by addition of RNase A (10 mg/L). Precipitation of the DNA was with 1 volume of isopropanol at -20 ° C. The pellet was washed with 75 % ethanol twice and dissolved in 50-100 µl TE buffer.

Sequence analysis was by Source Bioscience, Berlin, GATC-Biotech AG, Konstanz, Microsynth AG, Lindau, and Eurofins MWG Operon, Ebersberg.

2.2.2 PCR procedure

The standard PCR reactions was done by the thermoblock UNO (Biometra, Göttingen), using the GoTaq polymerase (Promega, USA) (primers in Table 3 and Table 4) and OneTaq-polymerase from NEB.

The quantitative RT-qPCRs were performed by lightcycler LC 480 from Roche, using the 2x SensiFASTTM SYBR No-ROX Kit from Bioline (Luckenwalde) (primers in Table 1 and Table 2).

2.2.3 Software and databases analysis software for DNA sequence files

Sequencing chromatograms were analyzed with 4Peaks for Mac OS X from Mekentosj BV, and Geospitza FinchTV for Windows.

Multiple DNA sequence alignment was by ClustalW 1.83 (http://www.genebee.msu.su/clustal/ [18.11.2012], by A.N. Belozersky, Institute of Physico-Chemical Biology, and Moscow State University.

Restriction sites analysis of the αβ-region was performed by NEBcutter2 (Vincze *et al.*, 2003), and Webcutter2 (http://rna.lundberg.gu.se/cutter2/)

Sequence transformations for the software reverse complement were used http://www.bioinformatics.org/sms/rev_comp.html [18.11.2012]. The copyright of this program is registered to Paul Stothard (Associate Professor of Bioinformatics and Genomics, Department of Agricultural, Food and Nutritional Science (AFNS), Canada).

For control and evaluation of RT-qPCRs, LightCycler ® 480 Software release 1.5.0 SP4 was used.

2.3 Biochemical analysis of benzoxazinoids

2.3.1 DIMBOA extraction

Benzoxazinoid extraction was executed as described by Grün *et al.*, 2005. Briefly 200 mg to 450 mg fine ground material was suspended in 3.3 volumes H_2O (w/v) and incubated for 1 h at RT. A pH of 2 was adjusted by adding approx. 30 - 35 µl 3 M HCl per 450 mg fresh weight. After incubation for 5 min at 65 ° C the debris was pelleted by centrifugation for 10 min at 17 500 x g. The supernatant was extracted with 0.6 ml ethyl acetate per 450 mg tissue. The supernatant was collected and the aqueous fraction was extracted two additional times. The unified organic supernatants were evaporated in the speed vacuum dryer at room temperature until completely dryness (about 45 minutes) and the pellet dissolved in 110 µl methanol. Insoluble material was removed by centrifugation for 15 min at 17 500 x g. The solution was stored at -20 °C or subjected directly to HPLC analysis. In most cases 100 µl

extract was analyzed.

2.3.2 HPLC analysis

The separation was carried out on a LiChroCART ® 250-4 LiChrospher ® 100 RP-18e (5µm) column (Merck, Darmstadt) and using Dionex 2284 Softron SP2. As mobile phases 0.3% formic acid and methanol were used at a flow rate of 1.0 ml/min according to the program described in Table 5. An upper pressure limit of 350 bar was set.

Table 5: HPLC-Program for the analysis of DIMBOA.

Time[min]	0.3 % Formic acid [%]	Methanol [%]
-3	80	20
1	80	20
5	60	40
11	57.5	42.5
12	0	100

Retention time for DIMBOA is 9.4 minutes.

Analysis was done with the chromatography data system Chromeleon version 6.80, HPLC chromatograms and the UV-VIS absorption spectra were analyzed. For calculating the relative benzoxazinoid concentration respectively the areas of all identified peaks were summed and the concentration was calculated using a calibration curve. The values were normalized with the material weight.

2.4 Statistical analysis tools

Microsoft Excel was used to calculate the average values, standard deviations of the DIMBOA-content and late *Bx1* expression level**.**

R (http://www.r-project.org) helped to accomplish boxplot Graphs drawing and ANOVA analysis (Statistics: ANOVA Tukey HSD test, α=0.05).

Composite interval mapping (Jansen and Stam, 1994; Zeng, 1994), for IBM population and selected IBM sub-population, were performed by the **Windows QTL cartographer 2.5** by 1000 permutations (Basten *et al.*, 1994, 1997). A high-density genetic map, with 1435 markers combining the genotypes of IBM 302 population provided by the Maize mapping project (http://www.maizegdb.org/ibm302scores.html) with SNP scores from University of Missouri, was used (Ganal *et al.*, 2011).

3. Results

Benzoxazinoids are major defense chemicals in maize and DIMBOA is the predominant aglucone. The *Bx* genes that constitute the biosynthetic pathway in maize have been isolated (reviewed in Frey *et al.*, 2009). High concentrations of DIMBOA are characteristically found in young plants and older plants have greatly reduced levels that are not sufficient to control herbivores and microbial pathogens. A survey making use of the genetic diversity of maize that is represented in the NAM founder line panel (McMullen *et al.*, 2009b) was made to uncover maize lines with divergent DIMBOA expression. The aim of the thesis was to detect maize lines with "late" high DIMBOA-content and use QTL mapping for elucidation of the underlying molecular mechanisms. The DIMBOA-content might be influenced by genotype, environment and interaction between genotype and environment. To minimize the variation caused by environment, the plants were grown in growth chambers under controlled conditions. The chambers can accommodate plants about 4 weeks after imbibition. The morphology of the different inbred lines does not differ significantly at this time point. The analysis was restricted to 23-25 days old plants and the blade of the third leaf. The glucosidase reaction that generates the aglucone is fast and hard to control during benzoxazinoids isolation. For this reason the glucosidase reaction was allowed to proceed completely before isolation and analysis of DIMBOA that hence represents the sum of its glucoside and aglucone.

The actual concentration of DIMBOA in a tissue is the product of biosynthesis, metabolization and transport processes. Hence the expression of the biosynthesis genes (*Bx* genes) might have major influence on DIMBOA-content. There is indication that the expression of the *Bx* genes is not only dependent of proximal promoter elements, but might generally be greatly influenced by global elements like chromatin modifications. These mechanisms cannot be revealed by standard methods like promoter reporter gene fusions. DIMBOA-content might hence be used as readout for *Bx* gene activity and might allow identifying regulators that contribute to *Bx* gene expression by QTL mapping.

3.1 Characterization of the NAM core population

The NAM core population comprises about 80% of the genetic variability present in maize. DIMBOA was extracted from the seedlings and 3rd fully developed leaf of 24 days after imbibition (dai) in 20 lines of the NAM core population and quantified with HPLC. All lines have high benzoxazinoid content in the young root and shoot (4 days after imbibition, data not shown) and hence are competent in benzoxazinoid biosynthesis. DIMBOA-content of 24 dai "late" plants is significantly reduced compared to seedlings (about 10% of the seedlings' content). However, lines B97, Ky21, Mo17, M162W, M37W, and MS71 have a significantly higher DIMBOA-content (Figure 4). In these lines, the benzoxazinoids content could be sufficient for pest control (>1.5 mM).

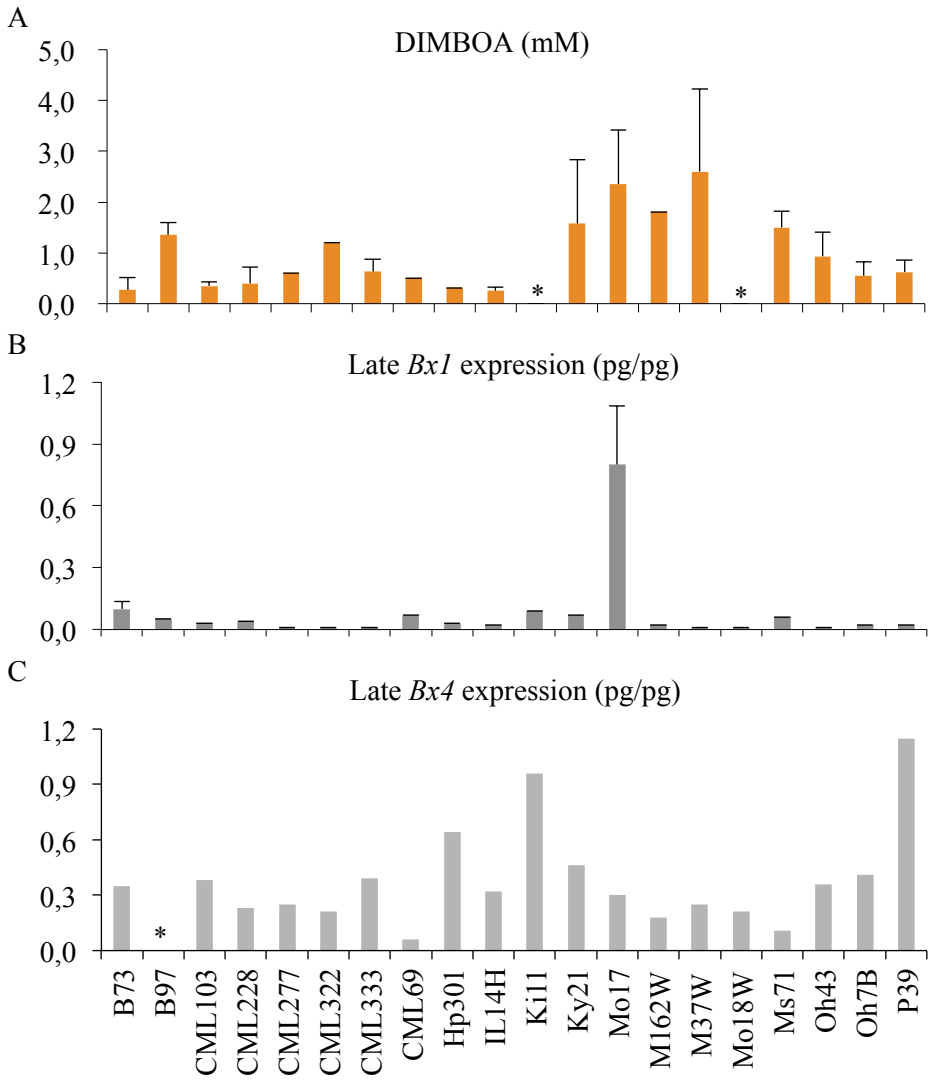


Figure 4: Analysis of NAM parental lines. A: DIMBOA-content; B: *Bx1* expression level; C: *Bx4* expression level. Plants are harvested at 24 days after imbibition (dai), the blade of the third leaf is taken for the analysis. The DIMBOA concentration is given in relation to the fresh weight. The gene expression of the *Bx* genes is normalized to *GAPDH*. Line Mo18W was not analyzed for DIMBOA-content and line B97 was not analyzed for *Bx4* expression, indicated with *.

To assay a correlation between DIMBOA-content and *Bx* gene expression, general primers for *Bx1* (Bx1QgeneralFW, Bx1QgeneralREV) and *Bx4* (Bx4QgeneralFW, Bx4QgeneralREV) were designed for the NAM lines based on the Panzea Database of NAM mRNA raw sequences (http://www.panzea.org/database/; Table 2). No general primers for *Bx2, Bx3*, and *Bx5* expression analysis were available for the NAM core population, but primers for these genes were designed to fit for B73 and Mo17 (Table 1).

Bx1 expression assays in late plants indicated that Mo17 is the unique line with high late *Bx1* transcription level among analyzed NAM lines, the expression level is at least 10-times higher compared to all other lines (Figure 4A). Highest *Bx4* transcript levels are present in the line P39 (1.13 pg *Bx4*/pg *GAPDH*), lowest in CML69 (0.09 *Bx4*/pg *GAPDH*), but the differences

between the lines are not as pronounced as for *Bx1* (mean value is 0.39, standard deviation is 0.27 (Figure 4B). The *Bx4* expression differences are not reflected in the DIMBOA-content (Figure 4C).

3.2 Characterization of *Bx* gene expression in B73, Mo17 and hybrid lines

Mo17 and B73 represent late DIMBOA-content and *Bx1* expression extremes. These two lines are of specific interest as models for heterosis since best-parent heterosis is displayed for many traits, e.g. biomass and kernel yield and provide many resources, e.g. the reference maize genome is B73, and efficient mapping populations based on Mo17 and B73 parents are available.

Transcription levels of the core pathway genes leading to DIMBOA-Glc (*Bx1*, *Bx2*, *Bx3*, *Bx4*, *Bx5*, *Bx8, and Bx9*) were assayed by realtime reverse transcriptase PCR (RT-qPCR) for the seedling, 3rd leaf of 14 dai and 24 dai plants of B73 and Mo17. The primers fit for both lines. In biological replicates, the time series were taken from seed aliquots that were imbibed in parallel. Differences in relative amounts may mainly be due to seed quality and slight differences during growth that might lead to minor differences in developmental stage. DIMBOA-content was determined in parallel (Figure 5). In both lines DIMBOA levels are highest in the seedling and decrease with age (Figure 5A). At all time-points the level is higher in Mo17, however, the difference increases from less than twofold in the seedling to at least fivefold in 24 dai plants. The expression pattern of the *Bx* genes in the seedling is almost identical for Mo17 and B73. All gene expression levels are decreased at 14 dai, but the *Bx1* transcript level is significantly high in Mo17. This difference in *Bx1* expression is consistent for 24 dai plants. At this stage *Bx1* transcript is hardly detectable in B73 plants. In the 24 dai Mo17 plants, *Bx1* expression is even higher than in 14 dai plants of parallel biological replicates (Figure 5B).

Next, the hybrids were added to the analysis. The DIMBOA-content and *Bx1* expression pattern was analyzed for parental lines, B73×Mo17 and Mo17×B73. Values are similar for both hybrids; hence no maternal effect for DIMBOA-content and *Bx1* expression exists. For the lines and hybrids and the different developmental stages a correlation between DIMBOA-content and *Bx1* transcript level was detected (Figure 6).

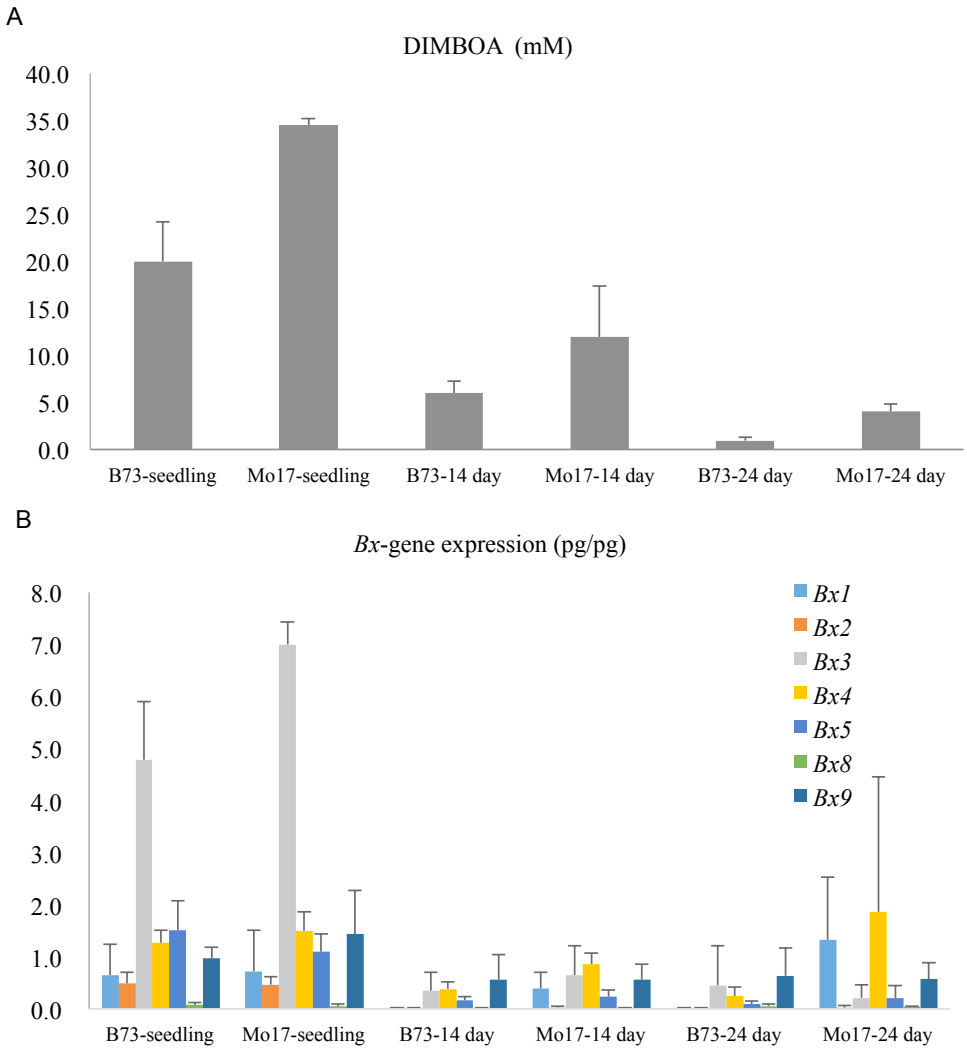


Figure 5: Characterization of DIMBOA-content (A) and *Bx* gene expression (B) in different developmental stages. Material used: Seedling 4 days: shoots; 14 days: leaf 2 plus leaf 3 of 14 days' plants; 24 days: leaf 3 from 24 dai plants. The DIMBOA concentration is given in relation to the fresh weight. The gene expression of the *Bx* genes is normalized to *GAPDH*. (pg/pg, normalized to *GAPDH*).

As in the previous experiments the DIMBOA level of B73 is lower than the one of Mo17 in the seedling. Mo17 and the hybrids are not significantly different in DIMBOA-content (Figure 6A) and *Bx1* expression levels in the seedlings are equivalent for lines and hybrids (Figure 6 B). A different picture is displayed in 24 dai plants: Mo17 has about 10-folds the DIMBOA-content compared to B73 (Figure 6D) and in contrast to seedlings the hybrids do not reach the high Mo17 DIMBOA level but are close to the mid-parent value. Neither in seedlings nor in older plants is best-parent heterosis displayed for DIMBOA-content. *Bx1* levels in both 24 dai hybrids are significantly higher than in B73 but reach only about 25% of the Mo17 level. Hence the *Bx1* transcript level of hybrids is below the additive value in older plants. The finding indicates the existence of one or more *trans*-negative factors influencing *Bx1* steady state levels in the older hybrids.

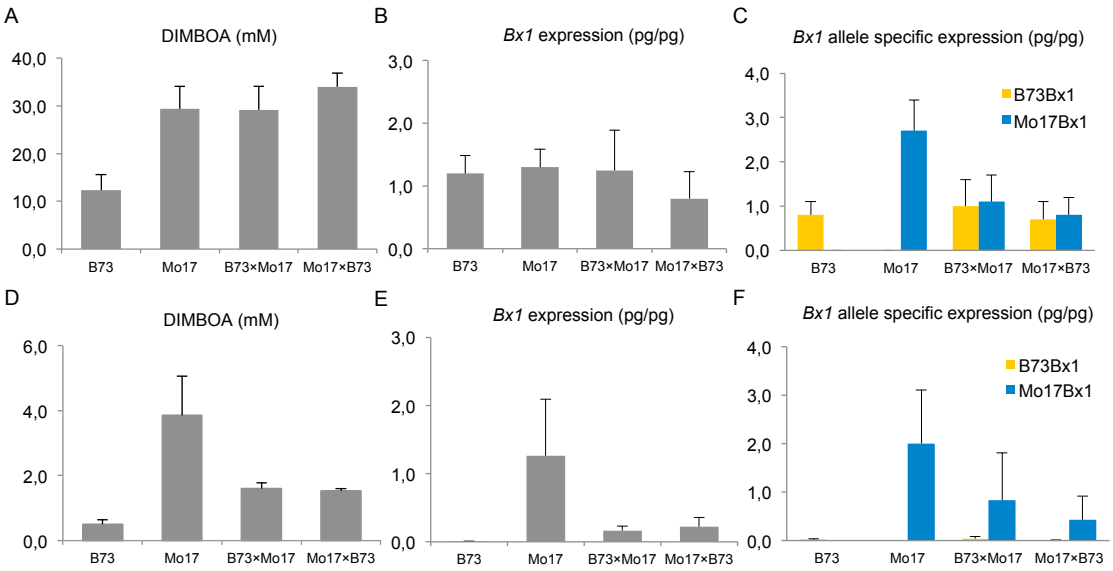


Figure 6: Analysis of DIMBOA-content and *Bx1* gene expression in B73, Mo17 and hybrids. A, D: DIMBOA-content. B, E: *Bx1* gene expression in relation to *GAPDH*. C, F: allele-specific *Bx1* gene expression, normalized to *GAPDH*. A, B, C: seedlings as in Figure 5. D, E, and F: 24 dai plants as in Figure 5. Three biological replicates were analyzed.

Allele specific primers (Table 1) were designed to analyze the expression of the parental *Bx1* alleles separately in the hybrids. Because allele specific and general analysis of *Bx1* by RT-qPCR is different with respect to amplification standards, fragment length and sequence, the sum of the allele specific transcript levels and the results for the general primer that theoretically should be identical differs. Systematically the value for the general analysis is less by a factor of 1 to 2 for Mo17 and the hybrids. However, the data demonstrate clearly that both alleles are expressed equally in the seedlings (Figure 6C). By contrast, in the late plants, the Mo17 alleles are exclusively expressed (Figure 6F). This indicates that a *cis*-element is influencing the expression of the gene. In the hybrids contribution of *trans*- and *cis*-effects can be approximated. If no *trans*-factors are effective the expression of the alleles should be the half of the parental lines. In case of the *Bx1* B73 allele the value is decreased by the factor about 0.05, the *Bx1* Mo17 allele is lowered by the factor about 0.17. The contribution of the *cis*-factor is given in the comparison of both alleles in the hybrid that contributes the same *trans*-factors for both alleles. This factor is about 60 when the Mo17 allele is compared to the B73 allele. Both *cis*- and *trans*-factors contribute to *Bx1* gene expression in the hybrids. While hybrid combinations of *trans*-factors are negative for both alleles, the Mo17-allele of *Bx1* is connected to a positive *cis*-element.

3.3 QTL analysis and fine mapping

B73 and Mo17 inbred lines have different DIMBOA-content and different *Bx1* steady state level in 24 dai 3rd leaf. Therefore, the lines can be taken as parents for QTL analysis for these traits. The two lines were used recently to generate the IBM 302 recombinant inbred line population (Lee *et al.*, 2002). The generation of the IBM 302 population included four

generations of random mating of the F2 generation.

3.3.1 QTL analysis with IBM 302 population

High quality data for late DIMBOA-content viz. at least three biological replicates with moderate standard deviations, were generated for 267 IBM302 lines. It turned out that DIMBOA-contents within IBM 302 have continuous values as expected for a quantitative trait. There is large variation among the lines, the lowest line (MO0382) has a DIMBOA level close to the detection limit, while the highest line (MO334) is 6 times higher than the high parent Mo17 (Figure 26, Appendix). The values of 92 lines are similar to B73 or lower, 91 lines are similar to Mo17 or higher, and 86 lines are between B73 and Mo17 (Figure 26, Appendix).

The scores of the IBM 302 population provided by the Maize Mapping Project (http://www.maizegdb.org/ibm302scores.html, Schaeffer *et al*., 2011) and SNP scores from University of Missouri were used to construct a high-density genetic map with 1435 markers. The Windows QTL cartographer version 2.5 software packages were used to detect QTLs. Composite interval mapping (CIM) was used to map QTLs for DIMBOA-content in IBM 302 lines. Three consistent QTLs, *QTL1* (chromosome 1), *QTL4-1* (chromosome 4), *QTL5* (chromosome 5) and a small QTL *(QTL4-2* chromosome 4*)* have been detected and are consistent in 1000 permutations (Figure 7).

The major QTL (*QTL4-1*) is on the short arm of chromosome 4, in the same position as the *Bx* gene cluster. It explains 21.5% of the phenotypic variation; the corresponding marker is PUT-163a-94477564-4893. Its major peak is around 200kb upstream of *Bx1* (position 3.109.838, based on AGI's B73 RefGen_v2) with a LOD value about 15.2.

Table 6: QTL analysis in IBM302 population (Composite interval mapping (CIM)).

		QTL cartographer						
QTL	Chromosome	Marker interval[a]	Confidence interval(bp)	Peak Location	LOD	Threshold[b]	Additive effect	R2 (%)[c]
QTL1	1	isu146-bnlg2057	176,512,769-185,017,139	umc1972	4.2	3.4	0.67	5.6
QTL4-1	4	SYN17726-SYN1073	2,765,157 -3,452,186	PUT-163a-94477564-4893	15.2	3.4	-1.27	21.5
QTL4-2	4	npi386-PZE-104031917	22,710,598- 38,725,625	PZE-104029222	2.7	3.4	-0.49	3.3
QTL5	5	umc126a-npi458a	190,689,469 -200,494,124	php20531	4.3	3.4	-0.7	6.1

a: Markers on both sides of the interval; b: Threshold calculated by 1000 permutations; c: Amount of phenotypic variation explained by each identified QTL.

The *QTL1* on chromosome 1 is detected with the marker umc1972; the LOD value is 4.2. *QTL5* on chromosome 5 is located at php20531 and has a LOD of about 4.3. *QTL4-2* is downstream of the *Bx* gene cluster; its corresponding marker is PZE-104029222. *QTL1*, *QTL5*, and *QTL4-2* account for 5.6%, 6.1% and 3.3% of the phenotypic variation, respectively. For *QTL4-1*, *QTL4-2* and *QTL5*, Mo17 alleles increase the trait value, while at

QTL1 the positive contribution is from B73 (Table 6). In the following the analysis was concentrated on the major *QTL4-1*.

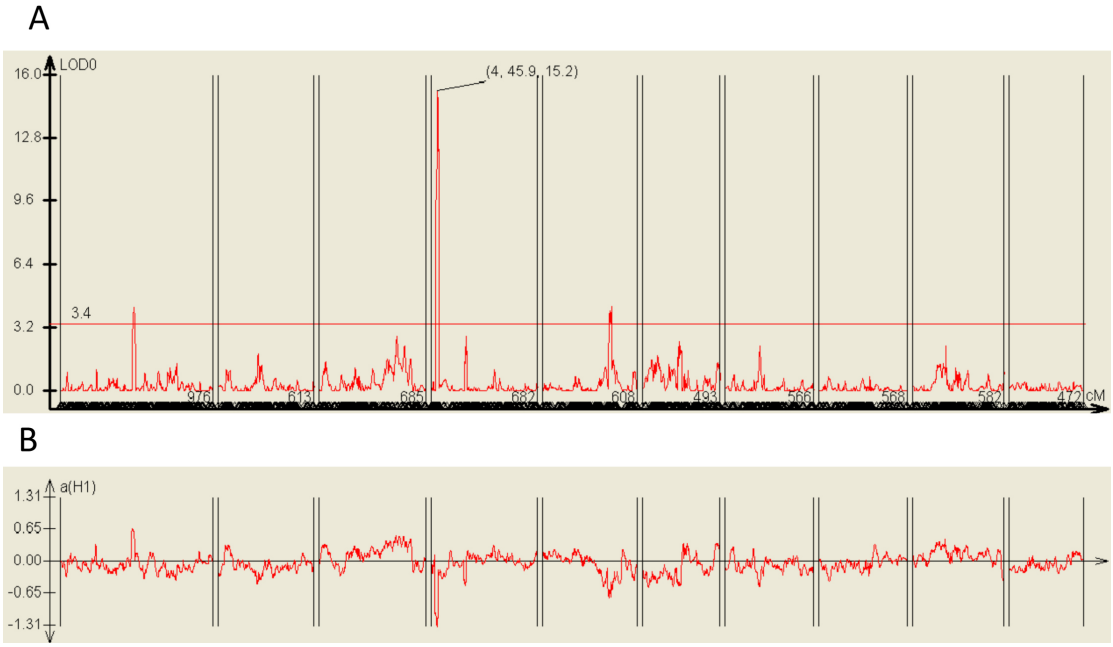


Figure 7: Composite interval mapping analysis. A: LOD scores for the IBM 302 population evaluated for DIMBOA-content. Three values are above the threshold of 3.4 in 1000 permutations. The major peak is localized on chromosome 4 at the position of the *Bx* gene cluster. Its LOD value is 15.2. B: Additive effect of the parental genomes. Values above zero indicate B73 contribution; values below zero indicate contribution by Mo17.

3.3.2. Construction of the fine mapping population and mapping of *QTL4-1*

3.3.2.1 Development of markers to reveal polymorphisms between B73 and Mo17

More markers were needed for the fine mapping of the *QTL4-1* high confidence interval (3.046.940 bp- 3.257.533 bp on chromosome 4, based on AGI's B73 RefGen_v2). To develop new markers, unique sequence regions with line-differences had to be identified. To do this 5 to 10 kb long consecutive stretches of the region were subjected to BLAST analysis (Schaeffer *et al.*, 2011). Unique (sub-) regions defined in this way were PCR-amplified on BAC-target DNA from B73 and Mo17. The PCR fragments were subjected to sequence analysis and analyzed for alteration between the lines. The designed primers are usually between 19 bp and 25 bp with a GC-content between 40% and 60%.

Thirteen markers (*Bx1*-gene, αβ marker, 2kb, 3kb, 5kb, 10kb, 32kb, 43kb, 83kb, 137kb, 143kb, 148kb, and 210kb, the latter ones named according to the distance to the translation start codon of *Bx1*) were designed in this region (Table 7, Figure 9). The region between the markers 83kb and 137kb has many gaps in the reference sequence and the search for unique sequences for marker design failed.

Table 7: Markers developed for fine mapping.

Name	Position (bp)	Detection	Type
Bx1-gene	3.256.736-3.256.843	in gel	plus/minus
2kb	3.254.405-3.254.897	sequencing	SNP
3kb	3.253.475-3.253.979	sequencing	SNP
5kb	3.250.998-3.251.489	in gel	plus/minus
10kb	3.246.208-3.246.410	in gel	plus/minus
32kb	3.224.206	sequencing	SNP
43kb	3.214.107-3.214.245	sequencing	indel
83kb	3.173.538-3.173.540	sequencing	indel
137kb	3.119.173-3.119.728	sequencing	indel
αβ marker	3.115.056-3.115.570	sequencing	duplication
143kb	3.113.527-3.113.362	in gel	indel
148kb	3.108.448-3.108.129	in gel	plus/minus
210kb	3.045.994-3.046.270	in gel	plus/minus

The name of the marker indicates the position upstream of *Bx1* or the specific feature. The position of the marker is based on the version 2 (v2) of the B73 sequence. The position of the fragment flanking the primers or the position of the SNP is indicated.

Four different types of markers were developed based on the sequence information gained in this way: plus/minus markers (Figure 8), indel (insertion and deletion) markers, SNPs, and a Mo17-specific duplication (αβ-region, Table 7). In the case of indel markers it was possible to design primers for the amplification of fragments of different size that can be discriminated by agarose gel or SDS-page gel analysis. 43kb and 143kb are this type of markers. A diagram of the maker distribution is given in Figure 9.

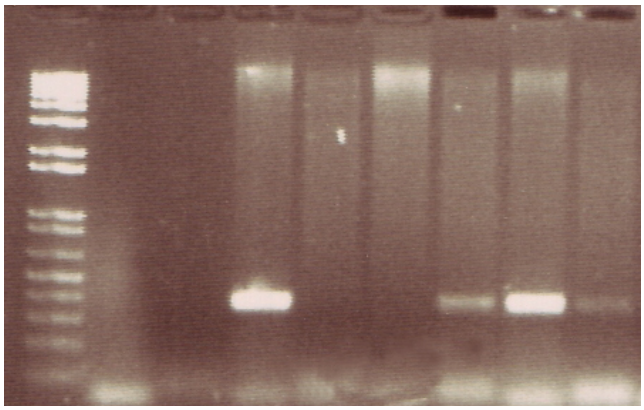


Figure 8: Example for the use of +/- markers. The genotype of B73, Mo17 and the hybrid MO038xB73 is determined. Lane 1, 3, 5, 7: 210kb marker, B73- specific primers. Lanes 2, 4, 6, 8: 210kb marker, Mo17-specific primers**.** Lane 1, 2: No DNA control. Lane3, 4: B73 genomic DNA. Lanes 5, 6: Mo17 genomic DNA, Lanes 7, 8: MO038×B73 hybrid DNA. M: DNA-size marker 1kb Plus.

The markers 5kb, 10kb, 148kb and 210kb were designed as +/- markers, for each marker two pairs of primers were designed from B73 and Mo17 allele separately (Figure 8). Markers that have only one or a few dispersed SNPS, were detected by Sanger sequencing, e.g. 2kb, 3kb, 32kb, and 83kb markers (Table 2, Table 7).

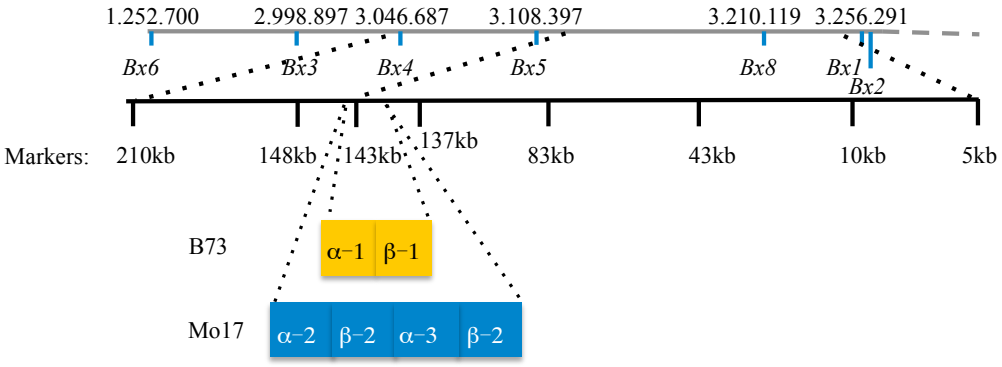


Figure 9: Distribution of markers of *QTL4-1*. The center of *QTL4-1* is flanked by the marker 210kb and 5kb. The positions of the markers and the *Bx* genes are indicated.

The αβ marker represents a 3874 bp stretch of unique DNA that is duplicated in Mo17 (between 3.112.492 bp and 3.116.365 bp on chromosome 4, based on MGSC B73 RefGen_v2). The sequence is formally divided into two parts, α and β. The β-part begins at 2629. The conformation in B73 is α−1/β−1, and Mo17 α−2/β−2/α−3/β−2 (Figure 10). The alleles β−1 and β−2, α−1 and α−2 have only minor differences. The Mo17 β−2 has one 4 bp insertion, two 1 bp deletions and 9 SNPs, α−2 differs from the B73 allele by a 21 bp deletion and 2 SNPs. By contrast, α−3 is greatly altered (Figure10A). The first 735 bp and the last 1106 bp have homologies of 84 and 94% with 5 and 4% gaps, respectively. Between these conserved regions major rearrangements took place. About 500 bp of the B73-sequence are missing, a sequence element of about 350 bp present in B73 is altered (87 to 95% homology) and present in tandem duplication. A part of all three α alleles can be PCR amplified using the primer pair AlphaF and AlphaR. The resulting fragments were sequenced after cloning. When Mo17 genomic DNA and Mo17 BAC DNA are used as templates the amplified PCR

fragments consist of a mixture of α-2 and α-3. Sequencing leads to double labeling at the site of divergence of the two duplicates in Mo17 (see appendix). The fragment amplified with the Alpha primer pair includes the region where β-1 and β-2 differs and hence can be used to recognize recombination within the αβ region. The organization of the duplicate was verified by Southern blot on B73 and Mo17 BACs (Figure 10), and then further confirmed by sequencing of genomic DNA of Mo17. The 3.6 kb region contains a mixture of unique sequences and small repetitive elements. There are no predicted genes in this region.

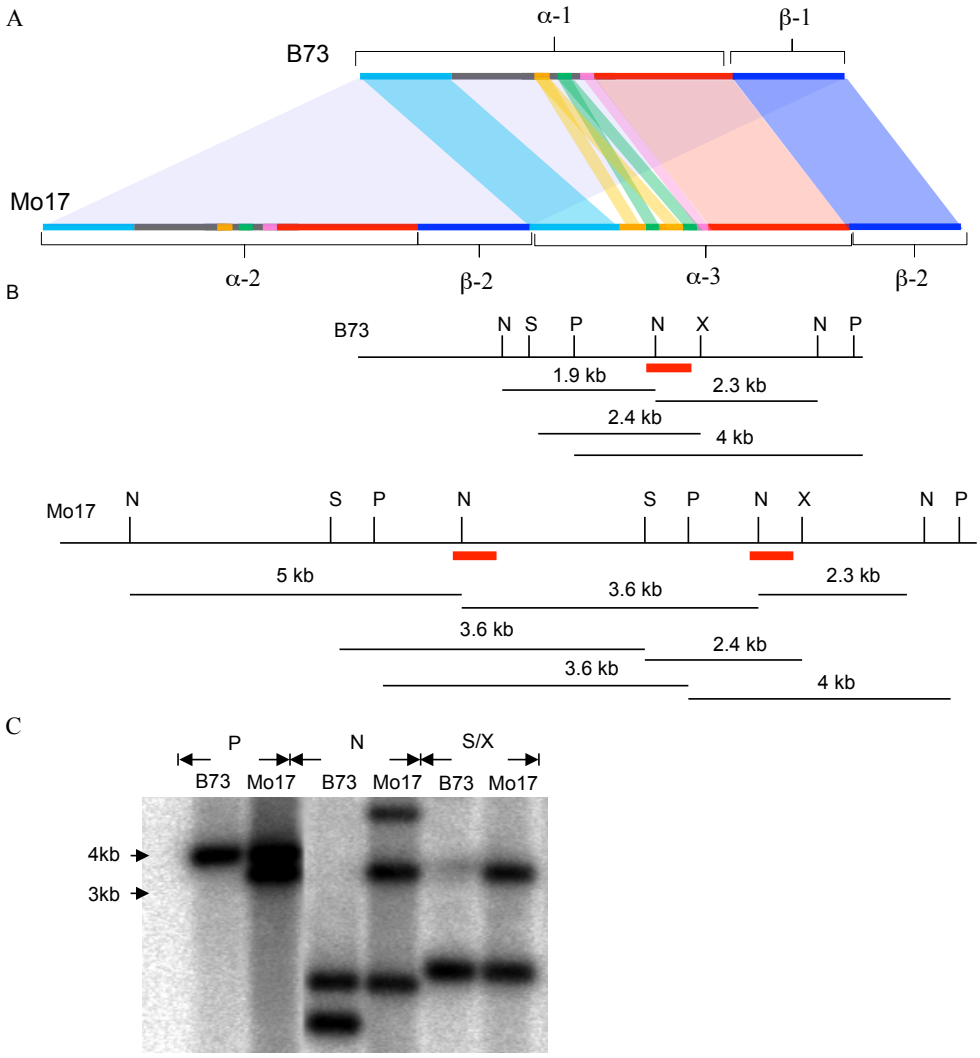


Figure 10: Organization of the αβ-region. A: Diagram of the αβ–region in Mo17 and B73. A duplication of the whole region is present in Mo17. The β-alleles have a homology of 99%, the alleles α-1 and α-2 differ by a 21 bp deletion and 2 SNPs in the blue region. The α-part can be subdivided in regions with different degree of homology between the alleles α-1 and α-3. Blue and grey 99%, orange: 84%, green: 92%, pink: 98%, red: 87%. Gray lines indicate regions that are not present in α-3. B: Restriction map. C: Southern analysis of B73 and Mo17 BAC DNA using a 600 bp fragment indicated by the short red line as probe. An extra band of 3.6 kb is present in Mo17. N: NcoI; S: SalI; P: PstI; X: XhoI.

3.3.2.2 Determination of recombinants in the fine mapping population

For *QTL4-1* fine mapping, MO038 was selected as parental line in the cross with B73. MO038 is Mo17 genotype at *QTL4-1* (0 bp- 4.908.092 bp*)*, *QTL4-2* (B73, 2.756.157-37.101.516; Mo17, 37.101.516-38.725.625), B73 genotype at *QTL1*, and Mo17 at *QTL5*. MO038 has high DIMBOA-content and *Bx1* expression level in 24 dai plants (Figure 11A, B). In the hybrid MO038×B73, allele specific expression is displayed (Figure 11C).

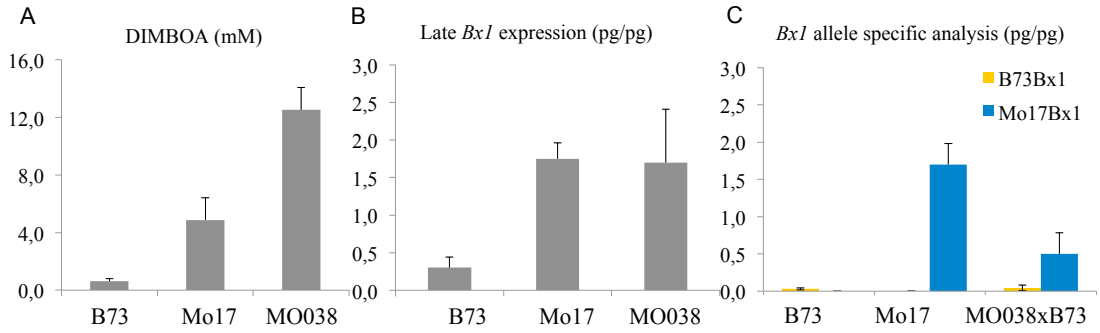


Figure 11: Characterization of the inbred line MO038. A: DIMBOA-content in the third leaf of 24 dai plants; B: *Bx1* level in 24 dai plants. C: Allele specific *Bx1*. The DIMBOA concentration is given in relation to the fresh weight. The gene expression of the *Bx* genes is normalized to *GAPDH*. (pg/pg, normalized to *GAPDH*).

Markers 5kb and 210kb were used for initial genotyping. Twenty recombination events were detected in the analysis of 1500 gametes. After two generation self-pollinated, the F4 progenies were used to characterize the recombinants (Figure 12).

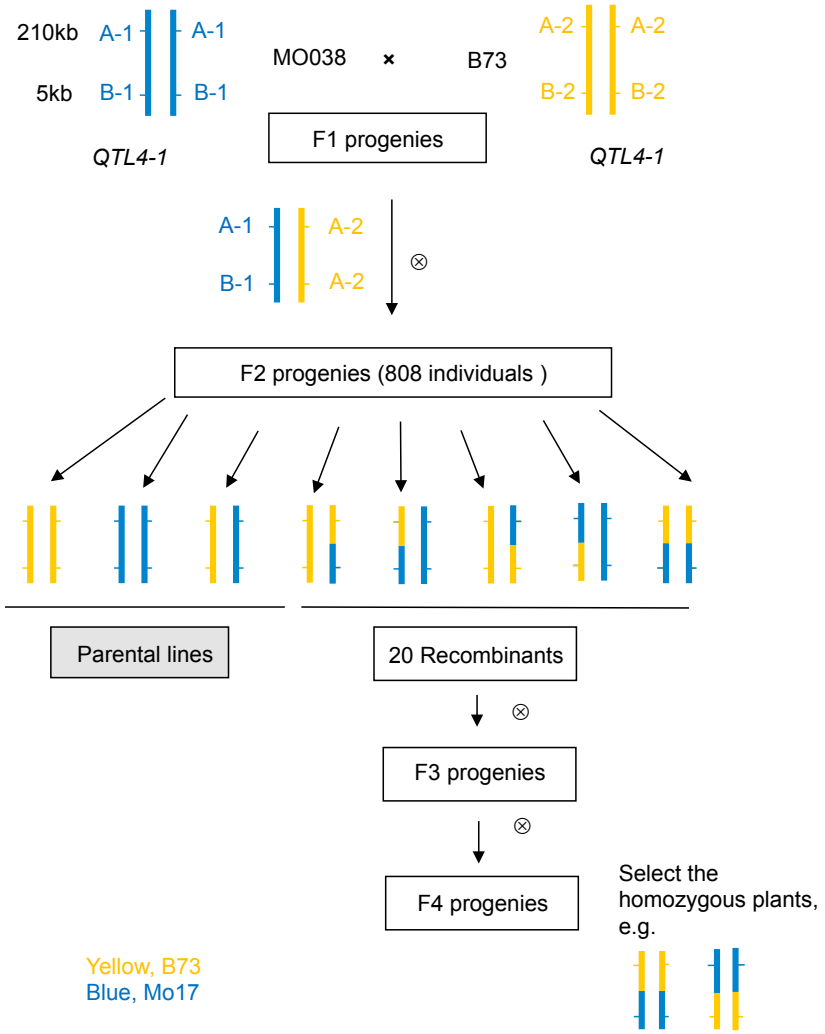


Figure 12: Fine mapping of *QTL4-1*. The inbred line MO038 was crossed with B73. Recombinants were screened in the resulting F2 and 20 recombinants of 5 different types were detected. Homozygous plants of recombinants were selected in the F4 progenies. The B73 like region is indicated in yellow color; Mo17 like region is indicated in blue color.

3.3.2.3 Genotype and phenotype of the recombinants

Marker 5kb, 10kb, 32kb, 43kb, 83kb, 137kb, αβ marker, 143kb, 148kb, and 210kb were used to genotype all the recombinants (Table 7). Two recombination hotspots were revealed. The borders are given by the markers 210kb and 148kb, 143kb and αβ. In the first recombination hotspot, there are six recombination events within 40 kb; 13 recombination events are in the second recombination hotspot within the 6 kb region (Table 8, Figure 13). There is no recombination event in the 135 kb between the αβ marker and 5kb.

The recombinants can be arranged into six groups (Figure 13). Roughly generalizing, group 1, group 2, and group 3 have B73 genotype at *Bx1* and up to different positions around the αβ marker. The genotype further upstream (towards the tip of the chromosome) is Mo17. Group 4, group 5 and group 6 have the inverse arrangement.

	Marker →	IDP8302	210kb	148kb	αβ marker				137kb	83kb	43kb	10kb	5kb	*Bx1*-gene
	Genes →		*Bx4, Bx5*		α	β	α	β			*Bx8*			*Bx1*
	Pos. (bp) →	2.424.753	3.045.994	3.108.322	3.112.492- 3.116.365				3.119.173	3.173.538	3.214.204	3.246.637	3.251.017	3.256.755
Group # ↓	Recomb. Lines ↓													
1	772													
	794													
	676-1*													
2	662													
	512													
	533													
	604													
3	543													
	325													
	676-2*													
4	451													
	729													
	329													
	259													
	92													
5	97													
	109													
	372													
	497													
6	222													

Figure 13: Genotype of recombinants. The recombinants are arranged in groups (#1-6). The number of individual recombinant lines is indicated. The position of the marker (Pos.) is given in the header based on the B73 v2 sequence, and the genes found in the region are indicated. The color code indicates Mo17 (blue), B73 (yellow) and hybrid (green) genotype. White color indicates the genotypes that not confirmed by analysis. Recombinant 676* is a double recombinant. It gives rise to two homologous recombinants (676-1* and 676-2*), which belong to two groups.

All recombinants were further propagated by selfing and crossing with the parental lines. Homozygous and heterozygous lines with respect to the *Bx1* allele were established. DIMBOA-content and *Bx1* expression were determined for both types of plants 24 dai; the heterozygous plants were analyzed additionally for allele-specific expression (Table 8).

Table 8: *Bx1* transcript levels in hybrids of recombinants.

Group	Recombinants #	*Bx1* allele of recombinant parent	Parental line in the cross	Allele expression in hybrid	*Bx1* expression of homozygous recombinants
Group 1	772, 794, 676-1	B73	Mo17	Mo17	low
Group 2	662, 512, 604, 533	B73	Mo17	Mo17	low
Group 3	325, 543, 676-2	B73	Mo17	Mo17	low
Group 4	451, 729, 92, 329, 259	Mo17	B73	Mo17	high
Group 5	97, 109,372, 497	Mo17	B73	Mo17	high
Group 6	222	Mo17	B73	Mo17 (decreased)	Medium

Among all group 4 and group 5 have the highest DIMBOA-content and *Bx1* expression level. Group 6 (line 222) has relatively low DIMBOA-content, low *Bx1* expression and reduced allele specific expression. The groups 1 and 3 to 5 consist of at least 3 independent recombinants that were crossed several times. Group 2 represents recombinations within the αβ-region joining upstream Mo17-sequences and downstream B73 sequences; in one recombinant α-2 and β-2 are joined, three recombinants have α-2/β-1 conformation (Figure 14). There is no significant difference in DIMBOA-content and late *Bx1* gene expression in these different types of group 2 recombinants. All of them are low with respect to these values. The original recombinants differ in the genotype outside of the investigated region. Since different lines with the same configuration within the *QTL4-1* region are available and several generations of backcross were applied, it can be supposed that unlinked regions influencing DIMBOA-content and *Bx1* gene expression will be randomly found in the progeny of the groups used for analysis.

Differences in DIMBOA-content are less pronounced compared to *Bx1* transcript levels. High DIMBOA levels are recorded for group 5 representing 4 individual recombinants. The presence of the αβ duplicate and downstream Mo17 sequences seem to be required but not sufficient for generation of high DIMBOA levels. *Bx1* transcript levels of group 4 and 5 are up to 10 times higher compared to groups 1, 2, and 3.

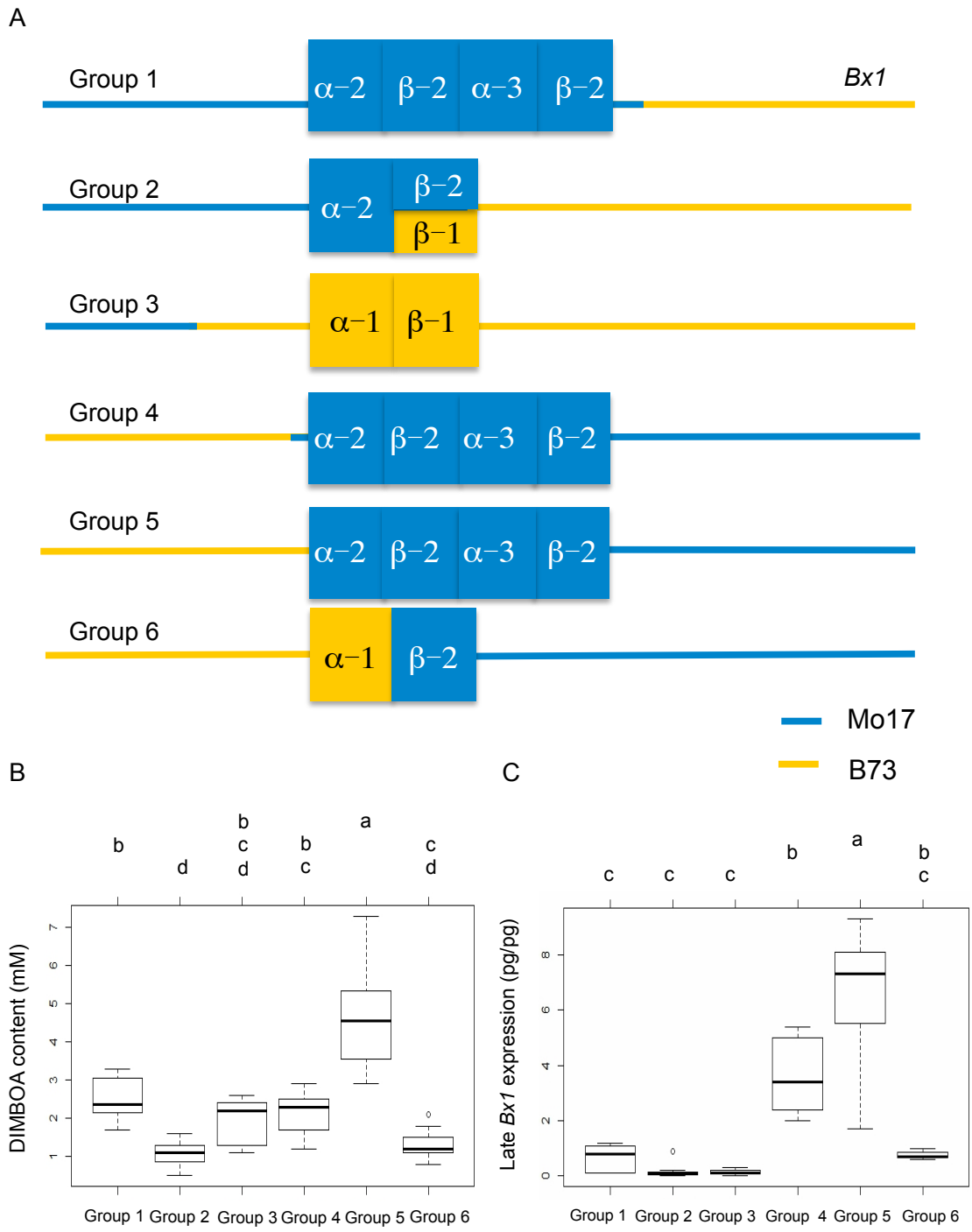


Figure 14: Genotype and Phenotype of 20 recombinants. A: The recombinants are grouped and breakpoints are color-coded as in Figure 13 (Mo17 blue, B73 yellow). Homozygous F4 progenies are analyzed. B, C: DIMBOA and *Bx1* expression level of 24 dai plants. The average *Bx1* expression level for B73 is 0.3, with standard deviation 0.1; for Mo17 the average level is 8.4, with standard deviation 5.5. Three to eleven biological replicates were analyzed for each genotype group. The results were statistically compared using ANOVA Tukey HSD test (α=0.05). The DIMBOA concentration is given in relation to the fresh weight. The gene expression of the *Bx* genes is normalized to *GAPDH*. (pg/pg, normalized to *GAPDH*).

This indicates that sequence elements downstream of the αβ region also influence the transcript level. The finding that in group 6 (222), the *Bx1* transcript level was decreased although Mo17 downstream sequence were present, indicates that the αβ duplication, which is missing in this recombinant, had great influence on *Bx1* transcription.

There may be a negative regulator for DIMBOA-content and *Bx1* transcript level upstream of the αβ-region and limited by the 210kb marker, because group 5 is higher in both respects compared to group 4. Group 6 has no αβ duplication, but the α-1 allele is connected to β-*2* (α-1/ β-*2*) and Mo17 downstream sequences (Figure 14), hence the recombination took place within the αβ-duplication. This can be taken as an indication that α–2 /β–2/ α–3/ β–2 structure in Mo17 is not only a recombination hotspot, but can also influences *Bx1* expression.

In order to define the *cis*-elements required for late allele specific expression, hybrid Mo17 and B73 genotypes with respect to *Bx1* were generated, by crossing of the recombinants with the respective parent line. Then the expression levels of the two alleles were determined. Only low late expression of the B73 *Bx1* allele of the hybrid was detected in recombinant groups 1-3. This demonstrates that neither the αβ duplicate nor sequences upstream thereof are sufficient to confer allele-specific expression. Allele specific expression of group 5 proves that no sequences upstream of the αβ duplicate are *cis*-active. Evidence for the requirement of the αβ-duplication for late expression of the *Bx1* gene located in *cis* comes from analysis of hybrids of the recombinant 222 (group 6) and the IBM line MO057 (see Chapter 3.6). Both lines are B73 in the αβ-duplication and Mo17 downstream including the *Bx1* gene. In both hybrids the expression of the Mo17 allele of *Bx1* is significantly reduced. These data indicate that αβ duplicate is necessary for maximum *Bx1* expression.

In summary, both the αβ–region and downstream sequences were important for *Bx1* expression level and *Bx1* allele specific expression, and regions upstream thereof do not increase the *Bx1* expression level in late plants. The genotype and phenotype of recombinants suggest that the αβ–region influences also the DIMBOA-content. The downstream limit for the *cis*-elements could not be determined by the analysis of the recombinants and was investigated using selected near-isogenic lines (NILs) and IBM302 lines (see 3.4 and 3.6).

3.4 Characterization of near-isogenic lines (NILs)

To assay the contribution of the αβ–region and downstream sequences to *Bx1* expression in a fixed genomic background, near-isogenic lines (NILs) were used. B73 and Mo17 were used by Nathan Springer lab to produce a set of NILs with small regions of introgression into both backgrounds (Eichten *et al.* 2011).

B184 and M31 are NILs of this collection that carry the introgression of the *QTL4-1* region. B184 is a B73-like NIL, containing 0 bp- 3.811.907 bp Mo17 sequence on the short arm of Chromosome 4. B184 also contains some Mo17 fragments on chromosome 1, chromosome 2, and chromosome 3. M31 is a Mo17- like NIL, containing from position 2.829.191 bp to

10.124.409 bp B73 sequence on the short arm of Chromosome 4 (Figure 15), and additionally small introgressions of B73 on chromosome 1 (not overlapping with the chromosome 1 region in B184), chromosome 3, and chromosome 7. DIMBOA-content, *Bx1* expression level and allele specific expression were analyzed in both B184 and M31.

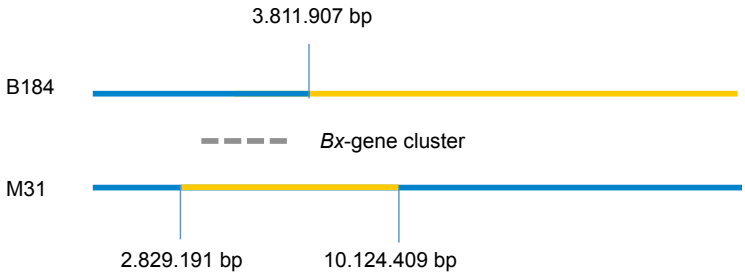


Figure 15: Genotype of B184 and M31 on chromosome 4. The dashed line indicates the *Bx gene* cluster. Mo17 genotype is indicated in blue, B73 genotype is indicated in yellow. The borders indicate recombination sites. The genotype downstream of the marker is given. The αβ-region is located at 3.115.056 bp- 3.115.579 bp, *Bx1* is located at position 3.256.736 bp – 3.256.843 bp.

Compared to B73, B184 has relatively high late DIMBOA-content and *Bx1* gene expression, but both values are not as high as in Mo17 (Figure 16). The NIL line M31 has a high DIMBOA-content, however, it has an extremely low *Bx1* expression level, even as low as B73. These data indicate that *QTL4-1* is important for a high *Bx1* expression level, but there are also further factors influencing the DIMBOA-content.

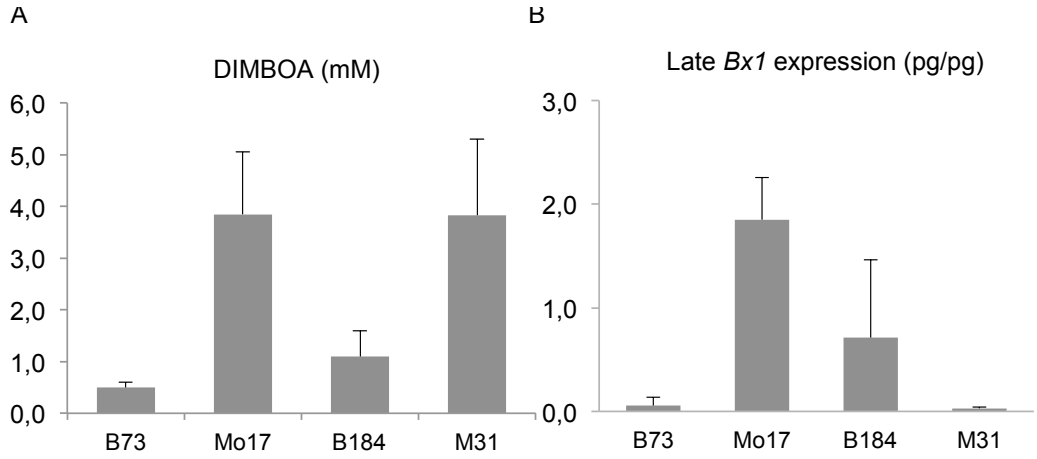


Figure 16: DIMBOA-content (A) and *Bx1* expression level (B) in 24 dai plants of NILs and parental lines. 3-4 biological replicates were analyzed for each line. The DIMBOA concentration is given in relation to the fresh weight. The gene expression of the *Bx* genes is normalized to *GAPDH*. (pg/pg, normalized to *GAPDH*).

To investigate the contribution of the αβ-region in the regulation of *Bx1* expression, progeny of B184 and Mo17 crosses were analyzed (Figure 17). B184 is Mo17-like in the *Bx* cluster region, which is hence present homozygous in the progeny of the cross. The late *Bx1* expression level in the hybrids (B184×Mo17 and Mo17×B184) is about twice as high as the

level in the B73×Mo17 and Mo17×B73 plants, but lower than 1/2 of the Mo17 expression level. Hence the B73 genome contributes factor(s) that suppress the *Bx1* expression of Mo17 allele in the hybrids (Figure 17).

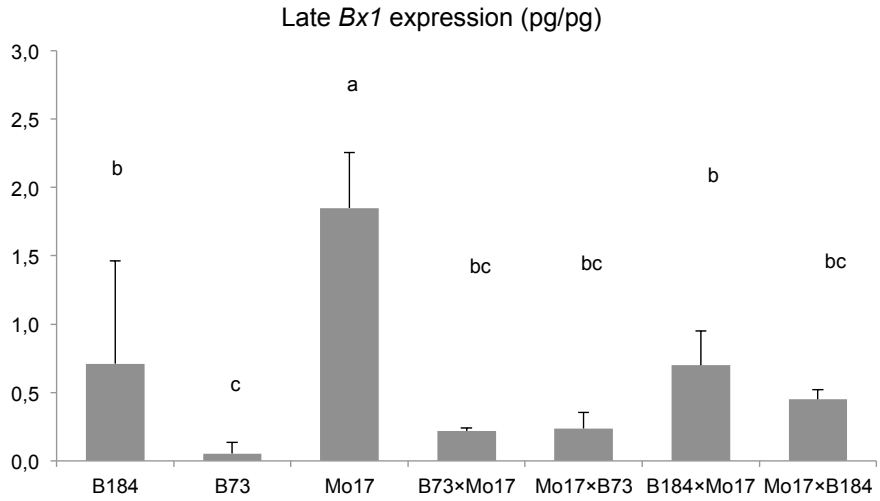


Figure 17: Characterization of B184/Mo17 hybrids. *Bx1* expression levels were determined for B184, hybrids and control lines in three biological replicates. Elevated *Bx1* expression levels are detected in B184×Mo17 and Mo17×B184, compared to B73×Mo17 and Mo17×B73. The results were statistically compared using ANOVA Tukey HSD test (α=0.05). All *Bx1* expression values are normalized to *GAPDH*.

The hybrids generated by crossing with B73 behave analogously. The hybrids B73×B184 have *Bx1* expression levels slightly but not significantly different to B73×Mo17 (Figure 18 A). Analysis of the allele specific expression level reveals that only the Mo17 allele derived from B184 in this case is expressed (Figure 18 B). These data further verify that B184 contains *cis*-elements influencing *Bx1* expression in 24 dai plants.

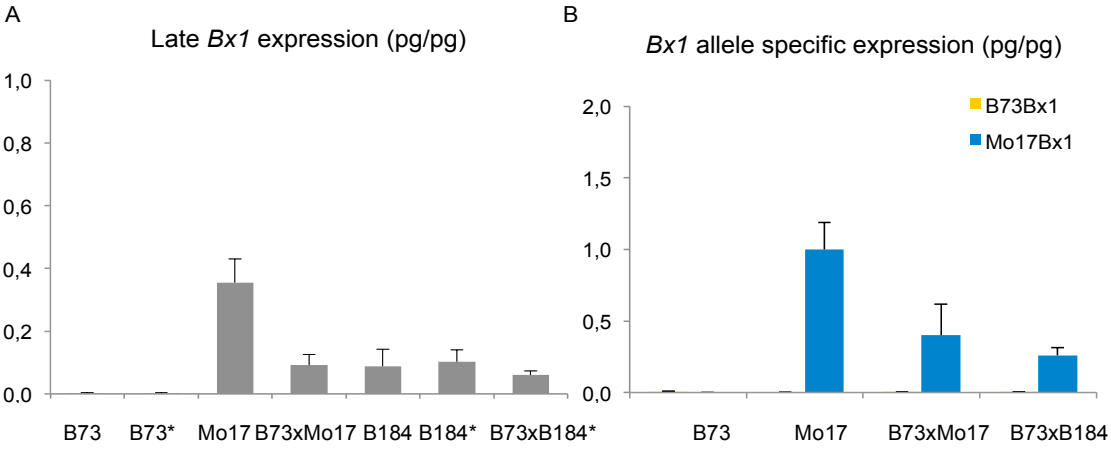


Figure 18: Characterization of B184/B73 hybrids. A: *Bx1* expression in B184, B73, Mo17 and hybrids. The * indicates segregating progeny of one maize cob (self-pollinated B73×B184). B: Allele specific expression of *Bx1* in B73×B184 and control lines. All values are normalized to *GAPDH*. The results in A were statistically compared by ANOVA Tukey HSD test (α=0.05).

In summary, the data demonstrated that the Mo17 sequence in B184 is sufficient to confer the allele specific *Bx1* gene expression. The B73 genome contribution of the B184 genome, on the other hand, includes the factors to decrease the *Bx1* transcript level. Hence, negative factor(s) of B73 genome are located outside of the introgressed genomic region in B184.

3.5 QTL analysis with the sub-population of IBM 302 population

The results from the recombinants and the NIL lines indicate that Mo17-conformation at *QTL4-1*, that means the αβ–region duplication and downstream sequences, are necessary for the high *Bx1* gene expression. Effects of *trans* regulatory factors will not be detected if B73 alleles are present in this region. Hence, a sub-population of IBM302, consisting of lines that are Mo17 at QTL4, was selected for further QTL analysis based on the original data for the RILs (Table in attachment). Composite interval mapping (CIM) was done by Windows QTL cartographer version 2.5 to map QTLs for the selected sub-population of IBM lines, using DIMBOA-content as phenotype.

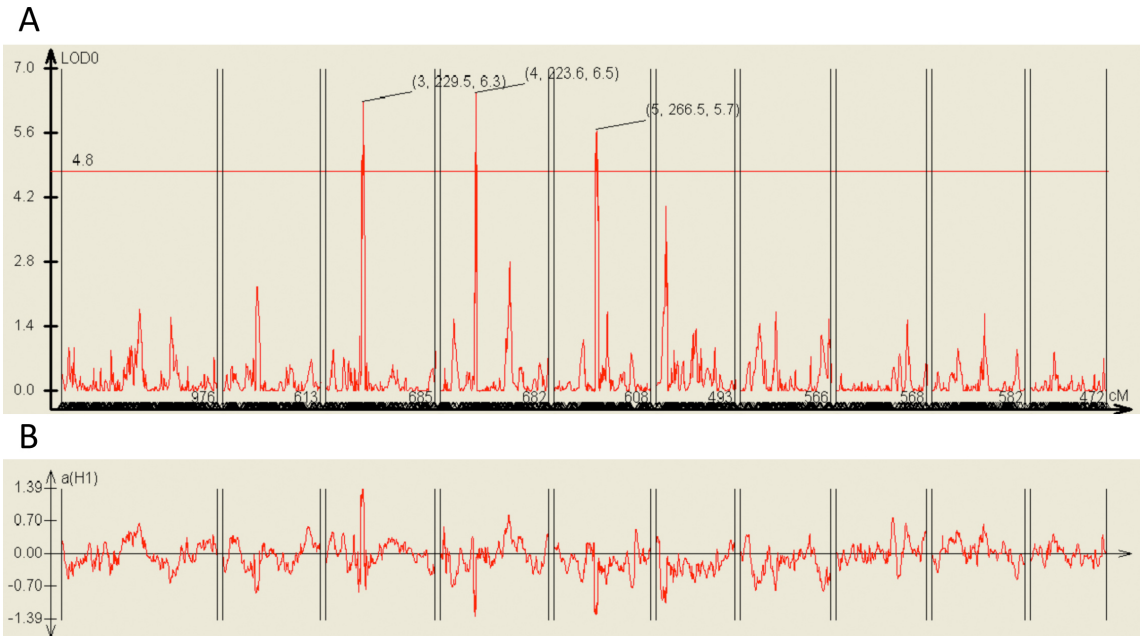


Figure 19: QTL analysis with a sub-population of IBM lines. Recombinant inbred lines that are Mo17 genotype at *QTL4-1* were selected. A: LOD scores for the sub-population evaluated for DIMBOA-content. Three values are above the threshold of 4.8 in 1000 permutations. B: Additive effect of the parental genomes. Values above zero indicate B73 contribution; values below zero indicate contribution by Mo17.Three QTLs, *sub-QTL4, sub-QTL3* and *sub-QTL5,* were detected with the sub-population by 1000 permutation. *sub-QTL4* is the only QTL with positive contribution by the B73 genome.

Three so-called sub-QTLs (sQTL) were detected that were stable in 1000 permutation. *sub-QTL4* co-localized with *QTL4-2*, the mini QTL from the complete IBM302 population analysis. Its LOD value goes up from 2.7 to 6.5, and the QTL can explain 18% of the phenotype variation. *sub-QTL3* and *sub-QTL5* explain 5.6% and 16.4% of the phenotype

variation separately (Figure 20, Table 9 in appendix). Only for *sub-QTL3* the positive contribution is from the B73 genome. The conformation at the sub-QTLs was taken into account in the analysis of selected lines (see 3.6).

3.6 Characterization and fine mapping with IBM lines

To increase the number of the lines with recombination points within the *QTL4-1* interval, we selected lines of IBM302 population based on the genotypes available from the database (MO038, MO141, MO024, MO035, MO067, MO276, MO033, MO161, MO331, MO057, MO061, MO076, MO039 and MO382).

The selected IBM RILs, the recombinants from the fine mapping, and the parental lines B73, Mo17 and MO038 were comprehensively genotyped in the *QTL4-1* interval and characterized for *Bx1* expression. As expected, all lines with Mo17 genotype at *QTL4-1* (MO038, MO141, MO024, MO035, MO067, MO276 and MO033) are higher in late *Bx1* expression compared to B73. All the others contain low *Bx1* transcript levels in 24 dai plants (Figure 20). MO057 can be added to group 6 of the recombinants that is represented by the recombinant 222 (Table 3, Figure 16A), both lines have no duplicate of the αβ-region but the recombinant conformation α-1/β-2 and Mo17 downstream sequences. MO057 has also lower *Bx1* expression level compare to MO038, MO141, MO024, MO067, MO276 and MO033, and moderate high expression level compared to MO161, MO331, MO076 and MO111. This finding further verifies that α−2/β−2/α−3/β−2 sequence in Mo17 has great influence on *Bx1* expression level in late plants.

The IBM lines' genotype and phenotype in detail can give a hint for the contributions of the sub-QTLs (Figure 20). Among the analyzed IBM lines, 7 lines (MO038, MO141, MO024, MO035, MO067, MO276, and MO033) belong to the sub-population for QTL analysis in Chapter 3.5. Of these MO038 and MO035 represent the two extremely high lines, and they are both Mo17 genotype at *sQTL4*. However, the *sub*-QTLs cannot easily explain (Figure 20B) the phenotypic variation of all the lines.

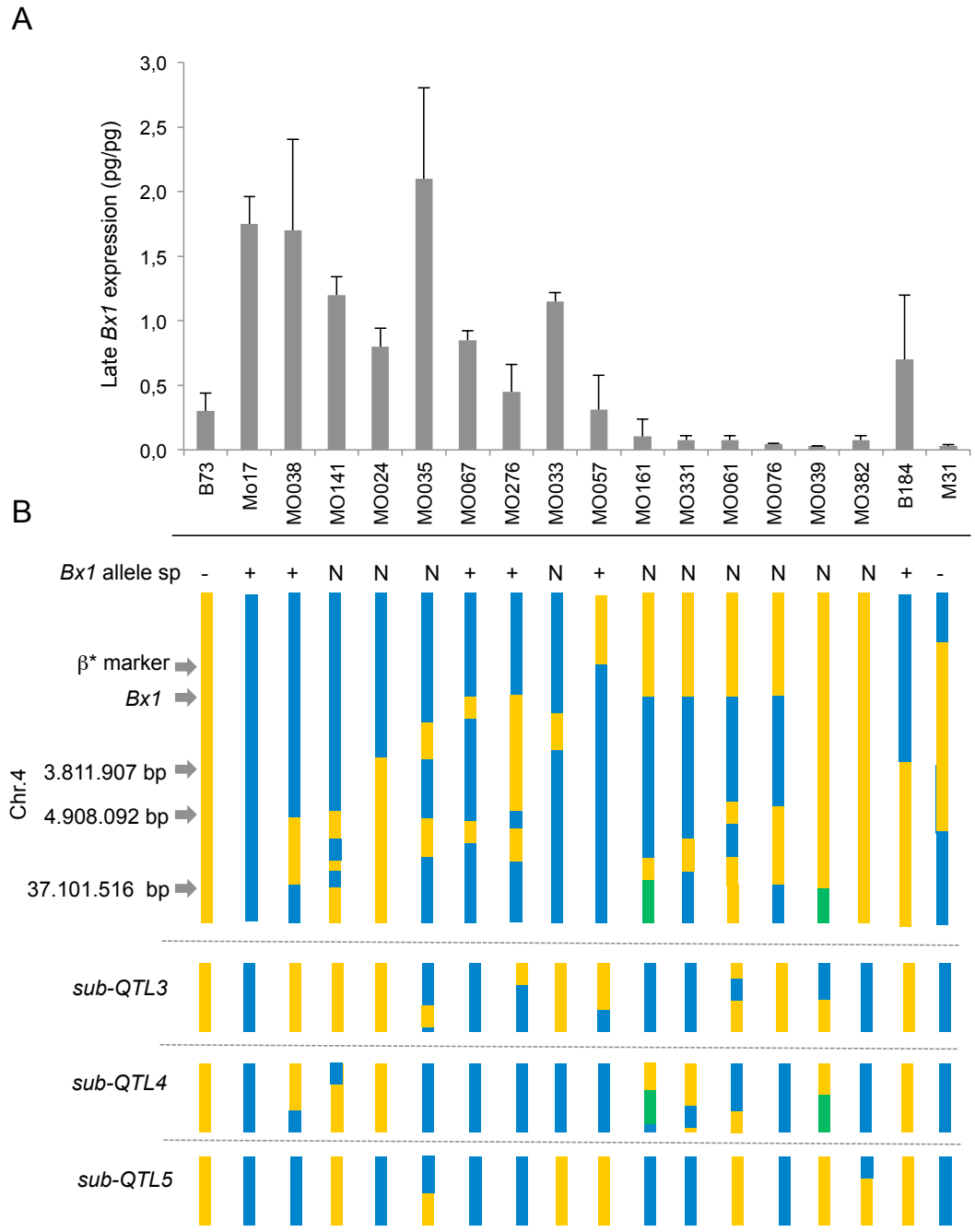


Figure 20: *Bx1* expression analysis and genotype of selected IBM lines. A: Relative *Bx1* expression level in 24 dai plants (pg/pg, normalized to *GAPDH*) B: Genotype of selected IBM lines. sQTLs from the sub-population analysis are included. Genotype Mo17 blue, B73 yellow, hybrid green. The genotype of the marker and downstream sequences is given. *Bx1* allele specific expression is indicated: + expression like Mo17; - allele expression like B73; N allele specific expression not analyzed.

It should be noted that the QTL analysis was based on DIMBOA-content and that the analysis has shown that there is no absolute correlation between DIMBOA-content and *Bx1* gene expression. Hence essential additional QTLs for late *Bx1* expression might have not been detected.

For selected IBM lines allele specific expression levels were analyzed in crosses with Mo17 or B73 (Figure 21). As mentioned before the sum of the allele specific transcript levels and the result for the general primers are slightly different. The relative differences of allele-specific expression in the lines, nevertheless, are significant. MO067 and MO276 have the B73 *Bx1* allele, and the Mo17 genotype upstream of the *Bx1* gene, and share common recombination breakpoints between 2kb and *Bx1* gene markers. Intriguingly, in hybrids of the two recombinants crossed with Mo17, the B73 allele and the Mo17 allele both are equally and highly expressed. Hence the *cis*-elements are present on the recombinant chromosomes.

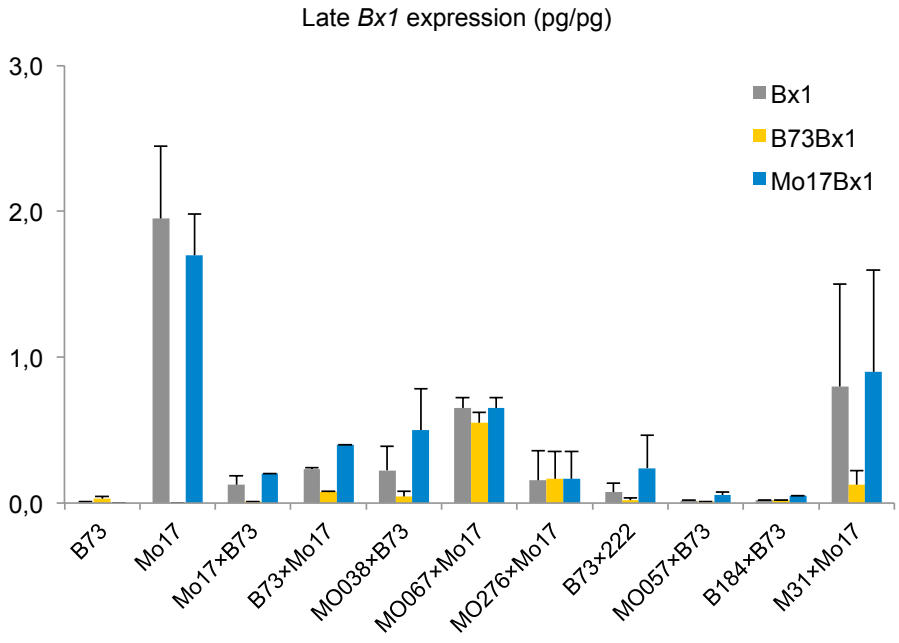


Figure 21: Allele specific expression in hybrids of IBM lines and recombinants. To get heterozygous conformation for the *Bx1* gene, parental lines and recombinants were crossed with either B73 or Mo17, as indicated. Two independent biological replicates were analyzed for total *Bx1* (grey column) and allele-specific expression (Mo17 blue, B73 yellow columns). All values were normalized to *GAPDH*.

Differences of the absolute values in the experiments (that cause the standard deviation) might be due to growth differences and experimental error in determination of absolute RNA levels. Both MO276 and MO067 have equally Mo17-conformation from the αβ-region to the *Bx1* locus. In each individual experiment, both alleles are identically expressed. However, late *Bx1* gene levels are different in the lines and this might be attributed to negative *trans*-factors contributed differentially by the RIL parents in the hybrids. Interestingly, the expression level of both alleles in MO067 is higher than the Mo17-allele expression in the Mo17/B73 hybrids.

MO067 and MO276 have B73 genotype from the breakpoint upstream of the *Bx1* gene downwards to position 3.477.883 and 4.908.092 (v2), respectively. The NIL B184 is B73 downward of position 3.811.907 (v2). All three lines demonstrate allele-specific expression

for *Bx1* (Figure 18, Figure 21). These results taken together delimit the major *cis*-element necessary for allele-specific *Bx1* expression between the αβ-region and *Bx1* locus.

3.7 Determination of allele specific expression for *Bx2*

Bx1 and *Bx2* are located in only 2.5 kb distance. Both genes have a similar expression pattern in different developmental stages in B73 (Figure 5B). The allele specific expression found for late *Bx1* in Mo17 might be shared by *Bx2*.

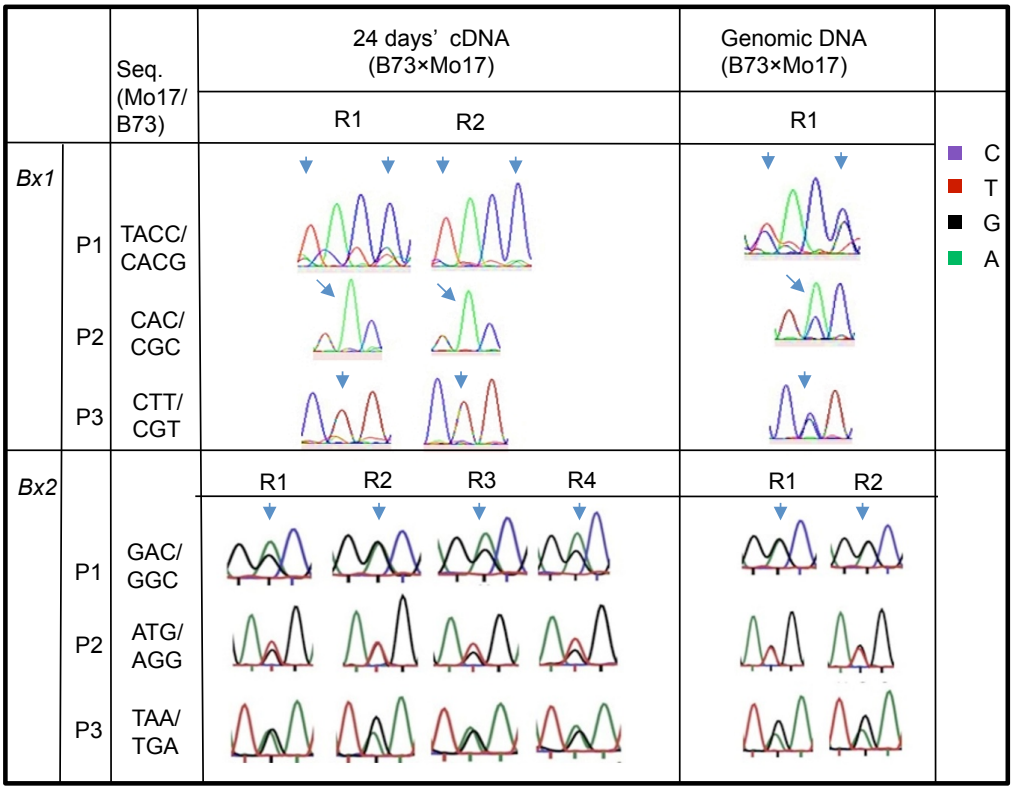


Figure 22: Analysis of *Bx1* and *Bx2* allele specific expression by sequencing. SNPs in three different positions (P1, P2, and P3) were selected for *Bx1* and *Bx2*, amplified and sequenced with primers fitting both alleles. Genomic DNA served as control. Two and 4 biological replicates were analyzed. The arrows point to the SNP in the sequencing chromatogram.

There is no sequence difference that would allow *Bx2* allele specific cDNA amplification and the possibility to determine allele specific values by qPCR. Alternatively, an analysis by sequencing of cDNA amplification was performed. Contamination by genomic DNA is excluded since intron sequences will lead to a fragment of larger size. The allele specific expression of *Bx1* and *Bx2* was analyzed in parallel and genomic DNA of hybrids was used as a control (Figure 22). Only the sequence of the Mo17 allele is present in the sequencing of the hybrid cDNA of 24 dai leaves in the case of *Bx1*. This result substantiates allele-specific analysis by RT-qPCR (Figure 6F). However, presents of both *Bx2* alleles in the cDNA are demonstrated in the sequence analysis. The result is similar to the analysis of genomic DNA

of the hybrids. Hence both alleles contribute equally to the expression of *Bx2*. Even though *Bx1* and *Bx2* are separated by only 2.5 kb, this analysis shows that the regulation of these two *Bx*-genes is independent.

3.8 Presence of the αβ-duplication in the European NAM population

The European NAM population ("EuNAM" population) was used to determine the allele frequency of the αβ-duplicate in maize. Twentyone EuNAM-lines were analyzed with AlphaF and AlphaR primers and modified primers (F18F2 and R18R2, Table 3 in Material and Methods; Bachelor thesis, David Dietl, 2012). Only in one line, FF0721H, a αβ-region with duplication structure like in Mo17 in EuNAM is detected. The exact sequence of the region was not determined. FF0721H has a low late DIMBOA-content (Bachelor thesis David Dietl, Figure 5). And the late *Bx1* expression level of this line is very low (Figure 23).

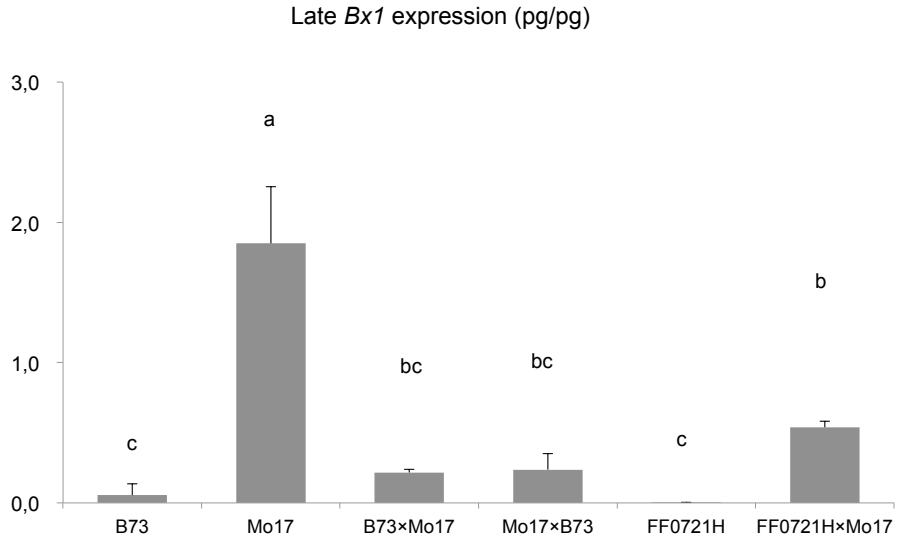


Figure 23: Characterization of the inbred line FF0721H. *Bx1* expression was analyzed in FF0721H, hybrids and control lines. Moderately elevated *Bx1* expression levels were detected in FF0721×Mo17 compared to B73×Mo17 and Mo17×B73. The results were statistically compared using Tukey HSD test (α=0.05). Relative *Bx1* expression was normalized to *GAPDH*. The data were collected parallel to the data set displayed in Figure 17.

To investigate the activity of αβ-duplicate in the regulation of *Bx1* expression, FF0721H was crossed with Mo17. Elevated *Bx1* expression level can be detected in FF0721H×Mo17 compared with the two hybrids B73×Mo17 and Mo17×B73 in 24 dai plants (Figure 23). This can be explained by positive *trans*-factor(s) contributed by Mo17 that increase expression of the FF0721H *Bx1* allele, maybe *via* the αβ-duplicate. Another possibility is that there is no negative *trans*-factor in the FF0721H genome that reduces Mo17 allele expression in Mo17/B73 crosses. The two alternatives can be distinguished by allele specific analysis. There is no sequence difference that could help to design qPCR primers that allow *Bx1* allele specific cDNA amplification and the possibility to determine allele specific values. A SNP between FF0721H and B73 was provided by Prof. C. -C. Schön's lab. An analysis by

sequencing of cDNA amplification analogous to the analysis of *Bx2* (Chapter 3.7, Figure 22) was also performed. Genomic DNA of hybrids was used as control. The result indicated that most of the transcription is contributed by the Mo17 allele in the hybrid cDNA of 24 dai leaves (Figure 24). This indicates that the negative *trans*-factor that reduces Mo17 allele expression in Mo17/B73 crosses is absent in the FF0721H genome.

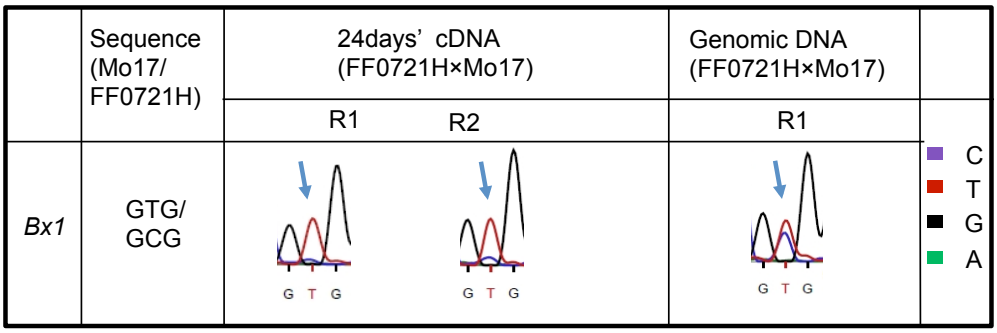


Figure 24: Analysis of *Bx1* allele specific expression in FF0721H×Mo17 by sequencing. One SNP in between FF0721H and Mo17 is selected for *Bx1* allele specific analysis. Two biological replicates were analyzed for the cDNA of 24 dai plants. The arrows point to the SNP in the sequencing chromatogram.

In conclusion, the αβ-duplicate is a rare conformation in both American and European maize lines. FF0721H is the only one with the duplicate in the EuNAM. Nevertheless, its *Bx1* expression in 24 dai leaf is low. Hence the duplication alone is not sufficient for elevated late *Bx1* expression as already seen in the NAM germplasm. Clearly the *trans*-factors contributed by B73 are missing in FF0721H.

4. Discussion

4.1 Regulation of benzoxazinoid content and *Bx*-gene expression in maize inbred lines and hybrids can be uncoupled

High DIMBOA-content is a juvenile trait. Accordantly, expression of the benzoxazinoid biosynthetic pathway genes is high in seedlings. The investigations showed that there is variability in DIMBOA levels at older stages within maize. The NAM founder line panel showed that DIMBOA-content in older plants (24 dai, leaf 3) is variable. The inbreds Ky21, M162, MS71, M37W and Mo17 have relatively high DIMBOA-content at this stage. In Mo17, the consistently high late DIMBOA level is connected with significantly elevated *Bx1* expression levels. However, Mo17 is unique within the NAM panel with respect to this feature. *Bx1* is the only gene that differs significantly in expression between Mo17 and the inbred B73 that has low late DIMBOA levels. An increase of *Bx1* transcript level was consistently detected between day 14 and 21 in Mo17 while a steady decrease takes place in B73. This result indicates that in Mo17 transcript reduction is not merely decelerated but *Bx1* transcription is reactivated and reaches levels similar to the seedling stage. Hence it was speculated that in Mo17 this prolonged high expression of the branchpoint gene of benzoxazinoids biosynthesis is the driving force of high late DIMBOA levels. Other mechanisms might be relevant in the case of the other late high DIMBOA inbred lines. The analysis of recombinants between Mo17 and B73 (Figure 14) showed that *Bx1* expression levels and DIMBOA-content are not strictly correlated. Uncoupling of these two traits is most obvious in the analysis of the NIL line M31 that has Mo17 genotype except for the tip of chromosome 4 that includes the *Bx* gene cluster. High late DIMBOA-content is found in this line although *Bx1* expression is very low (Figure 16). This finding reveals that even in Mo17 *Bx1* gene expression is not the only factor that has major effect on DIMBOA level. One possibility to explain this finding is that *Igl*, the herbivore inducible maize indole glycerol-phosphate lyase that has analogous enzymatic properties to *Bx1*, is up regulated in these lines. *Igl* transcript levels were analyzed for Mo17, B73, CML322, CML228, M162W, Ms71 and Ky21, and no correlation was revealed (data not shown). Besides biosynthesis, transport, accumulation in storage compartments, and degradation could further influence the DIMBOA concentration and the actual DIMBOA concentrations is a function of the rate of formation and the rate of decay. Meihls *et al.* (2013) identified three *O*-methyltransferases (*Bx10a-c*) that could convert DIMBOA-Glc to HDMBOA-Glc. Variation among the NAM lines concerning HDIMBOA-Glc content was analyzed by the authors. Most of the high DIMBOA-Glc lines in this study, e.g. Ky21, M162W, M37W, Mo17 and B97, which are also high DIMBOA lines in our data analysis, have quite low HDMBOA-Glc content in 2 weeks' leaf (Figure 3.1, Result). The negative correlation between these two benzoxazinoid species in the inbred lines leads to the speculation that high DIMBOA levels can be caused by reduced conversion to HDMBOA and further metabolization. However, although low *Bx1* gene expression does not exclude high DIMBOA, in our analysis no inbred or recombinant line was detected that has high *Bx1* transcript levels combined with low DIMBOA-content.

4.2 High *Bx1* gene expression in Mo17 is connected to *cis*-elements

A sequence element located about 140 kb upstream of *Bx1* is altered in Mo17 compared to B73. About 3.6 kb of non-repetitive DNA is duplicated. Formally the sequence can be divided into two parts. The first (α-) part (2628 bp) is greatly altered in the duplicate (overall 81% identity). The second (β-) part differs slightly and the altered version is found in both Mo17 duplicates (Figure 10). Duplication of the αβ-sequence element is rare in maize, within the NAM founder population it was detected only in Mo17, and within a diverse European population panel of 21 inbreds (EuNAM) an analogous duplication was again detected only in one line. The presence of the duplication is recognized and tested by sequence analysis of PCR-fragments amplified on genomic DNA templates; here the sequence differences in the fragment mixture lead to a defined break of the readable sequence due to double labeling. If exact copies of the region exist, these will not be detected.

Allele-specific high late *Bx1* expression was consistently detected by RT-qPCR (Figures 6, 18, 21, 24) and verified independently by sequencing (Figure 24). The comprehensive analysis of recombinant inbred lines and NIL lines based on Mo17 and B73 (Figure 14C, Figure 18B, and Figure 21A, B) revealed that the αβ-duplication of Mo17 is necessary for elevated late *Bx1* gene expression. In addition downstream sequence elements are required that could be located to the region between *Bx1* gene and the αβ region, comprising about 140 kb (Figure 14C, Table 8). Two recombinants, MO067 and MO276 (Figure 21A, B) demonstrate that the Mo17 allele of *Bx1* is not necessary for high late *Bx1* expression.

No recombinants were detected in this stretch of DNA in the IBM302 population and the mapping population generated in this work. Hence a further fine mapping of the second sequence component required for high late *Bx1* expression was not possible. At present there is no indication about the location and nature of the additional sequence element(s) required for high late and allele-specific *Bx1* expression. No high quality sequence data that would allow recognizing sequence duplications and alterations like the ones found for the αβ-sequence, are publically available for Mo17 in this region. No special features are displayed in the B73 sequence. Beside *Bx8* 16 gene models are annotated in the region, most of them are small (100 to 300 bp) and have no cDNA or EST support. The only exceptions are the gene models GRMZM2G085303 and GRMZM5G856095, where the latter represents the reverse complement of 380 bp at the 5'-terminus of the first transcript. The possibly encoded protein has similarity with a putative maize F-box containing protein but not with the F-box domain. Since the sequence is moderate repetitive, a comparative analysis in Mo17 and B73 based on PCR amplification was not possible and it is not clear whether this sequence element is conserved in Mo17.

In a survey for nucleosome occupancy within gene- and gene proximal- regions of maize classical genes (Fincher *et al.*, 2013) a significant deviation from expected pattern was detected about 600 bp upstream of the *Bx1* transcription start site for seedlings but not for adult leaves. Hence the detected loose chromatin conformation coincides with high *Bx1* expression. A stretch of 114 bp including this region was genotyped for Mo17, B73 and some

European inbred lines (David Dietl, Bachelor's Thesis, Technische Universität München, Freising, 2012). Five SNPs defining five haplotypes were detected. Mo17 and B73 differ with respect to a central dinucleotide pair. Whether this difference has an effect on gene expression is unclear; none of the recombinants separates this region and the *Bx1* gene.

Beside *Bx1* no other *Bx* genes of the cluster is significantly changed in expression level when B73 and Mo17 are compared. This is unexpected since *Bx5* is in close proximity to the αβ-region, *Bx8* is located between the αβ-region and *Bx1*, and *Bx2* is found only 2.5 kb downstream of *Bx1*. For *Bx2*, it was possible to demonstrate that not only the absolute transcript level is as low as in B73 but also both alleles are expressed at the same level in the hybrid in seedling and late plants. Hence the effect of the *cis*-elements has a sharp 5'-limit. Such an analysis was not possible for *Bx8* and *Bx5* since suitable SNPs are missing.

4.3 Expression of the Mo17 allele of *Bx1* is reduced by a *trans*-factor present in B73

The *Bx1* expression level of the Mo17-allele is repressed in both reciprocal B73/Mo17 hybrids, since it is clearly below the half value of Mo17 (Figures 6E). In the hybrids with the NIL-lines B184 and Mo17, the αβ duplicate and downstream region are homozygous like in Mo17. Here, the *Bx1* expression level in both reciprocal hybrids is higher than in the respective B73/Mo17 hybrids that have only one Mo17 *Bx1* allele, but even the half value of the homozygous Mo17 expression level is not reached. This leads to the speculation that the late *Bx1* expression is negatively influenced by one or more *trans*-factors contributed by B73. MO276 ×Mo17 and MO067×Mo17 hybrids both have the Mo17 αβ duplication to *Bx1* upstream region in homozygous conformation in connection with a Mo17 *Bx1* allele (from the Mo17 parent) and a B73 *Bx1* allele from the recombinant. Expression of both *Bx1* alleles is identical within one hybrid and significantly higher than the ones of the B73 allele in Mo17/B73 hybrids. Again the expression values are less than half of the Mo17 inbred level, and this reduction that can be ascribed to the postulated negative *trans*-factor(s) contributed by the B73 genome. The pattern of B73 genomic regions in NIL B184 and IBM302 RILs suggests that this negative factor is not in close proximity to the *Bx* gene cluster. In recent studies with B73/Mo17 hybrids it was shown that transcription rates in hybrids are either additive or biased to increased transcription rates (Qin *et al.*, 2013, Swanson-Wagner *et al.*, 2006). The reduction of the *Bx1* transcript level represents hence a rare result for hybrids of these two parents.

Given that the *Bx1* transcript level in Mo17 is one essential factor for high late DIMBOA, the finding that there is no hybrid vigor for this trait is easily explained since expression is restricted to the Mo17-allele and the level is even reduced by unknown factor(s) contributed by B73 in hybrids.

4.4 The αβ-region is a hotspot of recombination

Fine mapping of the chromosome 4 QTL (*QTL4-1*) uncovered two recombination hotspots within the *Bx* gene cluster. Six recombination events were detected within the 64 kb fragment

bordered by 148kb and 210kb markers, and 13 recombination events are close to and within the αβ-duplicate region (3.8 kb in B73 and 7.4 kb in Mo17). No recombination event was found in the span of 140 kb between the αβ-region and the proximal *Bx1* region (β-marker and the marker 5kb). In most organisms meiotic recombination events are not distributed evenly throughout the genome. Although environmental factors can also influence the recombination events (Xu *et al.*, 2006, McMullen *et al.* 2009b), special sequence features are believed to be the basis of high frequency of recombination. Variation of recombination frequencies was observed among NAM population families of maize (McMullen *et al.* 2009b), local (*cis*) genetic variation was supposed to influence the recombination rate. Components that influence recombination have been elucidated from fungi to mammals (reviewed by Goodstadt and Ponting, 2011). It is supposed that principal components are universal. Features of recombination are investigated comprehensively in *Schizosaccharomyces pombe*. Although factors determining the positions of recombination sites are not clearly understood, DNA breaks occur preferred in context of simple sequence motifs during meiosis (Steiner *et al.*, 2009; Wahls and Davidson 2012). A proposed sequence of events is binding of transcription factors to the motifs, followed by chromatin modification. Especially enrichment for trimethylated H3K4 histone is detected. The more open chromatin generated in this way is proposed to recruit components of the recombination machinery like the nuclease Spol1 (Goodstadt and Ponting, 2011). Since open chromatin conformation favours transcriptional activity, a link might exist between of increased recombination and enhancement of transcription. The αβ-region includes such 5-6 nucleotides recombination hotspots motives defined in yeast, but whether these are functionally relevant is unknown.

4.5 Models for *cis*-element dependent regulation of *Bx1* expression

The αβ-region, being located 140 kb upstream of *Bx1*, defines a distal sequence element involved in regulation of a target gene. Several examples of long-range regulatory sequences have been detected in maize recently. The insertion of a transposon of the Hopscotch family about 60 kb upstream was identified as a regulatory element for expression of the gene *teosinte branched 1* (Studer *et al.*, 2011). This insertion was essential in domestication of maize. QTL analysis identified the maize flowering time locus *vegetative to generative transition Vgt1*. *Vgt1* is or contains a long-range, *cis*-regulatory element(s) consisting of about 2 kb single copy DNA, 70 kb upstream of the *ZmRap2.7* gene (Salvi *et al.*, 2007). Paramutation has been recognized and investigated in maize for decades. The expression of paramutable alleles is affected by the status of the homologous allele and the status is inherited. The mechanism underlying paramutation is not completely elucidated but tandem repeats of non-coding DNA-sequences that are distal to the genes are essential for the effect (Stam *et al.*, 2002, Chandler 2013 and Haring *et al.*, 2010 for review). In all these examples chromatin modification of the upstream regulatory sequences is discussed. The hallmarks of differentially expressed chromatin are DNA methylation and histone modification. Changes of DNA and chromatin initiated at one site can spread and lead to expression or repression of distal genes. For *Vgt1,* it is speculated that transcription is regulated by microRNA (miRNA); miR172 has been shown to be involved in repression of the homologous *Arabidopsis* gene Rap2.7. This microRNA is also present in maize and regulation might be analogous (Salvi *et*

al., 2007). In paramutation, the copy number of tandem repeat sequences is positively correlated with the paramutagenic effect, presence of repeat-derived RNA, and is dependent on RNA-dependent RNA-polymerase MOP1 activity (Alleman *et al.*, 2006). Strikingly, tandem repeats are also present in transposable elements and centromeric sequences, which are all subject to RNA-directed transcriptional silencing (Lipman *et al.*, 2004, Villar *et al.*, 2009, Chan *et al.*, 2006). By contrast to these examples the duplication of the αβ-region is correlated with higher target gene expression. No data on microRNA derived from the αβ-region are available at present. Interestingly the region around the αβ-sequence comprising about 3.6 kb is hypomethylated in both B73 and Mo17 as displayed in the 5' methylcytosine database, that was established for B73 and Mo17 using the third fully expanded leaf of two weeks' seedlings (Eichten *et al.*, 2011). Such long stretches of hypomethylated DNA in non-coding regions are rare in this analysis. Only minor methylation differences are detected in the region between the two lines, however, the Mo17 αβ-duplication was not recognized and resolved. Since the developmental stage analyzed by Eichten *et al.* is between the 4 days' seedling and 14 days' leaves used in *Bx1* expression analysis, *Bx1* gene expression is expected to be moderate high in both lines and it would be interesting to see whether the epigenetic mark changes with development and is different between the two lines.

Besides spreading of chromatin status, two major models have been proposed to explain long-range communication between promoters and regulatory enhancer sequences, "looping" and "tracking". The difference of both processes is whether there are direct interactions between the enhancer and the promoter, or facilitated diffusion within the nucleus by an active ‘‘scanning” or ‘‘tracking’’ mechanism (Bulger and Groudine, 2011). The fact that at least two *cis*-elements are required for high late *Bx1* expression can be integrated in both models. In the case of tracking both elements would attract and facilitate the movement of positive regulator(s) towards the target gene and interaction with both elements is required for efficient activator location. In the looping model the *cis*-elements even can be integrated as nucleation sites for protein complexes that by interaction create a loop and bring transcription factors in proximity of the *Bx1* proximal promoter elements. The restriction of the *cis*-effect to *Bx1* can be established by insulators or by a specific interaction between transcription factor(s) bound to the distal *Bx1* promoter, and being not present in the promoters of other genes, especially *Bx2*. The *trans*-factor present in the B73 genome could be a repressing transcription factor that competes with the positive regulator or a function that counteracts either chromatin opening, or spreading of activators or formation of the loop. To better define the regulation mechanism, the status of DNA-methylation and chromatin modification in the αβ-region at different developmental stages has to be determined. Chromatin immune precipitation can reveal histone modification (Stam *et al.*, 2002) and physical interaction between sequences in the *cis*-element region (αβ-region to *Bx1*) can be tested by chromosome conformation capture (3C; Miele *et al.*, 2006). In this analysis the αβ-region can be supposed to be one interaction partner. Candidates for interaction are the region of the gene models GRMZM2G085303 and GRMZM5G856095 that are transcribed in both directions, and the 600 bp “open chromatin “ region upstream of *Bx1*. These can be directly assayed in a targeted approach. TSS and expression level-specific chromatin looping has been shown for epialleles

of the paramutagenic maize *B* locus using 3C (Louwers *et al.*, 2009)

Changes in chromatin conformation can be initiated by microRNA and binding of modeling enzymes (Alleman *et al.*, 2006; Erhard *et al.*, 2009; Regulski *et al.*, 2013). The αβ-region either could be the source or the target-site for interaction. In the αβ-region a gene model is annotated (GRMZM2G364237). This model has only one EST-support and reverse transcription based PCR-amplification with B73 and Mo17 leaf material failed (data not shown). However, this putative transcript maps to the αβ-region and includes the polymorphism between Mo17 and B73. Whether these transcripts have impact on *Bx1* expression is not known. Two mechanisms that would discriminate between the Mo17 and B73 conformation of the αβ-region and lead to different epigenetic modification are possible: A counting mechanism and sequence-specific interaction that recognizes either specifically one α allele or the β-2/α-3 sequence junction that only exists in Mo17. Recently the required geometry of tandem repeats in regulation of epialleles has been elucidated for the *B* locus of maize (Belele *et al.*, 2013, Brzeska *et al.*, 2010). Paramutagenic property has been ascribed to a subsequence of the tandem repeat that is GC-rich and speculatively would be prone to differential DNA-methylation. A major impact of unique junction sites has been excluded. Specific binding of the C-box binding protein (CBBP) CXC-domain protein to the tandem repeats of the *B* locus has been detected and the binding of CBBP-multimeres was proposed as a counting mechanism.

Analysis of the distal regions regulating downstream target genes in stable and transient transgenic approaches was successful in the case of paramutagenic *B* alleles and *tb1* (Belele *et al.*, 2013, Studer *et al.*, 2011). In both cases, the distance between the regulatory element and the target gene did not matter and could be shortened to 1 kb and smaller. Hence it can be proposed that the *cis*-elements of *Bx1* regulation in Mo17 can be further defined in analogous experiments. The generation of such αβ–region-minimal-promotor constructs unfortunately were outside of the scope of this thesis.

4.6 Regulation of biosynthetic gene clusters

Facilitated coordinated regulation of pathway genes was proposed to be a driving force in evolution of gene clusters in secondary metabolite biosynthesis (Osbourn, 2010a, b). Avenacoside biosynthesis in oats is restricted to plant root and all isolated biosynthetic genes are expressed in tissue specific manner. Genes of cyanogenic glucosides of *Lotus japonicus* also share a preferred expression in young tissue. Similar to the situation for the *Bx*-genes in maize, the glucosyltransferase has a more general expression pattern than the aglucone biosynthetic genes (Takos *et al.*, 2011). Although the expression pattern for genes in the cluster may differ in detail, stages of high coordinate expression for all members of different clusters are revealed. It is tempting to define possible master regulators that activate all pathway genes at these stages, however knowledge about whole cluster regulation is scarce. Recently a basic leucine zipper transcription factor has been described that appears to be a global regulator of induced diterpene synthesis in rice (Okada *et al.*, 2009). Knowledge about the regulation of biosynthetic pathways would provide a handle to alter the expression of

defense compounds either by smart breeding for superior gene combinations or in transgenic approaches. The genetic approach to elucidate *Bx* gene regulation by QTL mapping in inbreds with different DIMBOA-content did not reveal a master regulator. By contrast it was detected that single genes of biosynthetic clusters, e.g. *Bx1*, might be regulated individually and exclusively in different genomic context. At least two *cis*-elements are involved in the regulation and in different lines *trans*-factors might tune the expression rate (Figure 25).

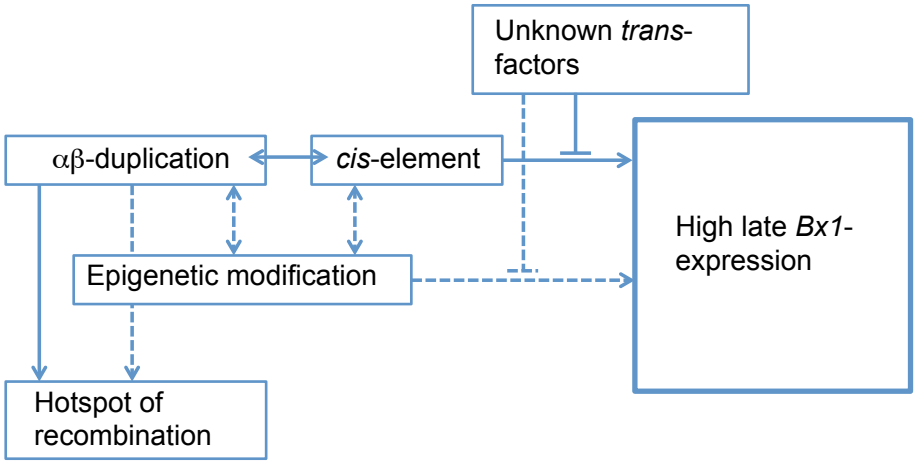


Figure 25: Factors involved in the establishment of high late DIMBOA concentrations in maize. High *Bx1* gene expression is one factor that supports high DIMBOA levels. *Bx1* expression is modulated in concert by the conformation of αβ-region located in *cis* and unknown *cis*-elements between this region and the *Bx1* gene. The *cis*-elements either attract activators or enhance *Bx1* expression by chromatin modification. Unknown *trans*-factors have negative impact on *Bx1* expression. The αβ-region is a hotspot of recombination. Whether this is due to DNA-sequence elements or chromatin modification is unclear. Likewise, it is unknown whether there is an overlap between the competence for recombination and *cis*-activation.

Extended high DIMBOA levels depend on several factors that have not completely been uncovered in this study. By contrast to other features like grain yield, the B73xMo17 hybrid does not show best parent heterosis for late DIMBOA-content, even in contrary, both values are reduced in the hybrids. However, the detection of extremely high late DIMBOA concentrations in IBM 302 inbred lines (Figure 26, appendix) demonstrates that there is potential to improve the concentration of the defense compound in older plants. Since in all analyses made in this study, no case was detected where high *Bx1* expression was connected with low late DIMBOA-content, prolonged activity of the biosynthetic branchpoint gene *Bx1* is probably contributing to the trait and might account for the about 20% contribution of *QTL4-1* to the trait. Initial screening for the effective conformation of the essential *cis*-element "αβ-*Bx1*-region" can be suggested for breeding efforts to improvement plant protection by the intrinsic defense system.

5. Summary

It is well documented that the secondary metabolites benzoxazinoids, are important defence compounds. DIMBOA (2,4-dihydroxy-7-methoxy-2*H*-1,4-benzoxazin-3 (4*H*)-one) is the major benzoxazinoid found in maize, concentrations up to 30 mM with respect to fresh weight are present in young seedlings. The biosynthesis of DIMBOA is elucidated, a set of biosynthetic genes *(Bx* genes) is found in a cluster on chromosome 4. In addition to direct correlations between DIMBOA-content and defence, many QTLs (Quantitativ Trait Loci) of insect and disease resistance in maize have been detected in close proximity to the *Bx* genes and might be ascribed to benzoxazinoids. However, benzoxazinoid biosynthesis generally is a juvenile trait, all *Bx* genes have expression maxima in the first week of germination, and high protective DIMBOA levels are rarely present in older maize plants. The aim of the study was to screen for prolonged maintenance of protective DIMBOA concentration in maize lines and to elucidate the underlying molecular mechanisms.

The panel of NAM (Nested Association Mapping) core parental lines was used to evaluate potential differences in late DIMBOA-content; about 80% of the genetic variation of maize is represented by these lines. Since the DIMBOA-content drops significantly about 20 dai (days after imbibition) in most lines, 24 dai plantlets were screened. Most lines have low DIMBOA-contents at this “late” time point. Mo17 appeared to be an exceptional line that has reproducibly high DIMBOA levels at this stage. B73 proved to belong to the class with low “late” DIMBOA-content. This difference in benzoxazinoid phenotype makes B73 and Mo17 attractive parental lines for QTL mapping. Using the IBM 302 population a major QTL was detected on chromosome 4 (*QTL4-1*) within the *Bx* gene cluster.

QTL4-1 accounts for about 20% of the Mo17 phenotype. Determination of *Bx* gene expression rates showed that *Bx1*, the branchpoint gene of the pathway from primary metabolism, has a significantly elevated expression rate in Mo17 compared to all other lines and it can be speculated that *QTL4-1* is connected with *Bx1* expression. Fine mapping of the *QTL 4-1* revealed that a rare genetic variation (αβ-duplicate), which locates 140 kb upstream of *Bx1*, is required for high and allele-specific late *Bx1* expression. Analysis of hybrids, recombinant lines, recombinant inbred lines (RILs), and near isogenic lines (NILs) indicates that in addition to the αβ-duplicate a yet undefined sequence element(s) between the αβ-duplicate and *Bx1* locus are required for this expression pattern. However, these *cis*-elements and the conferred elevated *Bx1* gene expression alone do not fully explain the high late DIMBOA-content in Mo17 as displayed in the analysis of NIL lines. Probably, concerted contribution of genes underlying the minor QTLs is required.

Zusammenfassung

Es ist seit langem bekannt, dass die Sekundärmetaboliten der Klasse der Benzoxazinoide wichtige Abwehrstoffe darstellen. DIMBOA (2,4-Dihydroxy-7-methoxy-2*H*-1,4-benzoxazin-3(4*H*)-on) ist das Hauptbenzoxazinoid, das in Mais gefunden wird; es werden im Keimling Konzentrationen im Bereich von 30 mM, bezogen auf das Frischgewicht, erreicht. Die Biosynthese von DIMBOA ist aufgeklärt, ein Set der biosynthetischen Gene (*Bx*-Gene) liegt in einem Cluster auf Chromosom 4. Neben der bestehenden direkten Korrelation zwischen DIMBOA-Gehalt und Abwehr, wurden auch viele QTLs (Quantitativ Trait Loci) für Insekten und Krankeitsresistenz in der Nähe der *Bx*-Gene lokalisiert und könnten der Wirkung des Benzoxazinoids zugeschrieben werden. Die Benzoxazinoidbiosynthese ist jedoch generell eine Eigenschaft der jungen Pflanze, alle *Bx*-Gene haben ein Maximum in der ersten Woche nach Keimung und schützende Konzentrationen von DIMBOA sind kaum in älteren Maispflanzen vorhanden. Das Ziel dieser Arbeit war nach verlängerter Präsenz schützender DIMBOA-Konzentrationen in Maislinien zu suchen und die zugrunde liegenden molekularen Mechanismen aufzudecken.

Zur Bestimmung möglicher Unterschiede in Bezug auf "späten" DIMBOA-Gehalt wurde die Kollektion der NAM (Nested Association Mapping)-Elternlienien eingesetzt; diese repräsentieren rund 80% der genetischen Variabilität von Mais. Da der DIMBOA-Gehalt ungefähr mit dem 20. Tag nach Keimansatz (dai, days after imbibition) in den meisten Linien abfällt, wurden 24 dai Pflänzchen untersucht. Die meisten Linien weisen zu diesem Zeitpunkt geringe DIMBOA-Gehalte auf. Die Linie Mo17 ist eine Ausnahme und hat in diesem Stadium reproduzierbar einen hohen DIMBOA-Gehalt. B73 dagegen gehört zu der Gruppe von Linien mit geringem „spätem" DIMBOA-Gehalt. Dieser Unterschied im Benzoxazinoid-Phänotyp macht B73 und Mo17 zu attraktiven Elternlinien für die QTL-Kartierung. Mit Hilfe der IBM 302-Kartierungspopulation wurde ein Haupt-QTL auf Chromosom 4 (*QTL4-1*) innerhalb des *Bx*-Genclusters identifiziert.

QTL4-1 macht etwa 20% des Mo17-Phänotyps aus. Die Bestimmung der *Bx*-Gen-Expressionsraten zeigte, dass *Bx1*, das Gen des Verzweigungpunkts zum Primärstoffwechsel, in Mo17 verglichen mit allen anderen untersuchten Linien signifikant erhöht exprimiert wird. Es kann spekuliert werden, dass der *QTL4-1* mit dieser erhöhten Expression in Verbindung steht. Feinkartierung des *QTLs 4-1* zeigte dass eine seltene genetische Variation (αβ-Duplikation) die 140 kb entfernt vom *Bx1*-Transkriptionsstart gefunden wird, notwendig für hohe und allel-spezifische späte *Bx1*-Expression ist. Analyse von Hybriden, recombinanten Linien, recombinanten Inzuchtlinien (RILs) und nahezu isogenen Linien (NILs) zeigt, dass die αβ-Duplikation und bisher nicht charakterisierte Sequenzelemente zwischen dieser und dem *Bx1*-Lokus notwendig sind, um die erhöhte Expression zu etablieren. Diese *cis*-Elemente und die vermittelte erhöhte *Bx1*-Expression allein erklären den hohen späten DIMBOA-Gehalt in Mo17 nicht vollständig, wie die Analyse der NIL-Linien zeigt. Wahrscheinlich ist dazu ein abgestimmter Beitrag der zugrunde liegenden kleineren QTLs notwendig.

6. References

Alipieva KI, Taskova RM, Evstatieva LN, Handjieva NV, Popov SS (2003). Benzoxazinoids and iridoid glucosides from four Lamium species. *Phytochemistry* 64:1413-1417.

Alleman M, Sidorenko L, McGinnis K, Seshadri V, Dorweiler JE, White J, Sikkink K, Chandler VL (2006). An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* 442:295-8.

Babcock GD, Esen A (1994). Substrate specificity of maize ß-glucosidase. *Plant Science* doi:10.1016.

Balint-Kurti PJ, Zwonitzer JC, Wisser RJ, Carson ML, Oropeza-Rosas MA, Holland JB, Szalma SJ (2007). Precise mapping of quantitative trait loci for resistance to southern leaf blight, caused by Cochliobolus heterostrophus race O, and flowering time using advanced intercross maize lines. *Genetics* 176:645-657.

Basten CJ, Weir B, Zeng ZB (1994). Zmap-a QTL Cartographer. In: Smith C, Gavora JS, Benkel B, Chesnais J, Fairfull W, Gibson JP *et al.* (eds) Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software. Published by the Organizing Committee, 5th World Congress on Genetics Applied to Livestock Production: Guelph, Ontario, Canada. pp 65–66.

Basten CJ, Weir B, Zeng ZB (1997). QTL Cartographer: A reference manual and tutorial for QTL mapping. Department of Statistics, North Carolina State University: Raleigh, NC.

Bauer E, Falque M, Walter H, Bauland C, Camisan C, Campo L, Meyer N, Ranc N, Rincent R, Schipprack W, Altmann T, Flament P, Melchinger AE, Menz M, Moreno-González J, Ouzunova M, Revilla P, Charcosset A, Martin OC, Schön CC (2011). Intraspecific variation of recombination rate in maize. *PLoS One* 6:e28334.

Belele CL, Sidorenko L, Stam M, Bader R, Arteaga-Vazquez MA, Chandler VL (2013). Specific tandem repeats are sufficient for paramutation-induced *trans*-generational silencing. *PLoS Genetics* 9:e1003773. doi: 10.1371.

Bok JW, Chung D, Balajee SA, Marr KA, Andes D, Nielsen KF, Frisvad JC, Kirby KA, Keller NP (2006). GliZ, a transcriptional regulator of gliotoxin biosynthesis, contributes to Aspergillus fumigatus virulence. *Infection and Immunity* 74:6761-6768.

Bommert P, Nagasawa NS, Jackson D (2013). Quantitative variation in maize kernel row number is controlled by the FASCIATED EAR2 locus. *Nature Genetics* 45:334-337.

Broman KW, Wu H, Sen S, Gary A (2003). Churchill: R/qtl: QTL Mapping in Experimental Crosses. *Bioinformatics* 19: 889-890.

Brooks TD, Willcox MC, Williams WP and Buckley PM (2005). Quantitative Trait Loci Conferring Resistance to Fall Armyworm and Southwestern Corn Borer Leaf Feeding Damage. *Crop Science* 45:2430-2434.

Brooks TD, Bushman BS, Williams WP, McMullen MD and Buckley PM (2007). Genetic Basis of Resistance to Fall Armyworm (Lepidoptera: Noctuidae) and Southwest Corn Borer (Lepidoptera: Crambidae) Leaf-Feeding Damage in Maize. *Journal of Economic Entomology* 100:470-1475.

Brzeska K, Brzeski J, Smith J, Chandler VL (2010). Transgenic expression of CBBP, a CXC domain protein, establishes paramutation in maize. *Proceedings of the National Academy of Sciences* 107:5516-5521.

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Oropeza Rosas M, Rocheford TR, Romay MC, Romero S, Salvo S, Sanchez Villeda H, da Silva HS, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen MD (2009). The genetic architecture of maize flowering time. *Science* 325: 714-718

Bulger M, and Groudine M (2011). Functional and mechanistic diversity of distal transcription enhancers. *Cell* 144:327–339.

Bullock WO, Fernandez JM, and Short JM (1987). XL1-Blue: a high efficiency plasmid transforming recA *Escherichia* coli strain with beta-galactosidase selection. *BioTechniques* 5:376-379.

Butrón A, Chen YC, Rottinghaus GE, McMullen MD (2010). Genetic variation at bx1 controls DIMBOA-content in maize. *Theoretical and Applied Genetics* 120:721-734.

Cambier V, Hance T, Hoffmann, E (2000). Variation of DIMBOA and related compounds in relation to the age and plant organ in maize. *Phytochemistry* 53:223-229.

Cardinal AJ, Lee M, Guthrie WD, Bing J, Austin DF, Veldboom, LR and Senior ML (2006). Mapping of factors for resistance to leaf-blade feeding by European corn borer (Ostrinia nubilalis) in maize. *Maydica* 51:93-102.

Chan SW, Zhang X, Bernatavichute YV, Jacobsen SE (2006). Two-step recruitment of RNA-directed DNA methylation to tandem repeats. *PLoS Biology* 4:e363.

Clark RM, Wagler TN, Quijada P, Doebley J (2006). A distant upstream enhancer at the maize domestication gene tb1 has pleiotropic effects on plant and inflorescent architecture. *Nature Genetics.*38:594-597.

Dafoe NJ, Thomas JD, Shirk PD, Legaspi ME, Vaughan MM, Huffaker A, Teal PE, Schmelz EA 2013. European corn borer (*Ostrinia nubilalis*) induced responses enhance susceptibility in maize. *PLoS One* 8:e73394.

Dietl D, Bachelor Thesis, 2012. Investigations of association of DIMBOA-content and *Bx* gene expression with the genotype of European maize inbred lines, Technische Universität München.

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

Eichten SR, Foerster JM, de Leon N, Kai Y, Yeh CT, Liu S, Jeddeloh JA, Schnable PS, Kaeppler SM, Springer NM (2011). B73-Mo17 near-isogenic lines demonstrate dispersed structural variation in maize. *Plant Physiology* 156:1679-1690.

EI-Lithy ME, Bentsink L, Hanhart CJ, Ruys GJ, Rovito D, Broekhof JL, van der Poel HJ, van Eijk MJ, Vreugdenhil D, Koornneef M (2006). New Arabidopsis recombinant inbred line populations genotyped using SNPWave and their use for mapping flowering-time quantitative trait loci. *Genetics* 172:1867-1876.

Erhard KF, Stonaker JL, Parkinson SE, Lim JP, Hale CJ, Hollick JB (2009). RNA polymerase IV functions in paramutation in *Zea mays*. *Science* 323:1201-1205.

Field B, Fiston-Lavier A-S, Kemen A, Geisler K, Quesneville H, Osbourn A (2011). Formation of plant metabolic gene cluster within dynamic chromosomal region. *Proceedings of the National Academy of Sciences* 108:16116-16121.

Field B, Osbourn AE (2008). Metabolic Diversification—Independent Assembly of Operon-Like Gene Clusters in Different Plants. *Science* 320:543-547.

Fincher JA, Vera DL, Hughes DD, McGinnis KM, Dennis JH, Bass HW (2013). Genome-wide prediction of nucleosome occupancy in maize reveals plant chromatin structural features at genes and other elements at multiple scales. *Plant Physiol*ogy 162:1127-1141.

Flint-Garcia SA, Thuillet AC, Yu J, Pressoir G, Romero SM, Mitchell SE, Doebley J, Kresovich S, Goodman MM, Buckler ES (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J*ournal 44:1054-1064.

Frey M, Chomet P, Glawischnig E, Stettner C, Grün S, Winklmair A, Eisenreich W, Bacher A, Meeley RB, Briggs SP, Simcox K, Gierl A (1997). Analysis of a chemical plant defense mechanism in grasses. *Science* 277:696-699.

Frey M, Huber K, Park WJ, Sicker D, Lindberg P, Meeley RB, Simmons CR, Yalpani N, Gierl A (2003). A 2-oxoglutarate-dependent dioxygenase is integrated in DIMBOA-biosynthesis. *Phytochemistry* 62:371-376.

Frey M, Kliem R, Saedler H, Gierl A (1995). Expression of a cytochrome P450 gene family in maize. *Molecular Genetics and Genomics* 246:100-109.

Frey M, Schullehner K, Dick R, Fiesselmann A, Gierl A (2009). Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants. *Phytochemistry* 70:1645-1651.

Frey M, Spiteller D, Boland W, Gierl A (2004). Transcriptional activation of *Igl*, the gene for indole formation in *Zea mays*: a structure-activity study with elicitor-active N-acyl glutamines from insects. *Phytochemistry* 65:1047-1055.

Frey M, Stettner C, Paré PW, Schmelz EA, Tumlinson JH, Gierl A (2000). An herbivore elicitor activates the gene for indole emission in maize. *Proceedings of the National Academy of Sciences* 97:14801-14806.

Ganal MW, Durstewitz G, Polley A, *et al.* (2011). A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *Genome Biology* 14:R103.

Gierl A, Frey M (2001). Evolution of benzoxazinone biosynthesis and indole production in maize. *Planta* 213:493-498.

Gigolashvili T, Berger B, Mock HP, Müller C, Weisshaar B, Flügge UI (2007). The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in Arabidopsis thaliana. *Plant Journal* 50:886-901.

Gonzalo M, Holland JB, Vyn TJ, McIntyre LM (2010). Direct mapping of density response in a population of B73 x Mo17 recombinant inbred lines of maize (*Zea Mays* L.). *Heredity* 104:583-599.

Goodstadt L, Ponting CP (2011). Is the control of recombination conserved among diverse eukaryotes? *Heredity* 106:710-711.

Grombacher AW, Russell WA, Guthrie WD (1989). Resistance to first-generation European corn borer (Lepidoptera: Pyralidae) and DIMBOA concentration in midwhorl leaves of the BS9 maize synthetic. *Journal of the Kansas Entomological Society* 62:103–107.

Grün S, Frey M, Gierl A (2005). Evolution of the indole alkaloid biosynthesis in the genus Hordeum: distribution of gramine and DIBOA and isolation of the benzoxazinoid biosynthesis genes from *Hordeum lechleri*. *Phytochemistry* 66:1264-1272.

Hamilton RH (1964). A Corn Mutant Deficient in, 4-Dihydroxy-7-methoxy-l, 4-benzoxazin-3-on with an Altered Tolerance of Atrazine. *Weeds* 12:27.

Haralampidis K, Bryan G, Qi X, Papadopoulou K, Bakht S, Melton R, Osbourn A (2001). A new class of oxidosqualene cyclases directs synthesis of antimicrobial phytoprotectants in monocots. *Proceedings of the National Academy of Sciences* 98:13431-13436.

Hartmann T (2007). "From waste products to ecochemicals: fifty years research of plant secondary metabolism." *Phytochemistry* 68: 2831-2846.

Holland JB (2007). Genezhetic architecture of complex traits in plants. *Current Opinion Plant Biology* 10:156-61.

Huffaker A, Kaplan F, Vaughan MM, Dafoe NJ, Ni X, Rocca JR, Alborn HT, Teal PE, Schmelz EA (2011). Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiology* 156:2082-2097.

Itkin M, Heinig U, Tzfadia O (2013). Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. *Science* 341:175-179.

Izawa T, Takahashi Y, Yano M (2003). Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and Arabidopsis. *Current Opinion Plant Biology* 6:113-120.

Jampatong C, McMullen MD, Barry BD, Darrah LL, Byrne PF, Kross H (2002). Quantitative trait loci for first and second generation EUropean corn borer resistance derived from the maize inbred Mo47, *Crop Science* 42:584-593.

Jansen RC and Stam P, 1994 High resolution mapping of quantitative traits into multiple loci via interval mapping. *Genetics* 136:1447-1455

Johanssen W (1909). Elemente der exakten erblichkeitslehre. *Fisher*, Jena.

Jonczyk R, Schmidt H, Osterrieder A, Fiesselmann A, Schullehner K, Haslbeck M, Sicker D, Hofmann D, Yalpani N, Simmons C, Frey M, Gierl A (2008). Elucidation of the final reactions of DIMBOA-glucoside biosynthesis in maize: characterization of *Bx6* and *Bx7*. *Plant Physiology* 146:1053-1063.

Kao CH, Zeng ZB, Teasdale RD (1999). Multiple interval mapping for quantitative trait loci. *Genetics* 152:1203-1216.

Kump KL, Bradbury PJ, Wisser RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA, Zwonitzer JC, Kresovich S, McMullen MD, Ware D, Balint-Kurti PJ, Holland JB (2011). "Genome-wide association study of quantitative resistance to southern leaf blight in the maize

nested association mapping population." *Nature Genetics* 43: 163-168.

Klun JA, Guthrie WD, Hallauer AR, Russell WA (1970). Genetic nature of the consentration of 2,4-dihydroxy-7-methoxy2*H*-1,4-Benzoxazin-3(4*H*)-one and resistance to the European corn borer in a diallel set of eleven maize inbreds. *Crop Science* 10:87-90.

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

Lee M, Sharopova N, Beavis WD, Grant D, Katt M, Blair D, Hallauer A (2002). Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. *Plant Molecular Biology* 48:453-461.

Lipman Z and Martienssen R (2004). The role of RNA interference in heterochromatic silencing. *Nature* 431:364.

Liu SC, Kowalski SP, Lan TH, Feldmann KA, Paterson AH (1996). Genome-wide high-resolution mapping by recurrent intermating using Arabidopsis thaliana as a model. *Genetics* 142: 247–258.

Louwers M, Bader R, Haring M, van Driel R, de Laat W, Stam M (2009). Tissue and Expression Level–Specific Chromatin Looping at Maize b1 epialleles. *Plant Cell* 21:832–842.

Lung'aho MG, Mwaniki AM, Szalma SJ, Hart JJ, Rutzke MA, Kochian LV, Glahn RP, Hoekenga OA (2011). Genetic and physiological analysis of iron biofortification in maize kernels. *PLoS One* 6:e20429.

Martin A. (2003). Charakterisierung einer Transkriptionsfaktorfamilie aus *Zea mays* und deren Funktion in der Expression der DIMBOA-Biosynthesegene. Dissertation, Technische Universität München.

McMullen MD, Frey M, Degenhardt J (2009a). Genetics and Biochemistry of Insect Resistance in Maize. In J. L. Bennetzen & S. C. Hake (Ed.), *Handbook of Maize – Its Biology* (271-289). New York: Springer.

McMullen MD, Kresovich S, Villeda HS, Li PB, Sun HQ, Flint-Garcia S, Thornsberry J, Acharya C, and Bottoms C (2009b). Genetic properties of the maize nested association mapping population. *Science* 325:737-740.

McMullen MD, Simcox KD (1995). Genomic Organisation of Disease and Insect Resistance Genes in Maize. *Molecular Plant-Microbe Interactions* 8:811-815.

Meihls LN, Handrick V, Glauser G, Barbier H, Kaur H, Haribal MM, Lipka AE, Gershenzon

J, Buckler ES, Erb M, Köllner TG, Jander G (2013). Natural variation in maize aphid resistance is associated with 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside methyltransferase activity. *Plant Cell* 25:2341-2355.

Meneghini MD, Wu M & Madhani HD (2003). Conserved histone variant H2A.Z projects euchromatin from the ectopic spread of silent heterochromatin. *Cell* 112:725-736.

Miele A, Gheldof N, Tabuchi TM, Dostie J, Dekker J (2006). Mapping chromatin interactions by chromosome conformation capture. *Current Protocols in Molecular Biology* 21:11.

Mugford ST, Qi X, Bakht S, Hill L, Wegel E, Hughes RK, Papadopoulou K, Melton R, Philo M, Sainsbury F, Lomonossoff GP, Roy AD, Goss RJ, Osbourn A (2009). A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. *Plant Cell* 21:2473–2484.

Mylona P, Owatworakit A, Papadopoulou K, Jenner H, Qin B, Findlay K, Hill L, Qi X, Bakht S, Melton R, Osbourn A (2008). *Sad3* and *Sad4* are required for saponin biosynthesis and root development in oat. *Plant Cell* 20: 201–212.

Niemeyer HM (1988). Hydroxamic Acids (4-hydroxy-1,4-benzoxazin-3-ones), defence chemicals in the *Gramineae*. *Phytochemistry* 27:3349-3358.

Niemeyer HM (2009). Hydroxamic Acids Derived from 2-Hydroxy-2*H*-1, 4-Benzoxazin-3 (4*H*)-one: Key Defense Chemicals of Cereals. *Journal of Agricultural and Food Chemistry* 57:1677-1696.

Nomura T, Ishihara A, Imaishi H, Ohkawa H, Endo TR, and Iwamura H (2003). Rearrangement of the genes for the biosynthesis of benzoxazinones in the evolution of Triticeae species. *Planta* 217:776-782.

Nomura T, Ishihara A, Yanagita RC, Endo TR, and Iwamura H (2005). Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat. *Proceedings of the National Academy of Sciences* 102:16490-16495.

Okada A, Okada K, Miyamoto K, Koga J, Shibuya N, Nojiri H, Yamane H (2009). OsTGAP1, a bZIP transcription factor, coordinately regulates the inductive production of diterpenoid phytoalexins in rice. *The Journal of Biological Chemistry* 284:26510-26518.

Osbourn A (2010a). Gene Clusters for Secondary Metabolic Pathways: An Emerging Theme in Plant Biology. *Plant Physiology* 154:531-535.

Osbourn A (2010b). Secondary metabolic gene clusters: evolutionary toolkits for chemical innovation. *Trends in Genetics* 26:449-457.

Papadopoulou K, Melton RE, Leggett M, Daniels MJ, Osbourn AE (1999). Compromised disease resistance in saponin-deficient plants. *Proceedings of the National Academy of Sciences* 96:12923-12928.

Poland JA, Bradbury PJ, Buckler ES, Nelson RJ (2011). Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proceedings of the National Academy of Sciences* 108:6893-6898.

Qi X, Bakht S, Leggett M, Maxwell C, Melton R, Osbourn A (2004). A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proceedings of the National Academy of Sciences* 101:8233-8238.

Qin J, Scheuring CF, Wei G, Zhi H, Zhang M, Huang JJ, Zhou X, Galbraith DW, Zhang HB (2013). Identification and characterization of a repertoire of genes differentially expressed in developing top ear shoots between a superior hybrid and its parental inbreds in *Zea mays*. *Mol Genet Genomics* 288:691-705.

Regulski M, Lu Z, Kendall J, Donoghue MT, Reinders J, Llaca V, Deschamps S, Smith A, Levy D, McCombie WR, Tingey S, Rafalski A, Hicks J, Ware D, Martienssen RA (2013). The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Research* 23:1651-1662.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, and Allard RW (1984). Ribosomal DNA spacerlength polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences* 81:8014–8019.

Sakamoto T, Miura K, Itoh H, Tatsumi T, Ueguchi-Tanaka M, Ishiyama K, Kobayashi M, Agrawal GK, Takeda S, Abe K, Miyao A, Hirochika H, Kitano H, Ashikari M, Matsuoka M (2004). An overview of gibberellin metabolism enzyme genes and their related mutants in rice. *Plant Physiology* 134: 1642–1653.

Salvi S, Sponza G, Morgante M, Tomes D, Niu X, Fengler KA, Meeley R, Ananiev EV, Svitashev S, Bruggemann E, Li B, Hainey CF, Radovic S, Zaina G, Rafalski JA, Tingey SV, Miao GH, Phillips RL, Tuberosa R (2007). Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proceedings of the National Academy of Sciences* 104:11376-11381.

Schaeffer ML, Harper LC, Gardiner JM, Andorf CM, Campbell DA, Cannon EK, Sen TZ, Lawrence CJ. (2011). MaizeGDB: curation and outreach go hand-in-hand. Database 2011:bar022.

Schnable JC, Freeling M (2011). Genes identified by visible mutant phenotypes show increased bias toward one of two subgenomes of maize. *PLoS One* 6: e17855.

Schnable JC, Freeling M, Lyons E (2012). Genome-wide analysis of syntenic gene deletion in the grasses. *Genome Biology and Evolution* 4:265-277.

Schnable PS, Ware D, Fulton RS et.al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112-1115.

Sekhon RS, Lin H, Childs KL, Hansey CN, Buell CR, de Leon N and Kaeppler SM (2011). Genome-wide atlas of transcription during maize development. *Plant Journal* 66: 553–563.

Sicker D, Frey M, Schulz M, Gierl A (2000). Role of natural benzoxazinones in the survival strategy of plants. *International Review of Cytology* 198:319-346.

Soshnikova N, Duboule D (2009). Epigenetic temporal control of mouse *Hox* genes in vivo. *Science* 324:1320-1323.

Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, Galjart N, de Laat W (2006). CTCF mediates long-range chromatin looping and local histone modification in the *β*-globin locus. *Genes and Development* 20:2349-2354.

Stam M, Belele C, Dorweiler JE, Chandler VL (2002). Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. *Genes and Development* 16:1906–1918.

Stam M (2009). Paramutation: a heritable change in gene expression by allelic interactions in trans. *Molecular Plant* 2:578-588.

Steiner WW, Steiner EM, Girvin AR and Plewik LE (2009). Novel Nucleotide Sequence Motifs That Produce Hotspots of Meiotic Recombination in Schizosaccharomyces pombe. *Genetics* 2:459-469.

Steiner WW, Davidow PA, Bagshaw AT (2011). Important characteristics of sequence-specific recombination hotspots in Schizosaccharomyces pombe. *Genetics* 187:385-396.

Strauss J, Reyes-Dominguez Y (2011). Regulation of secondary metabolism by chromatin structure and epigenetic codes. *Fungal Genetics and Biology* 48:62-69.

Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011). Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nature Genetics* 43:1160-1163.

Studier FW and Moffatt BA (1986). Use of Bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. *Journal of Molecular Biology* 198:113-130.

Sue M, Nakamura C, Nomura T (2011). Dispersed benzoxazinone gene cluster: molecular

characterization and chromosomal localization of glucosyltransferase and glucosidase genes in wheat and rye. *Plant Physiology* 157:985-997.

Swaminathan S, Morrone D, Wang Q, Fulton DB, Peters RJ (2009). CYP76M7 is an ent-cassadiene C11 α-hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. *Plant Cell* 21:3315-3325.

Swanson-Wagner RA, Jia Y, DeCook R, Borsuk LA, Nettleton D, Schnable PS (2006). All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. *Proceedings of the National Academy of Sciences* 103:6805-6810.

Takos AM, Knudsen C, Lai D, Kannangara R, Mikkelsen L, Motawia MS, Olsen CE, Sato S, Tabata S, Jørgensen K, Møller BL, Rook F (2011). Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in Lotus japonicus and suggests the repeated evolution of this chemical defence pathway. *Plant Journal* 68:273-286.

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES (2011). Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nature Genetics* 43:159–162.

Tinker NA, Mather DE (1995a). Methods for QTL analysis with progeny replicated in multiple environments. Journal of Agricultural Genomics Volume 1, Published with permission from CAB International at http://wheat.pw.usda.gov/jag/.

Tinker NA, Mather DE (1995b). MQTL: software for simplified composite interval mapping of QTL in multiple environments. Journal of Agricultural Genomics Volume 1, Published with permission from CAB International at http://wheat.pw.usda.gov/jag/.

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002). Looping and interaction between hypersensitive sites in the active β-globin locus. *Molecular Cell* 10:1453-1465.

Turner GW and Croteau R (2004). Organization of monoterpene biosynthesis in Mentha. Immunocy tochemical localizations of geranyl diphosphate synthase, Limonene-6-Hydroxylase, Isopiperitenol dehydrogenase, and pulegone reductase. *Plant Physiology* 136: 4215-4227.

Utz H, Melchinger A (1996). PLABQTL: a program for composite interval mapping of QTL. *Journal of Agricultural Genomics* 2:1–5.

Villar CB, Erilova A, Makarevich G, Trosch R, Kohler C (2009). Control ofPHERES1 imprinting in Arabidopsis by direct tandem repeats. *Molecular Plant* 2: 654–660.

Vincze T, Posfai J and Roberts RJ (2003). NEBcutter: A program to cleave DNA with

restriction enzymes. *Nucleic Acids Research* 31: 3688-3691.

von Rad U, Hüttl R, Lottspeich F, Gierl A, Frey M (2001). Two glucosyltransferases are involved in detoxification of benzoxazinoids in maize. *Plant Journal* 28:633-642.

Wallace JG, Larsson SJ, Buckler ES (2013). Entering the second century of maize quantitative genetics. *Heredity*.2013.6.

Wang JW, Xu T, Zhang LW, Zhong ZM and Luo SM (2007). Effects of methyl jasmonate on hydroxamic acid and phenolic acid content in maize and ist allelopathic activity to *Echinochloa crusgalli* (L.). *Allelopathy Journal* 19:161-170.

Wang S. Basten CJ and Zeng ZB (2012). Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm).

Wegel E, Koumproglou R, Shaw P, Osbourn A (2009). Cell Type-Specific Chromatin Decondensation of a Metabolic Gene Cluster in Oats. *Plant Cell* 21:3926-3936.

Wilderman PR, Xu M, Jin Y, Coates RM, Peters RJ (2004). Identification of *syn*-imara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. *Plant Physiology* 135: 2098–2105.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007). An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLoS one* 2: e718. doi:10.1371.

Winzer T, Gazda V, He Z, Kaminski F, Kern M, Larson TR, Li Y, Meade F, Teodor R, Vaistij FE, Walker C, Bowser TA, Graham IA (2012). A *Papaver somniferum* 10-Gene Cluster for Synthesis of the Anticancer Alkaloid Noscapine. *Science* 336:1704-1708.

Xu X, Hsia AP, Zhang L, Nikolau BJ and Schnable PS (2006). Unequal Sister Chromatid and Homolog Recombination at a Tandem Duplication of the *a1* Locus in Maize. *Genetics* 173: 2211–2226.

Yu J, Holland JB, McMullen MD, Buckler ES (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics* 178: 539-551.

Zeng ZB (1994). Precision mapping of quantitative trait loci. *Genetics* 136: 1457-1468.

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE (2007). Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. *PLoS Biology* 5:e129.

Zhang N, Gibon Y, Gur A, Chen C, Lepak N, Höhne M, Zhang Z, Kroon D, Tschoep H, Stitt M, Buckler E (2010). Fine quantitative trait loci mapping of carbon and nitrogen metabolism enzyme activities and seedling biomass in the maize IBM mapping population. *Plant Physiology* 154:1753-1765.

Zheng Y, Szustakowski JD, Fortnow L, Roberts RJ, Kasif S (2002). Computational identification of operons in microbial genomes. *Genome Research* 12:1221–1230.

Zwonitzer JC, Coles ND, Krakowsky MD, Arellano C, Holland JB, McMullen MD, Pratt RC and Balint-Kurti PJ (2010). Mapping Resistance Quantitative Trait Loci for Three Foliar Diseases in a Maize Recombinant Inbred Line Population – Evidence for Multiple Disease Resistance? *Phytopathology* 100:72-79.

7. Appendix

Figures:

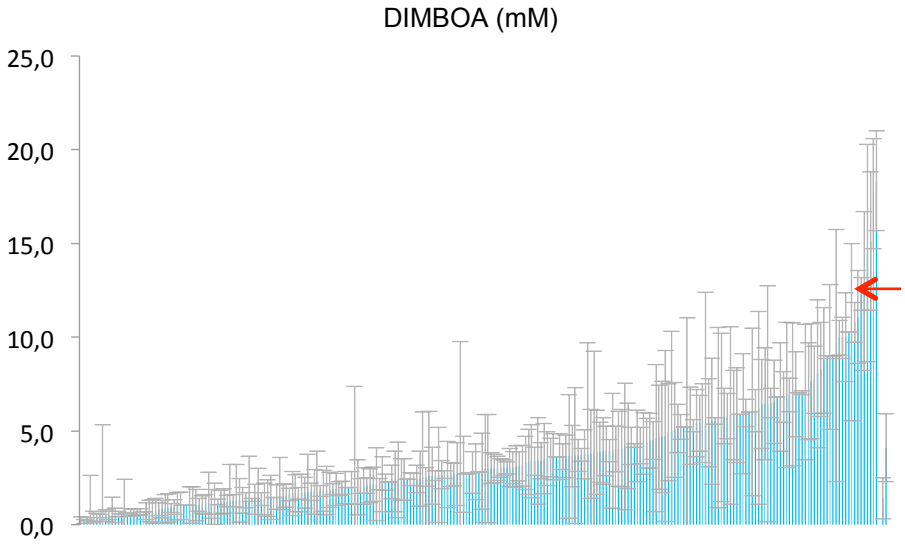
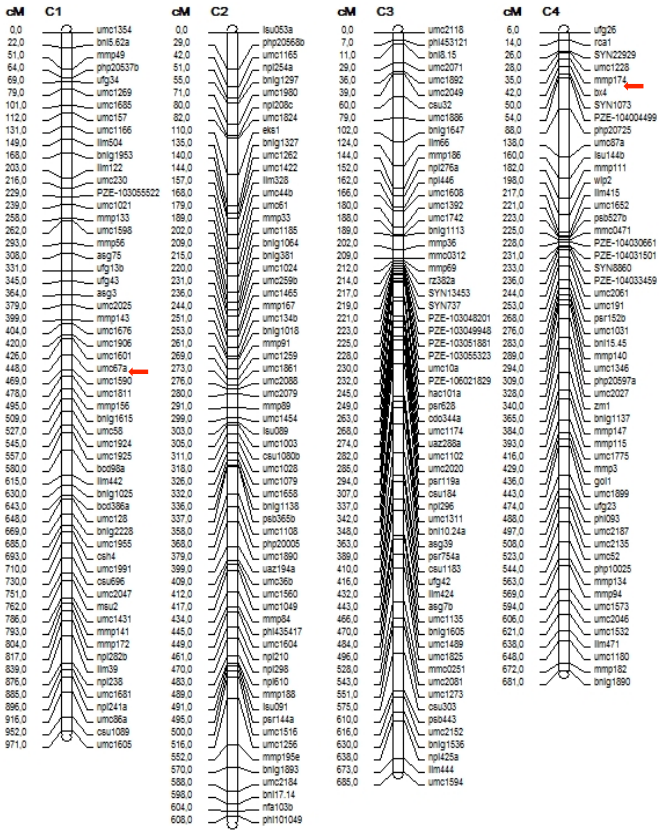


Figure 26: DIMBOA amount variation of IBM population. There is large variation among the lines in 24 dai plants. The red arrow indicated the line selected (MO038) to construct the population for fine mapping. The parental lines B73 and Mo17 are given as the second last and last to the right.



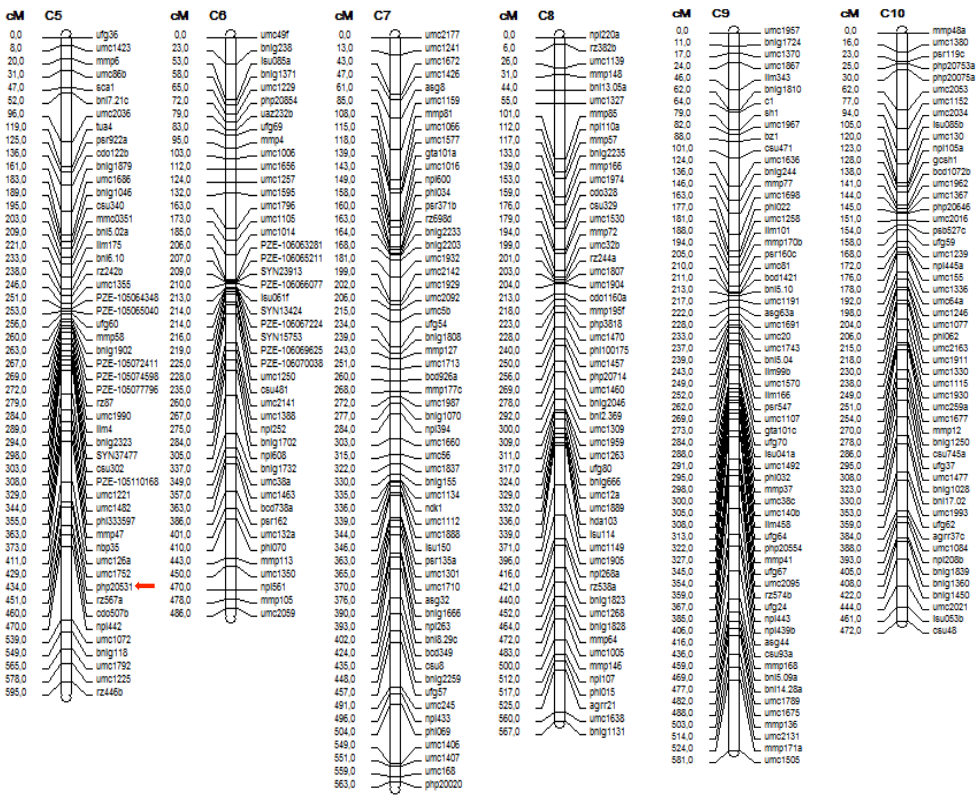
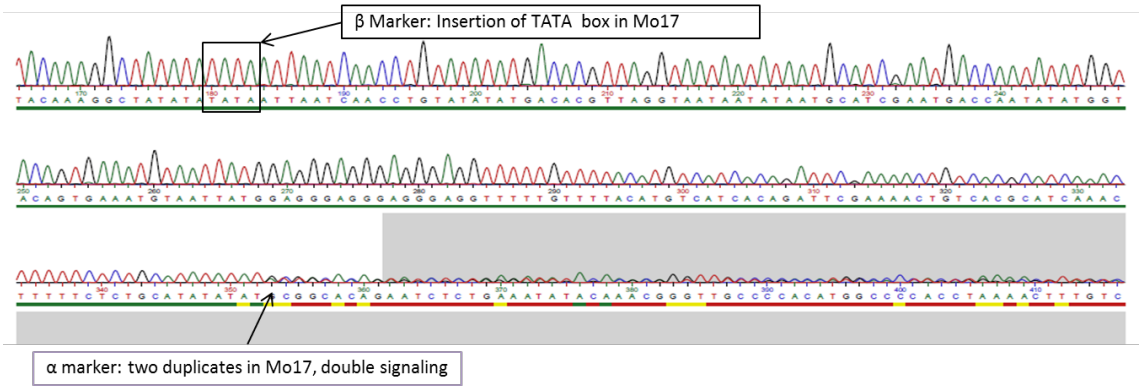


Figure 27: Chromosome graphs and position of detected QTLs in IBM population. The red arrows indicate the QTL positions. The marker for QTL on chromosome 1 is umc1972, for QTL on chromosome 4 is PUT-163a-94477564-4893, for QTL on chromosome 5 is php20531.

A



B

```
3_T7          G---AAAAAGTTTTACG----TACGTGCATAATTGCATAATATATAATCATCTGATAACG 113
8_T7          G---AAAAAGTTTTACG----TACGTGCATAATTGCATAATATATAATCATCTGATAACG 110
11_T7         ACAAAAAAAGTTTTACGGACGTACGTGCATAATTGCATAATATATAATCATCTGACAACG 120
1_T7          ACAAAAAAAGTTTTACGGACGTACGTGCATAATTGCATAATATATGATCATCTGACAACG 119
              .   *************    ************************.********* ****

3_T7          TCTCAG--AGGTGCATGTGCGAGAAAAC-CTTGTCTAATTCCTAGCTACTTTGTACCATA 170
8_T7          TCTCAG--AGGTGCATGTGCGAGAAAAC-CTTGTCTAATTCCTAGCTACTTTGTACCATA 167
11_T7         TCTCAGCTAGGTGCATGTGCGACAACCTGCTTGTCTAATTCCTAGCTACTTTGTAGCATA 180
1_T7          TCTCAGCTAGGTGCATGTGCGACAACCTGCTTGTCTAATTCCTAGCTACTTTGTAGCATA 179
              ******  ************** **..  ************************** ****

3_T7          TATATATGCAGAGAAAAAGTTTGATGCGTGACAGTTTTCGAATCTGTGATGAATCACTAG 230
8_T7          TATATATGCAGAGAAAAAGTTTGATGCGTGACAGTTTTCGAATCTGTGATGAATCACTAG 227
11_T7         TATA--TGCAGAGAAAAAGTTTGATGCGTGACAGTTTTCGAATCTGTGATGAATCACTAG 238
1_T7          TATA--TGCAGAGAAAAAGTTTGATGCGTGACAGTTTTCGAATCTGTGATGATTCCCAAT 237
              ****  ************************************************:**.*:*
```

Figure 28: Sequence signature of αβ markers in Mo17. A Sequence chromatogram of the Mo17 sequence. (PCR product of AlphaF and AlphaR in Mo17). The 4 bp indel is indicated. The arrow point to the region where the Mo17 duplicate deviates and causes double labeling; B Sequence alignment of the B73 sequence and the duplicate found in Mo17.

Table 9: QTL analysis with the sub-population of lines Mo17 at *QTL4-1*.

		QTL cartographer						
sQTL	Chromosome	Marker interval [a]	Confidence interval(bp)	Peak Location	LOD	Threshold [b]	Additive effect	R2 (%) [c]
sQTL3	3	bnlg1816-umc1449	176,512,769-185,017,139	umc10a	4.2	4.8	0.67	5.6
sQTL4	4	npi386-PZE-104031917	22,710,598- 38,725,625	PZE-104029222	6.5	4.8	-1.3	18.1
sQTL5	5	psr167-mmp60	65,115,102- 145,865,325	PZE-105071013	5.7	4.8	-1.3	16.4

a: Markers on both sides of the interval; b: Threshold calculated by 1000 permutations; c: Amount of phenotypic variation explained by each identified QTL.

Alignments (ClustalW2):

Sequence 1: Alignment of *Bx1* upstream sequences

```
B73     CATTTTCTCAAGGATATCGATGTTATCCATAAGGTATTTCTTGAACTTCTTATATTTCCC 540
Mo17    CATTTTCTCAAGGATATCGATGTTATCCATAAGGTATTTCTTGAACTTCTTATATTTCCC 71
        ************************************************************

B73     TTCGACATTTATATTCCATCCTTTCAACATTTTTTTGTTCAATCTTTTTTGTTTTTTTCC 600
Mo17    TTCGACATTTATATTCCATCCTTTCAACATTTTTTTGTTCAATCTTTTTTGTTTTTTTCC 131
        ************************************************************

B73     TTTCCAAACATCGATACATTTCCTGCTCCTCACAGGTAAGGACGAGCTTTCAAAAAACCT 660
Mo17    TTTCCAAACATCGATACATTTCCTGCTCCTCACAGGTAAGGACGAGCTTTCAAAAAACCT 191
        ************************************************************

B73     TCTGCTTTAAAGTCAGGTCTGAGCCTCCAGCAAAGCTCACATATCTAAAGTCCCTCTTCT 720
Mo17    TCTGCTTTAAAGTCAGGTCTGAGCCTCCAGCAAAGCTCACATATCTAAAGTCCCTCTTCT 251
        ************************************************************

B73     TAGTTGGGACAGAGTCAGTGCTAAGACACATGGGAACATGACCAGAAAAAAAAAATCATA 780
Mo17    TAGTTGGGACAGAGTCAGTGCTAAGACACAGGGGAACATGACCAGAAAAAAAAA-TCATA 310
        ***************************** ********************** *****

B73     TTTAGCCCAGAGACAACAATATTCTTGTACTGCAAGTCTCGTTATGGGCTAGCAAAGGAA 840
Mo17    TTTAGCCCAGAGACAACAATATTCTTGTACTGCAAGTCTCGTTATGGGCTAGCAAAGGAA 370
        ************************************************************

B73     TCTACCCAACTTCTCAAATGTGTTGGGATGTCAAGTATATAGACTATTCATCAGTTCCAA 900
Mo17    TCTACCCAACTTCTCAAATGTGTTGGGATGTCAAGTATATAGACTATTCATCGGTTCCAA 430
        ************************************************** *******

B73     CTCTATCAAACTGTGCAGCTCAATTATAGAGTTGAATAAAGTGCTCCATCTATTTGTTCT 960
Mo17    CTCTATCAGACTGTGCAGCTCAATTATAGAGTTGAATAAAGTGCTCCATCTATTTGTTCT 490
        ******** ***************************************************

B73     TATCCTCATATTTGGTTAAGATATTAAAATCACCTCCCACCAACATTTAAAGTGCACCAT 1020
Mo17    TATCCTCATATTTGGTTAAGATATTAAAATCACCTCCCAACAACATTTAAAGTGCACCAT 550
        ************************************* ********************

B73     TTAAAGTGGCTCGCGAGCACCAAACCGCTGAAAACCGGAAATGTTTAGCACGTTGGCAGC 1080
Mo17    TTAAAGTGGCTCGCGAGCACCAAACCGCTGAAAACAGGAAATGTTTAGCACGTTGGCAGC 610
        ********************************** ************************
```

```
B73      GGGACCCTTTTCTATCTCATCGTGTTCTTCGTTGTCCACCACGGCCCACGGGCCAACGCT 1140
Mo17     GGGACCCTTTTCTATCTCATCGTGTTCTTTGTTGTCCACCACGGCCCACGGGCCAACGCT 670
         ***************************** ******************************

B73      CCTCCATCCTGTAGTGTAGAGTATATTCCATTTGCGACCGAGCCGAGCATCGATCCAGCC 1200
Mo17     CCTCCATCCTGTAGTGTAGAGTATATTCCCGTTGCGACCGAGCCGAGCATCGATCCAGCC 730
         *****************************  ******************************

B73      ACACTGGCCACTGCCAGCCAGCCATGTGGCACTCCTACGTATACTACGTGAGGTGAGATT 1260
Mo17     ACACTGGCCACTGCCAGCCAGCCATGTGGCACTCCTACGTATACTACGTGAGGTGAGATT 790
         ************************************************************

B73      CACTCACATGGGATGGGACCGAGATATTTTACTGCTGTGGTTGTGTGAGAGATAATAAAG 1320
Mo17     CACTCACATGGGATGGGACCGAGATATTTTACTGCTGTGGTTGTGTGAGAGATAATAAAG 850
         ************************************************************

B73      CATTTATGACGATTGCTGAACAGCACACACCATGCGTCCAGATAGAGAAAGCTTTCTCTC 1380
Mo17     CATTTATGACGATTGCTGAACAGCACACACCATGCGTCCAGATAGAGAAAGCTTTCTCTC 910
         ************************************************************

B73      TTTATTCGCATGCATGTTTCATTATCTTTTATCATATATATATAACACATATTAAATGAT 1440
Mo17     TTTATTCGCATGCATGTTTCATTATCTTTTATCATATATATATAACACATATTAAATGAT 970
         ************************************************************

B73      TCTTCGTTCCAATTTATAATTCATTTGACTTTTTTATCCACCGATGCTCGTTTTATTAAA 1500
Mo17     TCTTCGTTCCAATTTATAATTCATTTGACTTTTTTATCCACCGATGCTCGTTTTATTAAA 1030
         ************************************************************

B73      AAAATATTATAATTATTGTTACTTTTTGTTGTAATATTGTTTAGCATATAATAAACTTTG 1560
Mo17     AAAATATTATAATTATTGTTACTTTTTGTTGTAATATTGTTTAGCATATAATAAACTTTG 1090
         ************************************************************

B73      ATACTAGTATGTTTCCGAGCAAAAAAAAATATTAATATTTAGATTACGAGCCCATTAATT 1620
Mo17     ATACTAGTATGTTTCCGAGCAAAAAAAAATATTAATATTTAGATTACGAGCCCATTAATT 1150
         ************************************************************

B73     AATTATATTCGAGACAAGCGAAGCAAAGCAAAGCAAGCTAATGTTGCCCCTGCTGTGCAT 1680
Mo17    AATTATATTCGAGACAAGCGAAGCAAAGCAAAGCAAGCTAATGTTGCCCCTGCTGTGCAT 1210
        ************************************************************

B73     GCAGAGGCCCGCTCTTGCTATAAACGAGGCAGCTAGACGCGACTCGACTCATCAGCCTCA 1740
Mo17    GCAGAGGCCCGCTCTTGCTATAAACGAGGCAGCTAGACGCGACTCGACTCATCAGCCTCA 1270
        ************************************************************
```

```
B73      TCAACCTCGACGAAGGAGGAACGAACGGACAGGTTGTTGCACAGAAGCGACATGGCT      1797
Mo17     TCAACCTCGACGAAGGAGGAACGAACGGACAGGTTGTTGCACAGAAGCGACATGGCT      1330
         ********************************************************
```

Sequence 2: Sequences of the 5 parts in Mo17 and the corresponding B73 sequences.

```
>Mo1
CATTGAGGGAATGGTGAGAATTATATGTAGAGTATTATTCTATAGGCACTTTGATTTAAAAATGATGTTCTGGTTTAAGATTTATC
ATACTAGTGTGTTATAGCAAACCACTTAATTAATTAGGAGAGAGTTATAGCAATTAATCTTGTAACTTTAGAAGTAATCTGCCTAG
>Mo2
AATTAAATGAGCTATAGTAGGGAATGGTGAGCTCTCCAACAGTATAATATATAGCTGCTGTGCTATAATAATACGTATAGGTAGCC
ACACTGCTGTCTCTCCTGGCCCAGCACCATACCGAGACAGCTTTATGAAATTTTATGAAATTTTTAAAGTTAGAACTTGGTAGCGA
>Mo3
ATCCACTCTAGTGGACATGAATCATTGAGGGAATGGTGGGAATTATATGTAGAGTATTATTATATAGCCACTTTGATTTAAAAATA
ATGTTCCGGTTTAAGATTTATCATACTAATTGACAGTTATAGCAAACCACTTAATTAATTGGGAGAGAGTTATAGCAATTAATCTT
>Mo4
GTAACTTTAGAAGTAATCTGCCGAGAATTAAATGAGCTATATATAGTAGGGAATGGTGAGCTCTCCAACAGTATAATATAGCTGCT
GTGCTATAATAATACGTATAGGTAGCCACACTGCTGTCTCTCCTGGCCCAGCACCATACCGAGACAGCTTTTTTTCATGCACAGTA
>Mo5
TTATTTGCTCCGTGCACTGTACAATCGCGTGTCGTATCTGATCAATAATCACGAGACAGAGACTGCCTTGATATCGATGGCAACGA
ACATCATGTGCCC
>Mo1
CATTGAGGGAATGGTGAGAATTATA TGTAGAGTATTATTCTATAGGCACTTTGATTTAAAAATGATGTTCTGGTTTAAGATTTAT
CATACTAGTGTGTTATAGCAAACCACTTAATTAATTAGGAGAGAGTTATAGCAATTAATCTTGTAACTTTAGAAGTAATCTGCCTA
G
>Mo3
ATCCACTCTAGTGGACATGAATCATTGAGGGAATGGTGGGAATTATATGTAGAGTATTATTATATAGCCACTTTGATTTAAAAATA
ATGTTCCGGTTTAAGATTTAT CATACTAATTGACAGTTATAGCAAACCACTTAATTAATTGGGAGAGAGTTATAGCAATTAATCTT
>B73
ATGAGTAGCCACATATATATAATGTTGAAGAGAATCGGTGCCTAGCTAGCTAGCCATATATGTAGAGTATTATTATATAGCCACTT
TGATTTAAAATAATGTTCTGGTTTAAGATTTATTTATCATAGTGTCATAGAAAATCACTATCGAGAGAATTGAGGGAGTGGTGAGC
TCTCCAACAATATAAT
>B73
GAGAGAATTGAGGGAGTGGTGAGCTCTCCAACAATA--TAATATAGCTGCTATGCTATGCTATAATACGTATACGTAGCCACTGCT
GTCTC TCCTGGCCCAGCACCATACCGAGACAGCTTT
>Mo2
GAGCTATAGTAGGGAATGGTGAGCTCTCCAACAGTATAATATATAGCTGCTGTGCTATAATAATACGTATAGGTAGCCACACTGCT
GTCTCTCCTGGCCCAGCACCATACCGAGACAGCTTT
>Mo4
TATATATAGTAGGGAATGGTGAGCTCTCCAACAGTATAATATAGCTGCTGTGCTATAATAATACGTATAGGTAGCCACACTGCTGT
CTCTCCTGGCCCAGCACCATACCGAGACAGCTTT
>Mo5
CGTGTCGTATCT GATCAATAATCACGAGACAGAGACTGCCTTGATATCGATGGCAACGAACATCATG
>B73
CGTGTCGTATCT GATCAATAATCACACGAGACAGAGACTCTGCCTAGCTTGACATCGATGGGCAACGAACATCATG
```

Sequence 3: Alignment of αβ-region of B73 and Mo17. The αβ-region of Mo17 was divided into 5 equal fragments to be aligned to the whole B73 sequence.

```
1Mo     ------------------------------------------------------------
3Mo     ------------------------------------------------------------
5Mo     ------------------------------------------------------------
B73     TTTCAGGGGTATTCCATGTGTTTAGGTAATGGGTTTAACCCAGTAGCTATGGGTGGATTT
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------

1Mo     ------------------------------------------------------------
3Mo     ------------------------------------------------------------
5Mo     ------------------------------------------------------------
B73     AGGTGGGTTATTTAGTTTAATTATTATATGGTGGGTTTGGTTGAAGAAAATAACTCATTT
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------

1Mo     ------------------------------------------------------------
3Mo     ------------------------------------------------------------
5Mo     ------------------------------------------------------------
B73     GTAGGGTTTGGGTGGGTTTGACCCTAATAGCAGGGCTAACAACATGACGCTTCCATGCTG
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------

1Mo     ------------------------------------------------------------
3Mo     ------------------------------------------------------------
5Mo     ------------------------------------------------------------
B73     CCGAAGAGCTTCTTCGCATATTCAAGCTCCTCCGCGCCCGCATATTCAAGCTTCTCCGTG
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------

1Mo     ------------------------------------------------------------
3Mo     ------------------------------------------------------------
5Mo     ------------------------------------------------------------
B73     CCGCCGCCGCCGCAGATTTCCTCCATGGCGCCGGCCGAGCTCCTCTGCCCTGGTGGTTCT
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------
```

```
1Mo     ------------------------------------------------------------
3Mo     ------------------------------------------------------------
5Mo     ------------------------------------------------------------
B73     CCTGACCGGTCGAGCTCCTCCGCGCCGCCGCAGCTGAGGGCCTCCACGTCACCCCTGGTG
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------

1Mo     ------------------------------------------------------------
3Mo     ------------------------------------------------------------
5Mo     ------------------------------------------------------------
B73     GCCATCGCAGCTTCTGTCGCGACGAGGAGGGGATATAGCGGATGAACAGTGAATTGAATC
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------

1Mo     ------------------------------------------------------------
3Mo     ------------------------------------------------------------
5Mo     ----------------------------------TTATTTGCTC----------------
B73     GCCAGCTTCAGAGTATACTGTTAATTTAATCGCCATATATGCTCGTAACAACCCGCTTTT
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------

1Mo     ----------------------------------CATTGAGGGAATGGTGAGAATTATA
3Mo     ------------ATCCACTCTAGTGGACATGAATCATTGAGGGAATGGTGGGAATTATA
5Mo     ------------------------------------------------------------
B73     ATGAGTAGCCACATATATATAATGTTGAAGAGAATCGGTGCCTAGCTAGCTAGCCATATA
2Mo     ------------------------------------------------------------
4Mo     ------------------------------------------------------------

1Mo     TGTAGAGTATTATTCTATAGGCACTTTGATTTAAAAATGATGTTCTGGTTTAAGATTTAT
3Mo     TGTAGAGTATTATTATATAGCCACTTTGATTTAAAAATAATGTTCCGGTTTAAGATTTAT
5Mo     ------------------------------------------------------------
B73     TGTAGAGTATTATTATATAGCCACTTTGATTTAAAATAATGTTCT GGTTTAAGATTTAT
2Mo     ------------------------------------------------------------
4Mo     ---------------------------------------------GTAACTTTA-GAAG---

1Mo     CATACTAGT---GTGTTATAGCAAACCACTTAATTAATTAGGAGAGAGT------TATAG
3Mo     CATACTAATTGACAGTTATAGCAAACCACTTAATTAATTGGGAGAGAGT------TATAG
5Mo     ------------------------------------------------------------
B73     TTATCATAGTGTCATAGAAAATCACTATCGAGAGAATTGAGGGAGTGGTGAGCTCTCCAA
2Mo     ---------------------AATTAAATGAGCTATAGTAGGGAATGGTGAGCTCTCCAA
4Mo     -----TAATCTGCCGAGAATTAAATGAGCTATATATAGTAGGGAATGGTGAGCTCTCCAA
```

```
1Mo     CAATTAATCTTGTAACTTTAGAAGTAATC-----TGCCTAG-------------------
3Mo     CAATTAATCTT-----------------------------------------------
5Mo     ---------------------------------------------------------
B73     CAATA--TAATATAGCTGCTATGCTATGCTATAATACGTATACGTAGCCACTGCTGTCTC
2Mo     CAGTATAATATATAGCTGCTGTGCTATAATAATACGTATAGGTAGCCACACTGCTGTCTC
4Mo     CAGTA--TAATATAGCTGCTGTGCTATAATAATACGTATAGGTAGCCACACTGCTGTCTC


1Mo     ---------------------------------------------------------
3Mo     ---------------------------------------------------------
5Mo     --CGTGCACTGTACA----------------------------ATCGCGTGTCGTATCT
B73     TCCTGGCCCAGCACCATACCGAGACAGCTTTTTTTTCATGCACACAAACGTGTCGTATCT
2Mo     TCCTGGCCCAGCACCATACCGAGACAGCTTTATGAAATTTTATGAAATTTTTAAAGTTAG
4Mo     TCCTGGCCCAGCACCATACCGAGACAGCTTTTTTTCATGCACAGTA--------------

1Mo     ---------------------------------------------------------
3Mo     ---------------------------------------------------------
5Mo     GATCAATAATCAC--GAGACA------GAGACTGCCTTGATATCGATGGC-AACGAACAT
B73     GATCAATAATCACACGAGACAGAGACTCTGCCTAGCTTGACATCGATGGGCAACGAACAT
2Mo     AACTTGGTAGCGA-----------------------------------------------
4Mo     ---------------------------------------------------------

1Mo     -------------------------------------------
3Mo     -------------------------------------------
5Mo     CATGTGCCC-----------------------------------
B73     CATGCATGTGCGCTCGTGGAGTAGCTTTCTTTCTTTCTTTATG
2Mo     -------------------------------------------
4Mo     -------------------------------------------
```

8. List of figures

9. List of tables

Acknowledgement

I am deeply grateful to Prof. Dr. Alfons Gierl for the keen interest in the process of the work, his great support, helpful discussions and offering the opportunities for the international exchange.

I would particularly like to thank Dr. Monika Frey for all the constructive instructions, her consideration in both lab work and daily life which make the working and living in Freising enrich and enjoyable, and also her great help with the German dissertation.

Many thanks to our excellent technician, Regina Hüttl and Peter Dobos, the dedicated students Nadine Nurasyid, Robert Fischer, Wolfgang Greiter, Jonathan Herlet, Marc Böhning, Markus Hippich, and David Dietl, who contributed to the work in this project.

Many thanks to Prof. Dr. Ramon A. Torres Ruiz, PD Dr. Erich Glawischnig and Dr. Lilla Römisch-Margl for their willingness to discuss with their professional experience. I also want to thank my colleagues, Dr. Regina Dick, Dr. Ralf Jonczyk, Stefan Lenk, Dr. Michaela Matthes, Zheng Yu, Manisha, Stefanie Mucha, Kamil Bakowski, Teresa Müller, Sebastian Wenz, Heidi Miller-Mommerskamp, Ottilie Peis, and the other students Miriam Hillenmeyer, Beiying Dai, Xiaomeng Yang and Alexandra Chapman, for a great working atmosphere and support.

I would like to specially acknowledge Carolin Ziegler for her constant help with all the administration work.

I appreciate to Prof. Dr. Michael McMullen for the detailed instruction of QTL analysis in the beginning, constant data support and helpful suggestions.

I want to thank Prof. Dr. Chris-Carolin Schön and colleagues from her lab, especially Stefan Schwertfirm for helping with DNA extraction for the fine mapping, Dr. Eva Bauer and Dr. Peter Westermeier for the data support, and Dr. Yu Wang for the suggestions about bioinformatics.

Specially thanks also to Prof. Dr. Claus Schwechheimer and colleagues from his lab, for the discussion in the project and for the generous support in student funding application.

I appreciate to Prof. Dr. Frank Hochholdinger for being my second mentor in the graduate school, and his great help in winter nursery. I thank Dr. Maike Stam for permission of the chance to have the internship in Amsterdam.

Thanks to Petra Wick and SFB924 student activity committee for the help with applying the SFB924 student activities' funding.

I would like to express my deepest appreciation to Prof. Dr. Ralph Hückelhoven for chairing the examination committee.

I am grateful for the constant support from my parents and my husband, Peng Huang.

Linlin Zheng
Liebigstr. 28a
85354, Freising

Lebenslauf

Zur Person

Geburtsdatum/-ort	01.02.1983 in Jilin, P.R. China
Nationalität	chinesisch
Familienstand	verheiratet

Ausbildung

02/2010-02/2014	Promotion und Fertigstellung der vorliegenden Dissertation am Lehrstuhl für Genetik, Wissenschaftszentrum Weihenstephan, Technische Universität München
09/2005-07/2008	Master of Science in Biological Science Tsinghua Universiät, China
09/2000-07/2004	Bachelor of Science in Biological Science Beijing Normal Universität, China
09/1997-07/2000	Nr.1 Mittelschule der Kreisstadt Lishu, China