Iowa State University
Digital Repository @ Iowa State University

Graduate Theses and Dissertations | Graduate College

2008

Genetic analysis of ear development and tassel architecture in maize (Zea mays L. ssp. mays)

Maria Laura Mayor
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/etd

Part of the Agronomy and Crop Sciences Commons

Recommended Citation
Mayor, Maria Laura, "Genetic analysis of ear development and tassel architecture in maize (Zea mays L. ssp. mays)" (2008). *Graduate Theses and Dissertations.* Paper 10636.

This Dissertation is brought to you for free and open access by the Graduate College at Digital Repository @ Iowa State University. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Digital Repository @ Iowa State University. For more information, please contact hinefuku@iastate.edu.

Genetic analysis of ear development and tassel architecture in maize (*Zea mays L. ssp. mays*)

by

Maria Laura Mayor

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Plant Breeding

Program of Study Committee:
Michael Lee (Major Professor)
Philip Dixon
Jode Edwards
Kendall Lamkey
Erik Vollbrecht

Iowa State University

Ames, Iowa

2008

Copyright © Maria Laura Mayor, 2008. All rights reserved.

TABLE OF CONTENTS

TABLE OF CONTENTS ii

LIST OF TABLES v

LIST OF FIGURES viii

ACKNOWLEDGMENTS x

ABSTRACT xii

CHAPTER 1: GENERAL INTRODUCTION 1

LITERATURE REVIEW 2
Shoot branching: prolificacy and tassel architecture 2
Environmental effects and hormonal regulation 3
Quantitative Trait Loci studies on prolificacy and tassel architecture 4
Candidate genes 5
Genetic designs and population structures for QTL mapping 6
DISSERTATION ORGANIZATION 7

CHAPTER 2: GENETIC ANALYSIS OF PROLIFICACY AND RELATED TRAITS IN A NEAR-ISOGENIC POPULATION OF MAIZE RECOMBINANT INBRED LINES 10

ABSTRACT 10
INTRODUCTION 11
MATERIALS AND METHODS 13
Plant material 13
Population development and phenotypic evaluation 14
Statistical analyses of phenotypic data 15
Genotypic evaluation 17
Genotypic analysis and genetic map construction 18
QTL detection 19
Assessment of pleiotropy and linkage for QTL 22
RESULTS 23
Phenotypic analysis 23
Genotypic analysis and genetic map construction 24
QTL detection 25
Assessment of pleiotropy and linkage for QTL 27
DISCUSSION 28
REFERENCES 34
APPENDIX 55

CHAPTER 3: GENETIC ANALYSIS OF TASSEL ARCHITECTURE IN A NEAR-ISOGENIC POPULATION OF MAIZE RECOMBINANT INBRED LINES 61

ABSTRACT 61
INTRODUCTION 62
MATERIALS AND METHODS 64
Plant material 64
Population development and phenotypic evaluation 65
Statistical analyses of phenotypic data 66
Genotypic evaluation 67
Genotypic analysis and genetic map construction 68
QTL detection 70
Assessment of pleiotropy and linkage for QTL 72
RESULTS 73
Phenotypic analysis 73
Genotypic analysis and genetic map construction 74
QTL detection 75
Assessment of pleiotropy and linkage for QTL 77
DISCUSSION 78
REFERENCES 82
APPENDIX 100

CHAPTER 4: GENERAL CONCLUSIONS 112

LITERATURE CITED 114

APPENDIX 1: GENETIC MAP CONSTRUCTION FOR THE C103/C103AP POPULATION OF RILs 120

APPENDIX 2: FORMULAE USED IN THE STATISTICAL ANALYSIS 129

PHENOTYPIC ANALYSIS 129
General Linear Models 129
Estimation of covariance parameters 130
Best linear unbiased estimators (BLUE) and predictors (BLUP) 131
Least Square Means (LS-means) 131
Phenotypic and Genetic correlation coefficients and standard errors 132
QTL ANALYSIS 135
Single-Marker analysis 135
Composite Interval Mapping (CIM) 136
Permutation tests 138

Multiple Interval Mapping (MIM) 139
Assessment of pleiotropy and linkage of QTL 141
REFERENCES 142

LIST OF TABLES

CHAPTER 2: GENETIC ANALYSIS OF PROLIFICACY AND RELATED TRAITS IN A NEAR-ISOGENIC POPULATION OF MAIZE (*Zea mays L. ssp. mays*)

Table 1: Monthly precipitation and GDU accumulation for the period 1987-2007 and for the growing season 2005-2007 at the Agronomy and Agricultural Engineering Research Center (AAERC) near Boone, Iowa (42°1´ N, 93°45´ W) 43

Table 2: Mean phenotypic values of C103, C103AP and 218 maize RILs and variance components for ears per plant (EPP), GDU to anthesis and plant height in the combined analysis of the 2005, 2006 and 2007 datasets 44

Table 3: Phenotypic (above diagonal) and genetic correlations (below diagonal) with their standard errors in parenthesis for ears per plant (EPP), GDU to anthesis and plant height in a maize population of 218 RILs derived from inbreds C103 and C103AP in the combined analysis of the 2005, 2006 and 2007 datasets 45

Table 4: Summary of QTL detected from a population of 218 maize RILs derived from inbred lines C103 and C103AP for ears per plant (EPP), GDU to anthesis and plant height (PH) 46

Table 5: Summary of epistatic interactions detected from a population of 218 maize RILs derived from inbred lines C103 and C103AP for ears per plant (EPP), GDU to anthesis and plant height (PH) 50

Table A1: Mean phenotypic values of inbreds C103, C103AP and 218 RILs and variance components for ears per plant (EPP), GDU to anthesis and plant height in years 2005, 2006 and 2007 55

Table A2: Phenotypic (above diagonal) and genetic correlations with their standard errors (in parenthesis) in years 2005, 2006 and 2007 for ears per plant (EPP), GDU to anthesis and plant height 56

Table A3: Regression coefficients from single-marker analysis at loci of chromosomes 2, 5 and 10 from a maize population of 218 RILs derived from inbreds C103 and C103AP. 57

Table A4: Summary of QTL detected by Multiple-trait Composite Interval Mapping in the three years of evaluation (2005, 2006 and 2007) for ears per plant (EPP), GDU to anthesis and plant height from a maize population of 218 RILs derived from the inbreds C103 and C103AP 58

CHAPTER 3: GENETIC ANALYSIS OF TASSEL ARCHITECTURE IN A NEAR-ISOGENIC POPULATION OF MAIZE RECOMBINANT INBRED LINES

Table 1: Summary of QTL for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) reported in previous and in the current study 88

Table 2: Monthly precipitation and GDU accumulation for the period 1987-2007 and for the growing season 2006-2007 at the Agronomy and Agricultural Engineering Research Center (AAERC) near Boone, Iowa (42°1´ N, 93°45´ W)..... 92

Table 3: Mean phenotypic values of maize inbreds C103, C103AP and 218 RILs, variance components and broad-sense heritability for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL)..... 93

Table 4: Phenotypic (above diagonal) and genetic correlations with their standard errors (in parenthesis) for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) in a maize population of 218 RILs derived from inbreds C103 and C103AP in the combined analysis of the 2005, 2006 and 2007 datasets..... 94

Table 5: Summary of QTL detected from a population of 218 maize RILs derived from inbred lines C103 and C103AP for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL)..... 95

Table A1: Mean phenotypic values of inbreds C103, C103AP and 218 RILs, variance components and broad-sense heritability for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) in years 2006 and 2007 100

Table A2: Phenotypic (above diagonal) and genetic correlations with their standard errors (in parenthesis) for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) in years 2006 and 2007..... 101

Table A3: Summary of QTL detected from a population of 218 maize RILs derived from inbred lines C103 and C103AP for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) from a population of 218 RILs in years 2006 and 2007..... 102

Table A4: Regression coefficients from single-marker analysis at loci of chromosomes 2, 5 and 10 from a maize population of 218 RILs derived from inbreds C103 and C103AP for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL)..... 106

Table A5: Summary of QTL detected by Multiple-trait Composite Interval Mapping in the three years of evaluation (2005, 2006 and 2007) for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) from a population of 218 RILs derived from inbreds C103 and C103AP 107

Table A1.1: List of loci that were not polymorphic between C103 and C103AP 120

Table A1.2: Polymorphic loci not included in the construction of the C103/C103AP genetic map. Genomic location based on the Maize database, number of alleles observed for C103 and C103AP, chi-square statistic and p-value for testing the 1:1 segregation ratio and genetic map annotation are shown 124

Table A1.3: Genomic location; observed and expected number of C103 and C103AP alleles for a 1:1 segregation ratio, chi-square statistics and its correspondent probability for SSR loci from the genetic map of the C103/C103AP population of RILs with significant departure from 1:1 segregation ratio 126

LIST OF FIGURES

CHAPTER 2: GENETIC ANALYSIS OF PROLIFICACY AND RELATED TRAITS IN A NEAR-ISOGENIC POPULATION OF MAIZE (*Zea mays L. ssp. mays*)

Figure 1: Histograms showing the phenotypic distribution of the number of ears per plant (EPP) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population 52

Figure 2: The genetic map from a maize population of 218 RILs derived from inbred lines C103 and C103AP using 123 SSR loci. The 1-LOD support interval of QTL is shown. Dotted lines represent genomic regions not polymorphic between C103 and C103AP and connect linkage groups of the same chromosome. SSR locus position (in cM) is shown on the left of each chromosome and its name is shown on the right 53

Figure A1: Histograms showing the distribution of GDU to anthesis from a population of 218 RILs derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population 59

Figure A2: Histograms showing the distribution of plant height from a population of 218 RILs derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population 60

CHAPTER 3: GENETIC ANALYSIS OF TASSEL ARCHITECTURE IN A NEAR-ISOGENIC POPULATION OF MAIZE RECOMBINANT INBRED LINES

Figure 1: Graphical representation of the measurements assessed on the tassels adapted from Upadyayula et al (2006a). TL tassel length, BZL branching zone length, CSL central spike length..... 97

Figure 2: Linkage map of the C103/C103AP population constructed with information from 218 RILs. Relative positions of loci are given in centimorgans at the left side of the linkage groups. Loci names are given at the right side of the chromosomes. QTLs with 1-LOD support intervals for the combined analysis are indicated. Dotted lines represent genomic regions not polymorphic between C103 and C103AP and connect linkage groups of the same chromosome..... 98

Figure A1: Histograms showing the phenotypic distribution of the tassel branch number (TBN) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred

line over replications. The value in parenthesis under "Mean" is the mean of the population 108

Figure A2: Histograms showing the phenotypic distribution of the tassel length (TL) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population..... 109

Figure A3: Histograms showing the phenotypic distribution of the central spike length (CSL) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population..... 110
Figure A4: Histograms showing the phenotypic distribution of the branching zone length (BZL) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population 111

APPENDIX 1: GENETIC MAP CONSTRUCTION FOR THE C103/C103AP POPULATION OF RILs

Figure A1.1: Total number of polymorphic and non-polymorphic SSR loci evaluated on each chromosome from the screening of C103 and C103AP inbreds. 128

ACKNOWLEDGMENTS

I am in debt with my advisor, Dr. Michael Lee, who not only guided me through my graduate program by providing advisory and mentoring, but also provided the environment that allowed me to develop my interests in genetics, statistics and plant breeding. These four years I spent as a member of Dr. Lee's group have been an invaluable learning experience for me. I am grateful to the committee members from my Program of Study for their guidance and invaluable suggestions during the development of my graduate program. Special thanks to Dr. Jode Edwards and Dr. Philip Dixon for their comments and suggestions on the statistical analysis of the data.

The influence that my former advisor at the Universidad Nacional de Rosario, Argentina; Dr. Liliana Picardi, had on my scientific career is greatly acknowledged. She was my first mentor and guide through the fascinating world of genetics. She transmitted me her affection and interest for science as well as encouragement.

Many thanks to fellow graduate students Kendra Meade, Victor Abertondo and Thanda Dhliwayo as well as the undergraduate students who helped me with data collection and field activities. Special mention to Thanda, I truly enjoyed our discussions on technical and philosophical aspects of plant breeding and genetics. I also extend my gratitude to Wendy Woodman-Clickemann and Anthony Assibi Mahama for their teaching, patience and help with the lab work. Special thanks to Assibi that read the drafts of my manuscripts and gave me useful comments.

Appreciation is extended to the Plant Science Institute and the Baker Center for Plant Breeding at Iowa State University that provided the financial support. Special thanks to Fulbright Commission for financially supporting part of my Ph.D. program.

This stage in my career would not have been accomplished without the unconditional support, love and caring of my family. To my parents, Clelia and Raul, I will be forever grateful. They encouraged me to pursue my dreams and taught me perseverance and determination in the course of achieving my goals in life, thank you! Finally, thanks to all my friends from here and there that supported me in numerous ways through my graduate program.

ABSTRACT

Yield potential of maize (*Zea mays* L.) has been increased significantly during the last century. Along with genetic gains for grain yield, changes in other traits have included an increase in the number of ears per plant (i.e. fewer barren plants) and a reduction in tassel size. The objectives of this study were 1) to identify Quantitative Trait Loci (QTL) associated with number of ears per plant (EPP), growing degree units to anthesis (GDU), plant height (PH) and tassel architectural traits, and 2) to evaluate the consistency of the QTL across environments. A population of 218 recombinant inbred lines (RILs) derived from two nearly isogenic inbreds, C103 and C103AP was evaluated for EPP, GDU, PH and four tassel architectural traits. The genetic map of 123 Simple Sequence Repeat (SSR) loci covered 894 cM. At least 5 novel regions for EPP were detected on chromosomes 2, 3, 6, 8 and 9. A region flanked by loci *umc1858* and *umc1309*, on chromosome 8 (bins 8.04-8.05; a bin is an arbitrary subdivision of the maize genome based on a set of core markers) had a major influence on EPP, PH and GDU to anthesis. With respect to tassel morphology, a total of 32 QTL were identified for tassel branch number (TBN), tassel length (TL), central spike (CSL) and branching zone length (BZL). The majority of these QTL were located on chromosomes 1, 2, 3, 4 and 8. The QTL for TBN, TL and CSL with strong association to the phenotypic variance were located in bins 2.01, 2.06, 2.08 and 9.03. In these bins candidate genes and QTL have not been identified; therefore, this is the first report of a biological function with respect to tassel morphology for those regions in the genome. Comprehensive descriptions of the QTL related to the traits evaluated in this study are provided in the individual chapters of this dissertation. Many results found have not been described previously in the literature and will contribute to the current knowledge. Finally, further

study of these regions is required for better understanding of the genetic factors affecting meristem initiation, maintenance and development in maize.

CHAPTER 1: GENERAL INTRODUCTION

Over the last century maize grain yields in the U.S. have increased steadily. Along with this genetic gain for grain yield, changes in other traits as a result of direct or indirect selection by the breeders have occurred. These changes include resistance to a wide variety of biotic and abiotic stresses (i.e, drought tolerance or higher plant density) as well as morphological and physiological changes that promote efficiency in growth, development, and partitioning (i.e. reduction in tassel size). There has been a significant improvement towards maize plants with more ears per plant (i.e. fewer barren plants); as a result each plant has the genetic potential to produce one or more grain-bearing ears at increasing plant densities (Duvick 2005).

Ears and tassels share a common developmental origin. Maize plant architecture is primarily determined by complex patterns of branching. After embryogenesis, three types of shoot apical meristems (SAMs) are formed: vegetative lateral SAMs, inflorescence lateral SAMs, and floral meristems. Vegetative lateral SAMs can form either tillers or secondary ear shoots when the axil is several nodes from the apex. The inflorescences of maize develop in terminal positions: the tassel from the primary SAM and the ears from the uppermost lateral SAMs. Likewise, a series of reiterative patterns of branching at the SAM and uppermost lateral axillary meristem define the establishment of the different inflorescence structures prior to the formation of the floral meristem (Evans and Barton 1997; Schmitz and Theres 2005).

Understanding the genetic architecture of complex traits may begin by identifying the genetic factors regulating these traits, mapping the subset of genetically varying QTL in

natural populations, and determining the molecular polymorphisms defining QTL alleles (Mackay 2004). QTL mapping studies are essential on this matter since they can provide some information on the minimum number of genomic regions that affect a trait, the distribution of gene effects, and the relative importance of additive and non-additive gene action, the relationship between genomic regions and the environment, and their impact on other related traits. The scope of inferences drawn from QTL studies is often restricted by the limitations of the experimental designs (i.e. number of parents and progeny). Nevertheless, recent advances in experimental and statistical procedures, including the simultaneous analysis of QTL that segregate in diverse germplasm, may ultimately lead to the identification of the specific DNA sequence variants underlying QTL (Holland 2007).

The objectives of this study were: 1) to identify QTL associated with the number of ears per plant, growing degree units to anthesis, plant height and tassel architectural traits, and 2) to evaluate the effect of the environment in the detection of QTL.

LITERATURE REVIEW

Shoot branching: prolificacy and tassel architecture

Maize has a natural tendency to be prolific (i.e. an increase in the number of ears per plant) by developing a visible potential ear at each node below the uppermost ear (Sass and Loeffel 1959). The development of secondary ears has an agronomic advantage over single-eared germplasm as prolific genotypes are less subject to barrenness under high plant populations, drought and low-nitrogen environments (Jampatong et al. 2000; Russell 1968; Russell and Eberhart 1968).

Tassels can affect grain yield by reducing light interception to the lower canopy as well as by utilizing carbohydrate resources (Duncan et al. 1967). In addition to its relevance for seed production, variation in tassel morphology could be an indicator of genetic diversity in both wild and cultivated maize (Orr and Sundberg 1994; Westerbergh and Doebley 2002).

Environmental effects and hormonal regulation

The effect of environment on the regulation of the branching patterns by the apical meristem has been extensively studied in maize. Competition among plants for light, water and nutrients affects the development of ears and tassels. Stressful environmental conditions that favor the dominance of the apical meristem lead to a concentration of resources in the main stalk and a suppression in the development of lateral branches (Bechoux et al. 2000; Earley et al. 1974; Motto and Moll 1982; Moulia et al. 1999; Sorrells et al. 1979; Xu et al. 2004).

Apical dominance, the process by which axillary buds of developing shoots are generally kept quiescent due to a growing apical meristem, is the central mechanism in the regulation of plant shoot architecture (Tamas 1995). Auxins have been proposed as an essential regulator in the initiation and development of the axillary meristem. Meanwhile, cytokinins have been proposed to mediate the action of auxin in apical dominance. In addition, abscisic acid was proposed to act as an inhibitor for axillary bud growth, although no further evidence has been reported (Wang and Li 2008). Recently, an additional second messenger of auxin, a carotenoid-derived long-range signaling pathway, involved in controlling axillary meristem outgrowth has been identified by genetic and grafting studies in *Arabidopsis*, pea and petunias. This pathway is represented by *MAX* genes in *Arabidopsis* (Schmitz and Theres 2005; Wang and Li 2008). Although *MAX*-dependent branching signals

are not attributed to any known phytohormone, it has been demonstrated that the *MAX* genes act as negative regulators of the polar auxin transport (Wang and Li 2008).

Quantitative Trait Loci studies on prolificacy and tassel architecture

Quantitative trait loci (QTL) mapping has detected several genomic regions associated with the development of secondary ears in populations from bi-parental crosses. Most (20 of 30) of the QTL detected were located to chromosome 1, 3 and 6 (Agrama and Moussa 1996; Austin and Lee 1996; Austin and Lee 1998; de Leon et al. 2005; Lauter and Doebley 2002; Veldboom and Lee 1994; Veldboom and Lee 1996). Changes for prolificacy in the Golden Glow maize population were suggested to be due to the *teosinte branched1* (*tb1*) locus on chromosome 1 (de Leon and Coors 2002). Further candidacy of *tb1* in the development of secondary ears was suggested by QTL analysis in the same population (de Leon et al. 2005).

Previous reports on QTL for tassel morphology (i.e. tassel branch number, length of the branching zone, central spike and total tassel length) identified regions on chromosomes 2, 4, 5 and 7 to be the most frequently detected for the individual variation of these traits (Table 1, Chapter 2). In four of these genetic studies, inbred B73 was the parent of the populations of RILs and the recurrent parent in the backcross population suggesting a narrow genetic background (Mickelson et al. 2002; www.maizegdb.org/MIP; verified 12/06/2007; Table 1). Only one of these studies simultaneously considered variation in traits affecting ear architecture (i.e. kernel row number, number of kernels per row and kernel number density) along with variation in tassel morphology (Upadyayula et al. 2006a). It is plausible that genetic factors influencing the expression of tassel architecture might affect ear development, since both traits rely on a common developmental origin.

Candidate genes

Maize inflorescences develop from the shoot apical meristem (SAM). The SAM gives rise to the terminal tassel, which has long branches and develops the male flowers. The ears are derived from axillary shoot meristems, have a prominent axis and develop the female flowers. The inflorescence meristems (IMs) of the tassel and the ear produce spikelet pair meristems (SPMs). The SPMs form a short branch, bearing two Spikelet Meristems (SMs), that produce a pair of floral meristems (FMs), though in the ear only one of these develops into a fertile flower (Bommert et al., 2005b).

The genetic control of maize inflorescence development has been partially elucidated through the analysis of mutant alleles with effects on the tassel and ear (Malcomber et al. 2006; McSteen and Leyser 2005). Some genes involved in axillary meristem initiation affect both vegetative and inflorescence branching, whereas genes controlling axillary meristem determinacy/identity are usually specific to either the vegetative or reproductive phase (McSteen and Leyser 2005).

The loci *Suppressor of sessile spikelet* (*Sos1*), *barren stalk1* (*ba1*) and *barren inflorescence2* (*bif2*) together with *tb1* affect the initiation and maintenance of inflorescence meristem (Gallavotti et al. 2004; McSteen and Hake 2001; Ritter et al. 2002). Also *Sos1*, *Bif1, bif2*; and *barren sterile* (*bs)* loci reduce inflorescence branches by affecting different steps in the development of the inflorescence meristem (Evans and Barton 1997). Other genes affecting inflorescence development are *fasciated ear2* (*fea2*), ramosas (*ra1, ra2* and *ra3*), *thick tassel dwarf* (*td1*) and *branched silkless* (*bd1*) (Bommert et al. 2005a; Bommert et al. 2005b; Bortiri et al. 2006; Vollbrecht et al. 2005).

Epistatic interactions between these candidate loci have been described to integrate genetic and hormonal control of axillary meristem initiation. For instance, *Bif1* exhibits an epistatic interaction with *ba1* and a synergistic interaction with *bif2*; playing an overlapping role with the latter one in the initiation of lateral organs. This suggests that *Bif1* functions as a regulator of auxin transport (Barazesh and McSteen 2008). Recently, Wu and McSteen (2007) demonstrated that *bif2* is expressed upstream and *ba1* is expressed downstream of auxin transport, integrating the genetic and hormonal control of axillary meristem initiation. However, interactions between these candidate genes and the environment have not been documented.

Integration of the information from these candidate loci and from QTL mapping studies would enhance our understanding of the genetic networks underlying ear and tassel development.

Genetic designs and population structures for QTL mapping

Near Isogenic Lines (NILs) are constructed by mating a donor parent to an inbred line (recurrent parent) to form the F1 generation. The resulting offspring is then backcrossed to the recurrent parent for several generations. After the backcrossing, the NIL is formed by at least one generation of selfing (Lynch and Walsh 1998; Muehlbauer et al. 1988).

Populations of near-isogenic lines (NILs) have been widely used to verify and fine-map QTL previously detected in other mapping populations in maize and other species. Also, the use of NILs has been proposed to simultaneously map, verify and incorporate QTL into adapted, elite genetic background (Graham et al. 1997; Salvi and Tuberosa 2005; Salvi et al. 2002; Sanchez et al. 2000; Szalma et al. 2007). Eshed and Zamir (1995) demonstrated the advantages of mapping QTL in nearly isogenic backgrounds by using introgression lines

(ILs) originated from a cross between *Lycopersicon pennellii* and cultivated tomato. In such highly isogenic background, the proportion of the phenotypic variability explained by a QTL is greatly increased along with an increase in the detection of minor QTL (Eshed and Zamir 1995; Zamir and Eshed 1998).

Populations of recombinant inbred lines (RILs) constitute a permanent set of individuals in which segregation is complete or nearly complete and can be used indefinitely for mapping. In addition, the RILs can be evaluated in multiple environments, thus providing an easy approach for measuring QTL × environment interactions. Since a genotype is represented by an inbred line, rather than an individual, a more accurate assessment of the genetic variance component can be made (Burr and Burr 1991; Burr et al. 1988). However due to more opportunities for meiotic recombination, this type of population has less linkage disequilibrium between alleles at the marker locus and alleles of the linked QTL compared to F_2/F_3 populations (Tanksley 1993).

DISSERTATION ORGANIZATION

This dissertation focuses on two aspects of shoot branching and development of maize inflorescence: ear development, and tassel architecture. Inflorescence development was addressed by analyzing prolificacy and tassel morphological traits in a population of 218 RILs generated by mating two near isogenic inbreds, C103 and C103AP. C103 is a single-eared inbred line that tends to abort its ears. Inbred line C103AP was obtained by four backcrosses with continuous selection for resemblance to C103 and for development of secondary earshoots (Duvick 1974).

The second chapter focuses on ear development, by analyzing prolificacy and two other related traits: growing degree units to anthesis and plant height. This chapter could be considered as an extension of the genetic studies on prolificacy in testcross hybrids from inbreds C103 and C103AP started by Duvick (1974). This author demonstrated that prolificacy can be transferred from exotic germplasm to a non-prolific elite inbred, C103, by continuous backcrossing while selecting for prolific plants. Also that study suggested that prolificacy could exhibit a relatively simple inheritance pattern influenced by a few genetic factors in some genetic contexts. Relative to C103, the converted prolific inbred, C103AP, requires more growing degree units to reach anthesis and is taller. These phenotypic changes did not interact with population levels or locations, suggesting linkage or pleiotropy of both traits to the prolific habit in the donor parent (Duvick 1974). The first objective of this chapter was to identify the underlying genetic factors affecting the number of ears per plant, growing degree units to anthesis and plant height, in the population of RILs. The second objective was to evaluate the consistency of the QTL across environments.

The third chapter addresses tassel architecture by evaluating four traits: number of main branches in the tassel, tassel length, central and branching zone length in the same population of RILs. When evaluating the population, we observed that both parents also differed in tassel architecture, while C103 had a lax tassel, C103AP showed a smaller and compact tassel. These observations ended in a characterization of tassel architecture and finally detection of regions in the genome associated with these traits. Since the development of tassel and ear are related events, it is plausible that prolificacy could be associated with changes in tassel morphology. The first objective was to associate phenotypic changes in tassel architecture with regions in the genome. A second objective of this chapter was to

detect interactions between QTL and assess the effect of the environment in the expression of these traits.

This dissertation includes two manuscripts to be submitted to *Theoretical and Applied Genetics*. The manuscripts are preceded by a General Introduction (Chapter 1) section and followed by a General Conclusions (Chapter 4) section. References cited in these two sections are listed after the General Conclusions.

CHAPTER 2: GENETIC ANALYSIS OF PROLIFICACY AND RELATED TRAITS IN A NEAR-ISOGENIC POPULATION OF MAIZE RECOMBINANT INBRED LINES

A paper to be submitted to Theoretical and Applied Genetics

M.L. Mayor, M. Lee and A.A. Mahama

ABSTRACT

Prolificacy and two related traits, GDU to anthesis and plant height, were investigated in a population of 218 recombinant inbred lines derived from two nearly isogenic inbreds, C103 and C103AP in three years. The objectives were: 1) to identify quantitative trait loci (QTL) associated with the number of ears per plant (EPP), growing degree units to anthesis (GDU) and plant height (PH), 2) to resolve the previously reported genetic relationships among EPP, GDU and PH for these inbreds and 3) to evaluate the consistency in the location of the detected QTL across environments. The genetic map of 123 SSR loci covered 894 cM. At least 5 novel regions for EPP were detected on chromosomes 2, 3, 6, 8 and 9. A region flanked by loci *umc1858* and *umc1309*, on chromosome 8 (bins 8.04-8.05) had a major influence on EPP, PH and GDU. Statistical evidence that favored pleiotropy instead of linkage for EPP and GDU was observed at QTL in bin 6.06, and for EPP and PH at QTL in bin 3.10. The statistical evidence of pleiotropic gene action concurs with previous

observations that C103AP contained genetic factors that increased EPP, GDU and PH relative to C103.

INTRODUCTION

The maize shoot may develop an ear at each node below the uppermost ear, prolificacy (Sass and Loeffel 1959). The development of secondary ears may have an agronomic advantage over single-eared germplasm as two-ear genotypes are less likely to exhibit barrenness under stressful conditions such as high plant populations, drought and low-nitrogen environments (Jampatong et al. 2000; Russell 1968; Russell and Eberhart 1968).

The expression of prolificacy relies on the genetic and environmental context. Previous studies emphasize that gene action for controlling prolificacy is mainly of the additive type (Hallauer and Miranda Fo. 1988; Motto and Moll 1983; Stucker and Troyer 1977). Duvick (1974) demonstrated that prolificacy can be transferred from exotic germplasm to a non-prolific elite inbred, C103, by continuous backcrossing while selecting for prolific plants. That study also suggested that prolificacy could exhibit a relatively simple inheritance pattern influenced by a few genetic factors in some genetic contexts. Relative to C103, the converted prolific inbred, C103AP, requires more growing degree units to reach anthesis and is taller. These phenotypic changes in the converted inbred did not interact with population levels or locations, suggesting linkage or pleiotropy of GDU to anthesis and plant height to the prolific habit in the donor parent (Duvick 1974).

Quantitative trait loci (QTL) mapping has detected several genomic regions associated with prolificacy in populations from bi-parental crosses. Most (20 of 30) of the

QTL detected were located to chromosome 1, 3 and 6 (Agrama and Moussa 1996; Austin and Lee 1996; Austin and Lee 1998; de Leon et al. 2005; Lauter and Doebley 2002; Veldboom and Lee 1994; Veldboom and Lee 1996). Associated changes for prolificacy in the Golden Glow maize population were attributed mainly to the *teosinte branched1* (*tb1*) locus on chromosome 1 (de Leon and Coors 2002). Further candidacy of *tb1* in the development of secondary ears was suggested by QTL analysis in the same population (de Leon et al. 2005).

The genetic control of maize earshoot development has been partially elucidated through the analysis of alleles with highly qualitative effects, mutants (Malcomber et al. 2006; McSteen and Leyser 2005). A number of genes have been characterized that affect identity, determinacy and maintenance of the inflorescence meristem and have been assembled in a developmental pathway (Wu and McSteen 2007). *Suppressor of sessile spikelet* (*Sos1*), *barren stalk1* (*ba1*) and *barren inflorescence2* (*bif2*) together with *tb1* affect the initiation and maintenance of the inflorescence meristem (Gallavotti et al. 2004; McSteen and Hake 2001; Ritter et al. 2002). Other genes may have a direct or indirect influence in the development of secondary ears such as *fasciated ear2* (*fea2*), *thick tassel dwarf1* (*td1*) and *barren fastigiate* (*baf*). However, a relationship between such loci and QTL for prolificacy and barrenness has not been established in elite maize germplasm, and other features of the shoot developmental network remain unsolved.

In this study, prolificacy and two related traits, GDU to anthesis and plant height, were investigated in a population of recombinant inbred lines (RILs) derived from two near-isogenic inbreds, C103 and C103AP. The advantages of developing populations of near-isogenic lines (NILs) to verify and fine-map QTL previously detected in other mapping populations have been exploited in maize and other species. Also, the use of NILs has been

proposed to simultaneously map, verify and incorporate QTL into adapted, elite genetic background (Graham et al. 1997; Salvi et al. 2002; Sanchez et al. 2000; Szalma et al. 2007). Besides the advantage of RILs for evaluation in multiple environments, most of the genomes of both parents are identical by descent; thus, the population of RILs would be segregating for smaller regions of the genome carrying the QTL of interest. In such isogenic backgrounds, the proportion of the phenotypic variability explained by a QTL may be increased because of the absence of other segregating QTL (Eshed and Zamir 1995, Lynch and Walsh 1998; Muehlbauer et al. 1988).

The first objective of this study was to identify genomic regions associated with the number of ears per plant and two related traits, growing degree units to anthesis and plant height, in a population of RILs derived from the mating of inbred lines C103 and C103AP. The second objective was to attempt to resolve the association among number of ears per plant, growing degree units to anthesis and plant height. The third objective was to evaluate the consistency in the location of the detected QTL across environments.

MATERIALS AND METHODS

Plant material

C103 is a single-eared inbred line that tends to abort ear development and produce barren stalks. This inbred belongs to the Lancaster Sure Crop Group (Jones and Everett 1949). An accession of a prolific Argentinean popcorn population (AP) was used as donor parent to increase the prolificacy of C103. Inbred line C103AP was developed by four backcrosses with continuous selection for resemblance to the recurrent parent, C103, and for development of secondary earshoots under plant populations of 40000 plants ha^{-1}. Then,

C103AP was compared to C103 in hybrid combinations at three plant densities: 24700, 49400 and 74100 plants ha^{-1}. Even at the lowest density, testcross hybrids from C103AP developed approximately 1.25 versus 1 ear per plant; when compared to C103 hybrids (Duvick 1974).

Population development and phenotypic evaluation

A population of 218 Recombinant Inbred Lines (RILs) was developed by mating inbreds C103 and C103AP. Each RIL was derived from a single F_2 plant by self-pollinating one plant per line for 7 consecutive generations until the $F_{7:8}$ generation, following the single-seed descent method. For the field experiments, seed for each RIL was increased by selfing three plants in the $F_{7:8}$ generation and harvesting them in bulk.

During 2005, 2006 and 2007, the 218 RILs and three other entries (one entry for C103 and two entries for C103AP) were evaluated in a 13 × 17 lattice design with 2 replications at the Agronomy and Agronomic Engineering Research Center (AAERC), near Boone, Iowa (42°1´ N, 93°45´ W). Annual and historic weather data were collected from an automated weather station located less than one kilometer from the field plots (Table 1). Monthly growing degree units (GDU) for the period of May to August for 2005 were 2% lower to the average for the period 1987-2007, whereas GDU for 2006 and 2007 exceeded the 20-year average by 4 and 9% respectively. The season with least precipitation (May through August) was 2006, followed by 2005 and 2007; all of them with less precipitation than the 20-year average (Table 1).

The field was fertilized prior to planting with 175 kg N ha^{-1} as urea. For weed control, 1.75 L ha^{-1} of Dual II Magnum (S-metolachlor/benoxacor) and 2.22 L ha^{-1} of atrazine [2-chloro-4-(ethylamine)-6-(isopropylamine)-s-triazine] were applied to the field. In 2005, the

first replication was planted on 10 May. The planting of the second replication was delayed by rain until 19 May. In 2006 and 2007 the experiment was planted on 10 May and 6 May, respectively. Each plot consisted of a row of 5.5 m long. The seeding rate was two kernels per hill. The distance between hills in a row was 0.56 m and between rows was 0.75 m. The experiments were hand-planted to achieve a final plant density of 24700 plant ha^{-1}. At growth stage V5 (Ritchie et al. 1996), each hill was thinned to one plant. The few tillers were removed at maturity and before harvesting ears from the main culms.

First, the number of main culms per plot was recorded, and then ears from all main culms in a plot were harvested in bulk. The number of ears per plant (EPP) was calculated as the total number of ears (including ears with and without kernels) on the main culms per plot divided by the number of plants in that plot. Growing degree units to anthesis (GDU) was determined when 50% of plants in a plot reached anthesis. The GDUs were calculated as: $\sum_{i=1}^{n}\left(\frac{T_{Max}+T_{min}}{2}-10°C\right)$, where T_{Max}= maximum daily temperature (in °C) , T_{min}= minimum daily temperature (in °C) and i=1,…, n days from planting to 50% anthesis. The minimum value for T_{min} was set at 10ºC and the maximum value for T_{Max} was set at 30ºC (Cross and Zuber (1972). After anthesis was complete, plant height (PH) was assessed by measuring the height of each main stalk in the plot, from the soil surface to the ligule of the flag leaf; and averaged per plot.

Statistical analyses of phenotypic data

Histograms of the mean phenotypic values from the population of RILs were generated for each trait. This provided an assessment of their distribution and located the mean phenotypic values of the RILs and parents. A combined analysis was completed for the

2005, 2006 and 2007 datasets. The homogeneity of variance was evaluated using the Levene's test on absolute residuals (Levene 1960). A mixed-effects linear model obtained by Restricted Maximum Likelihood (REML) was used to estimate the phenotypic values according to the model: $y_1 = \mu_1 + \tau_{1i} + \alpha(\tau)_{1ij} + \beta(\alpha\tau)_{1ijk} + \gamma_{1l} + (\gamma\tau)_{1il} + \varepsilon_{1ijkl}$; where μ_1 = overall mean for trait *1*; τ_{1i} = effect of the *i* environment (*i*=1, …, 3); $\alpha(\tau)_{1ij}$ =effect of the *j* replications (*j*= 1, 2); $\beta(\alpha\tau)_{1ijk}$ = effect of the *k* blocks within each replication (*k*=1, 2… 17); γ_{1l} = effect of the *l* RILs (*l*=1, 2…, 218); $(\gamma\tau)_{1il}$ = term to account for the genotype by environment interaction and ε_{ijkl} = error term. The analyses were performed using the MIXED procedure of SAS (SAS 2007, version 9.1.3) in which the effects of year, block, and replication within blocks were specified as fixed in the MODEL statement, whereas the effect of the RILs was specified as random. The traits EPP and GDU to anthesis showed heterogeneous residual variances. To account for this heterogeneity, the REPEATED statement with the GROUP=year option was used in the MIXED procedure of SAS (Littel et al. 2006; SAS 2008). Best linear unbiased predictors (BLUPs) were used to estimate the phenotypic value of each RIL (Bernardo 2002).

Least square means of C103, C103AP and the mean of the RILs were obtained by REML. Differences among the parents and the mean of the RILs population were assessed using Tukey's multiple-comparison test. Degrees of freedom were estimated using the Kenward–Rogers version of the Satterthwaite approximation (Kenward and Roger 1997).

Variance component estimates and their standard errors were obtained by REML and used to estimate the broad-sense heritability on a progeny-mean basis and its standard error for each year and a combined estimate across the three years of evaluation (Holland et al.

2003; Table 2). For the combined analysis, the broad-sense heritability on a progeny-mean basis was calculated as: $H^2 = \frac{\hat{\sigma}_\lambda^2}{\left(\frac{\hat{\sigma}_\tau^2}{e}\right) + \hat{\sigma}_\lambda^2 + \left(\frac{\hat{\sigma}_\varepsilon^2}{re}\right)}$ where: $\hat{\sigma}_\lambda^2$ = genetic variance of the trait; $\hat{\sigma}_\tau^2$ = variance due to environmental factors; $\hat{\sigma}_\varepsilon^2$ = error variance; *e*= number of environments and *r* = number of replications (Hallauer and Miranda Fo. 1988). Spearman rank-order correlation coefficients of the predicted values of the RILs in the three environments were obtained in order to determine whether the 'genotype × environment' interactions were due to variation in the magnitude or in the ranking of the genotypes.

Estimates of the phenotypic and genetic correlations were obtained using REML from the MIXED procedure in SAS. Approximate sampling variances and standard errors for the phenotypic and genotypic correlation estimates were obtained with the delta method (Holland 2006; Lynch and Walsh 1998).

Genotypic evaluation

Before anthesis, leaf tissue from 218 RILs, C103 and C103AP was collected and lyophilized. Each inbred was represented by bulked leaf tissue of 10 plants. Genomic DNA was extracted via the modified-CTAB protocol (Saghai-Maroof et al. 1984). The original donor parent, AP, was not available for genotypic evaluation. C103 and C103AP were evaluated at 890 Simple Sequence Repeat (SSR) loci distributed throughout the maize genome. These loci were selected from the maize database based on polymorphism among a standard set of inbred lines (www.maizegdb.org; verified 12/07/07). A high frequency of monomorphic loci between C103 and C103AP was expected since the latter was developed by four backcrosses and phenotypic selection for resemblance to C103. Polymorphism

between C103 and C103AP was detected at 174 SSR loci. A subset of 123 SSR loci, with clear and repeatable PCR products, and maximum coverage of the genome, was used to construct the linkage map.

The PCR reagents were: 10 X buffer (Tris-HCl 100 mM, pH 8.3, 500 mM KCl, 1mg mL^{-1} gelatin); 2.5 mM MgCl2; 0.1 mM dNTPs; 0.17 pmol μL^{-1} forward- and reverse-primers, 0.2 mg mL^{-1} BSA, 1X betaine, 0.5 units μL^{-1} *Taq* polymerase and 50 ng genomic DNA. A final reaction volume of 15 µL was achieved by adding ddH_2O. The touchdown PCR protocol was used and the cycling was as follows: activation of the polymerase for 1 min at 95 °C followed by twenty cycles of 94 °C for 1 min, annealing at 65 °C for 1 min, and extension at 72 °C for 2 min (the annealing temperature was decreased by 1 °C every 2 cycles down to 56 °C), then 21 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and extension at 72 °C for 2 min. PCR products were stored at 4 °C until electrophoresis in 4% MetaPhor® agarose gel. Gels were stained with ethidium bromide, photographed and scored visually.

Genotypic analysis and genetic map construction

The population structure most commonly used for QTL mapping possesses sufficient polymorphism and linkage disequilibrium to generate linkage maps suitable to scan the whole genome. However, when the parents of the mapping population are highly related, the polymorphism between the parents may not be sufficient to generate linkage maps that evaluate the entire genome. For instance, a genetic map with two or more linkage groups per chromosome has been generated for the study of prolificacy in the Golden Glow maize population (de Leon et al. 2005). In the present study, since the parents of the RILs population were highly related, several genomic regions were monomorphic, leading to a

genetic map with more than one linkage group per chromosome (Figure 2). These monomorphic loci could arise from regions that are identical by descent (IBD) between C103 and C103AP or regions that are alike in state between these inbreds. Since the AP parent was not available for evaluation, the monomorphic regions could not be resolved to be IBD or alike in state.

At each SSR locus, a chi-square goodness-of-fit test was performed to detect deviations from the expected 1:1 genotypic segregation ratio. The genetic map was constructed using MAPMAKER Version 3.0 (Lander et al. 1987). The SSR loci were placed in linkage groups based on a LOD score of 3 and a maximum Haldane distance of 50 cM. For each linkage group, the 'three point' command was used to pre-compute the likelihoods of all three-point combinations. The LOD threshold to exclude a three-point order from a linkage group was set to 4. The 'order' command was used to find the best linear order of loci in a given linkage group based on the informativeness criteria (2 cM minimum distance between any pair of loci, the 218 individuals being informative for each locus and loci being codominant). Since the 'order' command adds loci sequentially, once the linkage map was obtained, it was tested using the 'ripple' command and the best loci order(s) was defined based on a LOD threshold of 2. Unplaced markers were located using the 'build' command. Mistyping or misclassification in the genotyping was detected by the 'triple error detection on' command.

QTL detection

Composite Interval Mapping (CIM) was initially used to detect regions associated with each trait evaluated in each environment and in the combined analysis using WinQTL cartographer (Wang et al. 2007; Zeng 1994). The empirical LOD threshold for CIM was

determined by 1000 permutations for an overall significance level of 5% (Churchill and Doerge 1994). For the combined analysis, the LOD thresholds defined by the permutation tests were: 2.7 for EPP, 2.8 for GDU and 2.6 for PH. The forward and backward stepwise selection criterion was used in the regression model to select cofactors accounting for genetic background variation when mapping QTL by interval test ($p_{entry/remove} \leq 0.1$). The window size on either side of the loci flanking the interval was defined as 10 cM.

Results obtained from CIM were used to further explore the genetic regression models and search for epistatic interactions with Multiple Interval Mapping (MIM). The underlying theoretical bases to define the best genetic model in MIM were described by Kao and Zeng (1997), Kao et al. (1999) and Zeng et al. (1999). The Hannan and Quinn (1979) Bayesian Information Criteria (BIC) with the penalty function $c(n)= 2\ln(\ln(n))$, with n=218 was used to define the critical value for testing QTL effects and epistatic interactions to be included in the multiple regression genetic model.

The QTL detected by CIM were included in the MIM model if the LOD threshold was >2.4 and the minimum distance between QTL was 5 cM defining an initial genetic model. This initial model was tested for significance in MIM based on the critical value (Hannan and Quinn 1979) and the significant effects were retained. A new search for QTL was performed and the genetic map was scanned to determine the best position of the QTL. At that position, a likelihood ratio test was performed to retain or delete this QTL from the model based on the critical value. A search for epistatic effects was conducted among those pair-wise interaction terms not yet included in the model. To determine if an interaction term remained in the model, a likelihood ratio test on the effect was performed. The process was repeated until no more significant epistatic effects were found. Each QTL in the MIM genetic

model was re-evaluated for its significance. If the LOD for a QTL (marginal or epistatic) effect fell below the significance threshold conditioned on the other effects, the effect was removed from the model. However, if the marginal effect of a QTL, that has a significant epistatic effect with other QTL, fell below the threshold; this marginal effect was retained. Finally, the position of each QTL was optimized in a sequential order by evaluating the model likelihood for possible positions conditional on the positions of other QTL in the model. The position that maximized the model likelihood was chosen as the new position of the QTL (Tao et al. 2003).

Once the final genetic model was defined, the genetic variance was estimated by the sum of squares of the final QTL model; and the phenotypic variance was estimated by the total sum of squares. The estimate of the broad-sense heritability was approximated by the coefficient of determination (R^2) of the MIM model (Kao et al. 1999). The proportion of the total genetic variance explained by the final QTL model was obtained by dividing R^2 by the progeny mean broad-sense heritability of the trait (i.e. R^2/H^2) (Dudley 1993).

Several regions of the C103 and C103AP genomes were monomorphic. Chromosome 5 was evaluated at 90 SSR loci but only 4 of them were polymorphic (bins 5.00, 5.01 and 5.07; see Gardiner et al. 1993 for bin definition). Similar results were obtained for chromosome 10, in which only two of 54 SSR loci were polymorphic (bin 10.06). Since these loci could not be assembled into linkage groups, a reasonable alternative for simple detection of QTL linked to a locus was to perform single-marker analysis. Under this method, the distribution of the phenotypic values for each trait was examined separately for each locus and likelihood ratio tests were performed to detect association between the phenotypic variation and genotypic classes at each such locus (Wang et al. 2007).

Multiple Trait Composite Interval Mapping (Mt-CIM) was used to test QTL × environment interaction for each trait (Jiang and Zeng 1995). The empirical LOD threshold for Mt-CIM was determined by 1000 permutations for an overall significance level of 5% (Churchill and Doerge 1994). Hence, the LOD thresholds were: 4.3 for EPP and GDU, and 4.2 for PH. The forward and backward stepwise selection criterion was used to select cofactors accounting for genetic background variation when mapping QTL by the interval test ($p_{entry/remove} \leq 0.1$). The window size on either side of the flanking loci was defined as 10 cM. By testing the QTL × environment interaction, the null hypothesis of $b_1 = b_2 = \ldots = b_j$ at the chromosome regions where QTL have been suggested by the joint mapping is being tested; where b_j corresponds to the additive effect of a QTL in the jth environment (Jiang and Zeng 1995).

Assessment of pleiotropy and linkage for QTL

For each QTL, a one-LOD support interval was constructed (Lander and Botstein 1989). On a linkage group, QTL with non-overlapping one-LOD support intervals were considered as different QTL (Conneally et al. 1985, Lander and Bolstein 1989). Multiple-trait MIM (Mt-MIM) was used to perform a likelihood ratio test for the hypothesis of pleiotropy versus close linkage between two QTL with overlapping support intervals (Jiang and Zeng 1995).

Each QTL with an effect on one trait only, have positions symbolically specified by p_1 for the QTL having an effect on trait 1 and p_2 for the QTL having an effect on trait 2. If the two positions are in the same interval defined by SSR loci, p_1 and p_2 are then defined as the ratios of the recombination frequencies between a locus and the two positions, respectively, and between the two flanking SSR loci (Jiang and Zeng 1995). The hypotheses

can be formulated as H_0: $p_1 = p_2$, H_1: $p_1 \neq p_2$ where p_1 corresponds to position of the QTL for trait 1 and p_2 correspond to position of the QTL for trait 2. The test statistic is the likelihood ratio LR = -2 ln(L_0/L_1), where L_1 is the maximum of the likelihood for the alternative model (given by H_1) and L_0 is the maximum of the likelihood that corresponds to the null hypothesis of $p_1 = p_2$. The LR test statistic for fixed testing position (i.e., for fixed p) follows an approximate chi-square distribution with 1 degree of freedom (Jiang and Zeng 1995). The graphical display of the genetic map and QTL locations was constructed using MapChart v2.2 (Voorrips 2002).

RESULTS

Phenotypic analysis

C103AP had more ears per plant, was later flowering and taller than C103AP; similar to previous observations of their respective hybrid testcrosses at this planting density (Duvick 1974; Table 2). Significant differences were observed between C103 and C103AP for number of ears per plant (EPP) as well as for plant height (PH). The mean of the RILs tended to be closer to C103 with respect to GDU and PH. C103AP differed from the mean of the RILs for PH while both parents differed from the mean of RILs for EPP (Table 2). A similar trend in the phenotypic values of the parents and the RILs was observed in each year. For EPP, C103 and C103AP were located at the extremes of the population distribution; consequently none of the RILs exceeded the parental phenotype for this trait (Figure 1). However, for GDU and PH several lines consistently had lower values than C103 each year and in the combined analysis. A plausible explanation for the presence of these extreme phenotypes would be the

segregation, in individual genotypes, of combination of alleles from both parents that have effects in the same direction (complementary gene action; Rieseberg et al. 2003). This explanation for GDU and PH was supported by QTL results herein (Table 4).

Genetic and genotype × environment (G × E) variance components were significant for all traits (Table 2). Spearman rank-correlation coefficients varied from 0.62 (EPP for 2005-2006) to 0.90 (GDU for 2006-2007), suggesting that changes in the magnitude and ranking of the genotypes were involved in the G × E variance detected for EPP while changes in magnitude of the estimates were more influential for GDU and PH. The broad-sense heritability values, all greater than 0.80 for all traits, facilitate the detection of QTL with relatively large effect on the genetic variance of these traits, reduce their confidence intervals and thus improve the detection of linked QTL (Table 2; Kearsey 1998; Lande and Thompson 1990).

Genotypic analysis and genetic map construction

Considering that C103AP was developed after 4 backcrosses with selection for the C103 phenotype, C103 and C103AP were expected to be highly related (F=0.97). This degree of relatedness was reflected in the level of polymorphic SSR loci detected. When both inbreds were evaluated at 890 SSR loci, 174 loci were polymorphic. As a consequence, more than one linkage group was defined per chromosome. A distance of 23.5 cM between two loci was used as the minimum threshold to define a linkage group to avoid the detection of false QTL positions in regions of the genetic map that were not sufficiently mapped. The genetic map covered a total of 894 cM of the maize genome with an average distance of 7.3 cM between two loci (Figure 2). Linkage maps for chromosomes 5, 7 and 10

were not presented because very few polymorphic loci (4 loci in chromosome 5 and 2 loci in chromosome 10) were detected for those regions of the genome.

Forty-seven loci showed a significant departure from the expected 1:1 genotypic ratio. On chromosome 2 and 8, the loci with a significant departure from the 1:1 genotypic ratio had an excess of C103 alleles. On chromosome 4, the 6 loci with significant departure from the 1:1 genotypic ratio and a region in chromosome 9, from locus *umc1037* to locus *umc1571*, had an excess of C103AP alleles. Single-locus segregation distortion skews the genotypic frequencies from the expected mendelian segregation ratio and introduces a bias in the estimation of recombination fractions (Bailey 1949; Lorieux et al. 1995). Therefore, the segregation distortion observed in these regions could affect the position of the QTL detected.

QTL detection

The use of RILs allows the evaluation of the population in different environments, providing an estimation of the QTL effects, the environment and their interaction (Burr and Burr 1991, Lynch and Walsh 1998). Nine QTL were detected in the combined analysis for EPP (Table 4). The QTL in bins 6.06 and 8.04 accounted for 21.9% of the phenotypic variation for EPP. In bin 6.06, C103 increased EPP by 0.11 while, in bin 8.04, C103AP increased EPP by 0.16. QTL in or adjacent to, those bins and with the same parental allele were detected in each year (Table 4). From single-marker analysis, QTL for EPP were detected at locus *umc1325* (bin 5.00) in 2005 and locus *bnlg1028* (bin 10.06) in 2007 ($p<0.05$). QTL linked to locus *umc1766* (bin 5.01) were detected in each year of evaluation ($p<0.01$).

Seven QTL for GDU to anthesis accounting for 76% of the phenotypic variance were detected in the combined analysis; one of them (bin 8.05) explained 61% of the total genetic variance and was detected in all three years. C103AP alleles delayed anthesis by 28.1 GDU in the combined analysis (Table 4). Also, at locus *umc1766* of bin 5.01, genotypes significantly differed for GDU to anthesis in the combined analysis as well as in years 2005 and 2006, suggesting the presence of a QTL linked to this locus.

For plant height, 8 QTL were detected in the combined environment, accounting for 54.9% of the phenotypic variance (Table 4). The QTL of bin 8.05 most strongly contributed to the total genetic variance (15.4%) and was detected in all years. The C103AP alleles at that QTL had a positive additive effect for this trait (Table 4). Genotypic differences for plant height were detected at a QTL linked to locus *bnlg1028* (bin 10.06) in 2005 ($p<0.01$) and in the combined analysis ($p<0.05$).

With respect to EPP, the contribution of epistatic interactions to the genetic variance of the QTL model ranged from 2.5%, in 2007, to 7.6%, in 2006. Only one epistatic interaction was detected in 2005 for GDU to anthesis and accounted for 4.6% of the genetic variance. Epistatic interactions between bin 1.02 and bins 8.05-8.06 were also detected in all environments (Table 5). For PH, 5 epistatic interactions were detected in the combined analysis. Among them, additive × additive interactions between QTL 3 (bin 2.06) and 7 (bin 8.05); and QTL 1 (bin1.02) and 8 (bin 9.03), contributed the most to the genetic variance (7.8 %; Table 5).

Environmental conditions differed among the three years of evaluation; while GDUs were similar, there was relatively more variation for precipitation. The amount of rainfall in 2006 was lower than the two other seasons and mostly concentrated in July

(Table 1). These environmental conditions may have had an effect on the estimates of the variance components and the broad-sense heritability values. The magnitude of the genetic variance estimates was higher in 2005 for EPP and GDU when compared to the estimates obtained in 2006 and 2007. The contribution of the G × E variance component to the phenotypic variance was relatively low for all traits (Table 2). However, this environmental effect along with the G × E interaction was reflected in the consistency of QTL and epistatic interactions detected. The QTL with major effect on the genotypic variance detected for EPP (QTL 6 and 7), GDU (QTL 6) and PH (QTL 7) were detected in all years. For EPP, QTL 2, 6 and 7 identified in 2005 were no identified in similar location in 2006 or 2007. The same situation was observed for QTL 1, 2, 4, 5 and 8 from 2006, and QTL 5 and 6 from 2007. With respect to GDU, only 2 QTL detected in 2006 (QTL 2 and 5) and a QTL detected in 2005 (QTL 3) were not detected in similar locations in the other years. For PH, QTL 7 from 2005 was not detected in similar location in 2006 and 2007 (Table 4).

Assessment of pleiotropy and linkage for QTL

Phenotypic correlations between two traits can arise from environmental, developmental or genetic factors. The expression of two traits may be modified by the same environmental factors affecting the allocation of resources during development, either by influencing both traits in the same or opposite directions. Likewise, genetic correlations can arise either by close linkage or pleiotropy between loci affecting both traits. Positive phenotypic correlation between EPP, GDU and PH were detected on a previous study of testcross hybrids of C103 and C103AP (Duvick 1974). In this study, the phenotypic (r_p) correlations between EPP and GDU, and EPP and PH, were lower

than 0.3 in the combined analysis (Table 3). Also, GDU and PH had a low phenotypic correlation (r_P=0.25). Weak positive genetic correlations were observed for EPP and GDU, and EPP and PH height (r_G<0.3) (Table 3).

As suggested by their locations on the genetic map and by statistical tests (Jiang and Zeng 1995) there was statistical evidence of pleiotropy for EPP, GDU and PH in four genomic regions on chromosomes 3, 6, 8 and 9 (Figure 2 and Table 4). Evidence favoring pleiotropy was found for EPP and PH (p=0.3) in bin 3.10 and for EPP and GDU in bin 6.06 (p=0.5). Based on the 1-LOD supporting interval and the subsequent statistical test on the QTL positions, evidence favoring pleiotropy was detected for GDU (QTL 6), PH (QTL 7) and EPP (QTL 7) (p>0.1) in bin 8.05. Also, statistical evidence of pleiotropy was detected for PH (QTL 8) and EPP (QTL 9) in bin 9.03 (p=0.8). Although no statistical evidence of linkage between these traits was detected on chromosomes 3, 6, 8 and 9, these results cannot discard the presence of linkage.

The statistical evidence supporting the hypothesis of pleiotropic gene action concurs with previous observations that C103AP contained genetic factors that increased EPP, GDU to anthesis and PH relative to C103.

DISCUSSION

Prolificacy, GDU to anthesis, and plant height were investigated in a population of recombinant inbred lines (RILs) derived from two near-isogenic inbreds, C103 and C103AP. As a consequence of the highly isogenic background detected in the genotypic evaluation of both parents, a genetic map with more than one linkage group per chromosome was constructed covering only smaller regions at which parental alleles were segregating (Figure

2). Given the nature of the population, a search for QTL at fewer genomic regions was conducted. The number of QTL detected in the combined analysis of this population of RILs ranged from 5, for GDU in 2005 and 2006, to 9 QTL for EPP (Table 4).

A QTL with major effect on the genetic variance of EPP was detected in bin 8.04 (Table 4). From the maize database, QTL for EPP have not been detected in this bin, suggesting a new observation for this region. In a nearby region (bin 8.02), a QTL for EPP, *qepp6*, was detected in a related population (Veldboom and Lee 1996). Two similar genes of the MADS-box family, *ZMM18* and *ZMM29* are located in bin 8.04. These genes are expressed in male and female inflorescences and in developing kernels (Münster et al. 2001). In flowering plants MADS-box genes form a multigene family that controls diverse developmental processes including root, flower and fruit development. Recently, Danilevskaya et al. (2008) suggested that *ZMM4*, another member of the MADS-box family, may play roles in both floral induction and inflorescence development. Given the importance of these genes in developmental processes, it can be hypothesized that *ZMM18* and *ZMM29* directly or indirectly may influence the initiation, maintenance and growth of axillary meristems that constitute secondary ears in maize. However, detailed studies on these genes are needed to assess their candidacy for the development of secondary ears.

In the vicinity of bin 6.06, several QTL for EPP, *qepp2*, *qepp9, qepp13, qepp17, qepp20* and *qepp23*, were detected in studies of the Mo17/H99 population (Austin and Lee 1998; Veldboom and Lee 1996). In this study, the QTL for EPP in bin 6.06, add more evidence to the effect of this region in the development of axillary meristems in other environments and genetic contexts.

The major QTL detected in this study for EPP, QTL 6 and 7 were located to regions which are not described as containing candidate genes or other DNA sequences known to affect meristem initiation and maintenance in maize (McSteen and Leyser 2005). Such QTL suggest that other regions in the maize genome may also affect the control of axillary meristems. According to information from the maize database, other regions detected in this study are novel with respect to EPP; those include QTL 2, 4, 5, 7 and 9 (Table 4).

Even though the extent to which epistasis is involved in regulating complex traits is not known, the identification of these interactions for QTL may be a valuable starting point for a more thorough understanding of the genetic networks leading to the expression of such phenotype (Carlborg and Haley 2004). Several other studies demonstrate the importance of epistasis in the genetic variance of quantitative traits among progeny of line crosses (Carlborg et al. 2006; Dudley 2008; Malmberg et al. 2005; Peripato et al. 2002).

QTL mapping studies using conventional segregating populations have generally uncovered little evidence for epistasis (Holland 2007). Near Isogenic Lines may facilitate the detection of epistatic gene action of QTL because of an increase the power of detection of such interactions when compared to F_2 populations (Zamir and Eshed 1998). In the present study, additive × additive epistatic interactions between QTL that contribute in a range of 1.8 to 12.3 % to the genetic variance were detected in the combined analysis (Table 5). de Leon et al. (2005) detected additive by dominant epistatic interactions for EPP. As these authors suggested, the lack of detection of other type of interaction in that study could be a consequence of the low statistical power of the experimental setting and sample size rather than absence of such genetic effects.

Based on studies with double mutants, few interacting loci have been described and assembled in the shoot meristem network (Barazesh and McSteen 2008, Wu and McSteen 2007). Herein, only the epistatic interaction between QTL 1 and QTL 4 in 2005 involves a region (bin 2.04) in which *ba2*, a candidate gene for shoot development, was described (Table 5).

Several studies have investigated consistency of QTL across environments; however, it is likely that inconsistent QTL may represent the genetic factors responsible for genotype by environment interaction (Beavis and Keim 1996). The QTL with major effect on the genotypic variance detected for EPP, GDU and PH were detected in all environments. Nevertheless, QTL with minor effects on the genetic variance were not detected in some environments (i.e. for EPP, QTL 7 from 2005 and QTL 6 from 2007 were not located in the same position or close to in the combined environment) (Table 4). The inconsistent detection of these QTL could be due to QTL by environment interaction that eventually could contribute to the significant G x E variance (Table 2). Given the influence that environmental conditions may have in the initiation, maintenance and growth of the apical meristem, inconsistent QTL may contribute to the understanding and the regulation of this developmental pathway.

Knowledge of the mechanisms underlying the correlations between different traits is fundamental to understanding the degree of integration of the phenotype and to resolving constraints imposed on evolutionary processes (Lynch and Walsh 1998). Correlations between traits may have resulted from several sources, such as environmental conditions, shared developmental processes, alternative splicing, pleiotropy or close linkage between genetic factors. The low genetic correlations between EPP, GDU and PH (r_G<0.30) suggested

that few loci may be in close linkage or pleiotropic for those traits. However, four genomic regions, in bins 3.10, 6.06, 8.05 and 9.03 showed statistical evidence supporting the hypothesis of pleiotropy and had overlapping 1-LOD support intervals (Figure 2). The direction of pleiotropy may differ among genetic factors and, at least in principle, strong pleiotropy need not to result in a strong genetic correlation between traits since the pleiotropic effects from different loci may cancel each other (Lynch and Walsh 1998). Nevertheless, an improvement in the mapping resolution would be needed for resolving linkage and pleiotropy. Identification of recombinant progeny would suggest that the two traits were in close linkage. With respect to pleiotropy, to determine that this mechanism is the cause of the association between two traits would imply the identification and manipulation of the DNA sequence associated with the variation for those traits.

The transition from vegetative to reproductive growth affects male and female flowering time, leaf number and correlatively plant height (Irish and Nelson 1991). A very influential region, flanked by loci *umc1858* and *umc1309*, was located to bins 8.04-8.05 with major effects on EPP, GDU to anthesis and PH. Alleles from C103AP increased EPP, plant height and the GDUs to anthesis. Two QTL that affect transition from vegetative to generative states, *Vgt1* and *Vgt2* have been located to bins 8.05 and 8.04, respectively (Salvi et al. 2007; Salvi et al. 2002; Vladutu et al. 1999; Chardon et al. 2005). The phenotypic effect of *Vgt1* and *vgt2* are a reduction in days to anthesis, plant height and node number. The major QTL detected for GDU to anthesis in this study was located to the same bin as for *Vgt1*. Salvi et al. (2007) resolved *Vgt1* to a non-coding region positioned 70 kb upstream of an *Apetala2*-like transcription factor, *ZmRap2.7* and suggested that *Vgt1* acts as a cis-regulatory element of *ZmRap2.7*. The molecular mechanism of *Vgt1* action on *ZmRap2.7*

expression was not predicted. However, confirmation of the role of *Vgt1* in flowering time was obtained in an association study of 375 maize inbred lines (Ducrocq et al. 2008). Association of *Vgt1* and *vgt2* with the number of ears per plant was not described in these studies. Therefore, a novel observation associating GDU to anthesis and EPP is described herein for this genetic context.

Another region that deserves further study is in bin 6.06, in which statistical evidence of pleiotropy was found for EPP and GDU. In the vicinity of this bin, 6 QTL associated with EPP and QTL affecting heat units to pollen shed have been detected (www.maizegdb.org; verified 02/15/08).

A region in bin 3.10, where QTL for EPP and PH were identified, was unique relative to previous reports in maize as summarized in the maize database (www.maizegdb.org; verified 02/15/08). Auxins have been proposed as an essential regulator in the initiation and development of the axillary meristem mediated by the action of cytokinins (Wang and Li 2008). Recently an additional carotenoid-derived-long-range signaling pathway involved in the outgrowth of axillary meristems has been proposed by studies on *MAX* genes from *Arabidopsis*. One gene, *MAX1*, encodes a cytochrome P450-type enzyme. Several homologs of *MAX1* have been found in rice but no maize ortholog has been described in the literature (Schmitz and Theres 2005; Wang and Li 2008). Nevertheless, a homolog of cytochrome P450 (*cyp1*) was located in bin 3.10 (Chao et al. 1994) that might be involved in the expression of plant height and development of secondary ears. Further research is needed for a better understanding the potential role of this carotenoid pathway in the hormonal regulation of axillary meristem in maize as well as its relationship with plant height.

At least 5 novel regions for EPP were detected on chromosomes 2, 3, 6, 8 and 9. A region on chromosome 8, flanked by loci *umc1858* and *umc1309*, had a major influence not only on the increase of number of ears per plant but also on plant height and GDU to anthesis. A genomic region in bin 6.06 was pleiotropic for a QTL associated with flowering time and EPP. Two major QTL were detected for flowering time in accordance with previously described and cloned QTL. A QTL that increased plant height in bin 3.10 was consistently detected across environments and has not been described previously in the literature. The evidence of pleiotropic gene action for the traits evaluated here support previous hypothesis of regions in C103AP affecting EPP, GDU to anthesis and plant height (Duvick 1974).

REFERENCES

Agrama HAS, Moussa ME (1996) Mapping QTLs in breeding for drought tolerance in maize (Zea mays L.). Euphytica 91:89-97

Austin DF, Lee M (1996) Comparative mapping in $F_{2:3}$ and $F_{6:7}$ generations of quantitative trait loci for grain yield and yield components in maize. Theor Appl Genet 92:817-826

Austin DF, Lee M (1998) Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. Crop Sci 38:1296-1308

Bailey N (1949) The estimation of linkage with differential viability, II and III. Heredity 3:220-228

Barazesh S, McSteen P (2008) *barren inflorescence1* functions in organogenesis during vegetative and inflorescence development in maize. Genetics 179:389-401

Beavis W, Keim P (1996) Identification of quantitative trait loci that are affected by environment. In: Kang MS, Gauch HG (eds) Genotype-by-environment interaction. CRC Press LLC

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

Burr B, Burr FA (1991) Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. Trends Genet 7:55-60

Carlborg O, Haley CS (2004) Epistasis: too often neglected in complex trait studies? Nat Rev Genet 5:618-625

Carlborg O, Jacobsson L, Ahgren P, Siegel P, Andersson L (2006) Epistasis and the release of genetic variation during long-term selection. Nat Genet 38:418-420

Chao S, Baysdorfer C, Heredia-Diaz O, Musket T, Xu G, Coe EH (1994) RFLP mapping of partially sequenced leaf cDNA clones in maize. Theor Appl Genet 88:717-721

Chardon F, Hourcade D, Combes V, Charcosset A (2005) Mapping of a spontaneous mutation for early flowering time in maize highlights contrasting allelic series at two-linked QTL on chromosome 8. Theor Appl Genet 112:1-11

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971

Conneally PM, Edwards JH, Kidd KK, Lalouel J-M, Morton NE, Ott J, White R (1985) Report of the committee on methods of linkage analysis and reporting. Cytogen. Cell Genet. 40:356-359

Cross HZ, Zuber MS (1972) Prediction of flowering dates in maize based on different methods of estimating thermal units. Agron J 64:351-355

Danilevskaya ON, Meng X, Selinger DA, Deschamps S, Hermon P, Vansant G, Gupta R, Ananiev EV, Muszynski MG (2008) Involvement of the MADS-Box Gene *ZMM4* in floral induction and inflorescence development in maize. Plant Physiol 147:2054-2069

de Leon N, Coors JG (2002) Twenty-four cycles of mass selection for prolificacy in the Golden Glow maize population. Crop Sci 42:325-333

de Leon N, Coors JG, Kaeppler SM (2005) Genetic control of prolificacy and related traits in the Golden Glow maize population: II. Genotypic analysis. Crop Sci 45:1370-1378

Ducrocq S, Madur D, Veyrieras J-B, Camus-Kulandaivelu L, Kloiber-Maitz M, Presterl T, Ouzunova M, Manicacci D, Charcosset A (2008) Key impact of *Vgt1* on flowering time adaptation in maize: evidence from association mapping and ecogeographical information. Genetics 178:2433-2437

Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci 33:660-668

Dudley JW (2008) Epistatic interactions in crosses of Illinois High Oil x Illinois Low Oil and of Illinois High Protein x Illinois Low Protein corn strains. Crop Sci 48:59-68

Duvick DN (1974) Continuous backcrossing to transfer prolificacy to a singled-eared inbred line of maize. Crop Sci 14:69-71

Eshed Y, Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141:1147-1162

Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Enrico Pe M, Schmidt RJ (2004) The role of *barren stalk1* in the architecture of maize. Nature 432:630-635

Gardiner JM, Coe EH, Melia-Hancock S, Hoisington DA, Chao S (1993) Development of a core RFLP map in maize using an immortalized F_2 population. Genetics 134:917-930

Graham GI, Wolff DW, Stuber CW (1997) Characterization of a yield Quantitative Trait Locus on chromosome five of maize by fine mapping. Crop Sci 37:1601-1610

Hallauer AR, Miranda Fo. JB (1988) Quantitative genetics in maize breeding, Second edn. Iowa State University Press, Ames, IA

Hannan EJ, Quinn BG (1979) The determination of the order of an autoregression. J R Soc Stat 41:190-195

Holland JB, Nyquist WE, Cervantez-Martinez CT (2003) Estimating and interpreting heritability for plant breeding: an update. Plant Breeding Rev 22:9-112

Holland JB (2006) Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. Crop Sci 46:642-654

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

Irish EE, Nelson TM (1991) Identification of multiple stages in the conversion of vegetative to floral development. Development 112:891-898

Jampatong S, Darrah LL, Krause GF, Barry BD (2000) Effect of one- and two-eared selection on stalk strength and other characters in maize. Crop Sci 40:605-611

Jiang C, Zeng ZB (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111-1127

Jones DF, Everett HL (1949) Hybrid field corn. Conn Agric Exp Stn Bull, New Haven, CT

Kao C-H, Zeng Z-B (1997) General formulas for obtaining the MLEs and the asymptotic variance- covariance matrix in mapping quantitative trait loci when using the EM algorithm. Biometrics 53:653-665

Kao C-H, Zeng Z-B, Teasdale RD (1999) Multiple Interval Mapping for Quantitative Trait Loci. Genetics 152:1203-1216

Kearsey M (1998) The principles of QTL analysis (a minimal mathematics approach). J Exp Bot 49:1619-1623

Kenward MG, Roger JH (1997) Small sample inference for fixed effects from Restricted Maximum Likelihood. Biometrics 53:983-997

Lande R, Thompson R (1990) Efficiency of Marker-Assisted Selection in the improvement of quantitative traits. Genetics 124:743-756

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

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181

Lauter N, Doebley J (2002) Genetic variation for phenotypically invariant traits detected in teosinte: implications for the evolution of novel forms. Genetics 160:333-342

Levene H (1960) Contributions to probability and statistics: essays in honor of Harold Hotelling. In: Olkin Iea (ed). Stanford University Press, pp 278-292.

Littel RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O (2006) SAS for mixed models, second edition. SAS Institute Inc. , Cary, NC

Lorieux M, Goffinet B, Perrier X, León DG, Lanaud C (1995) Maximum-likelihood models for mapping genetic markers showing segregation distortion. 1. Backcross populations. Theor Appl Genet 90:73-80

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

Malcomber ST, Preston JC, Reinheimer R, Kossuth JE, Kellogg EA (2006) Developmental gene evolution and the origin of grass inflorescence diversity. Adv Bot Res 44:426-481

Malmberg RL, Held S, Waits A, Mauricio R (2005) Epistasis for fitness-related quantitative traits in Arabidopsis thaliana grown in the field and in the greenhouse. Genetics 171:2013-2027

McSteen P, Hake S (2001) *barren inflorescence2* regulates axillary meristem development in the maize inflorescence. Development 128:2881-2891

McSteen P, Leyser O (2005) Shoot Branching. Annu Rev Plant Biol 56:353-374

Motto M, Moll RH (1983) Prolificacy in maize: a review. Maydica 28:53-76

Muehlbauer GJ, Specht JE, Thomas-Compton MA, Staswick PE, Bernard RL (1988) Near-Isogenic Lines a potential resource in the integration of conventional and molecular marker linkage maps. Crop Sci 28:729-735

Münster T, Ursula Wingen L, Faigl W, Werth S, Saedler H, Theißen G (2001) Characterization of three GLOBOSA-like MADS-box genes from maize: evidence for ancient paralogy in one class of floral homeotic B-function genes of grasses. Gene 262:1-13

Peripato AC, de Brito RA, Vaughn TT, Pletscher LS, Matioli SR, Cheverud JM (2002) Quantitative Trait Loci for maternal performance for offspring survival in mice. Genetics 162:1341-1353

Rieseberg LH, Widmer A, Arntz AM, Burke JM (2003) The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. Philos T R Soc B 358:1141-1147

Ritchie SW, Hanway JJ, Benson GO (1996) How a corn plants develops. In: Serv ISU CE (ed) Spec Rep. Iowa State University, Ames, IA.

Ritter MK, Padilla CM, Schmidt RJ (2002) The maize mutant *barren stalk1* is defective in axillary meristem development. Am J Bot 89:203-210

Russell WA (1968) Testcrosses of one- and two-ear types of corn belt maize inbreds. I. performance at four plant stand densities. Crop Sci 8:244-247

Russell WA, Eberhart SA (1968) Testcrosses of one- and two-ear types of corn belt maize inbreds. II. stability of performance in different environments. Crop Sci 8:248-251

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci U S A 81:8014-8018

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 G-H, Phillips RL, Tuberosa R (2007) Conserved non-coding genomic sequences associated with a flowering-time Quantitative Trait Locus in maize. Proc Natl Acad Sci 104:11376-11381

Salvi S, Tuberosa R, Chiapparino E, Maccaferri M, Veillet S, van Beuningen L, Isaac P, Edwards K, Phillips RL (2002) Toward positional cloning of *Vgt1*, a QTL controlling the transition from the vegetative to the reproductive phase in maize. Plant Mol Biol 48:601-613

Sanchez AC, Brar DS, Huang N, Li Z, Khush GS (2000) Sequence tagged site marker-assisted selection for three bacterial blight resistance genes in rice. Crop Sci 40:792-797

SAS (2007) SAS/STAT Software. 9.1.3 edn. SAS Institute Inc. , Cary, NC.

SAS (2008) The MIXED Procedure. SAS/STAT 9.2 User's Guide. SAS Institute Inc., Cary, NC., pp 2559-2851

Sass JE, Loeffel FA (1959) Development of axillary buds in maize in relation to barrenness. Agron J 51:484-486

Schmitz G, Theres K (2005) Shoot and inflorescence branching. Curr Opin Plant Biol 8:506-511

Stucker RE, Troyer AF (1977) Some observations on inheritance of prolificacy. MNL 51:80-81

Szalma S, Hostert B, LeDeaux J, Stuber C, Holland J (2007) QTL mapping with Near-Isogenic Lines in maize. Theor Appl Genet 114:1211-1228

Tao Y, Zeng Z-B, Li J, Hartl DL, Laurie CC (2003) Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. II. Mapping hybrid male sterility loci on the third chromosome. Genetics 164:1399-1418

Veldboom LR, Lee M (1994) Molecular-marker-facilitated studies of morphological traits in maize. II: Determination of QTLs for grain yield and yield components. Theor Appl Genet 89:451-458

Veldboom LR, Lee M (1996) Genetic mapping of Quantitative Trait Loci in maize in stress and nonstress environments: I. grain yield and yield components. Crop Sci 36:1310-1319

Vladutu C, McLaughlin J, Phillips RL (1999) Fine mapping and characterization of linked quantitative trait loci involved in the transition of the maize apical meristem from vegetative to generative structures. Genetics 153:993-1007

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Hered 93:77-78

Wang S, Basten CJ, Zeng Z-B (2007) Windows QTL Cartographer 2.5 edn. Department of Statistics, North Carolina State University, Raleigh, NC.

Wang Y, Li J (2008) Molecular basis of plant architecture. Annu Rev Plant Biol 59

Wu X, McSteen P (2007) The role of auxin transport during inflorescence development in maize (*Zea mays*, Poaceae). Am J Bot 94:1745-1755

Zamir D, Eshed D (1998) Case history in germplasm introgression: tomato genetics and breeding using nearly isogenic introgression lines derived from wild species. In: Paterson A (ed) Molecular dissection of complex traits. CrC Press LLC, New York, pp 207-217

Zeng Z-B. KC-H, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. Genet Res Camb 74:11

Zeng ZB (1994) Precision mapping of Quantitative Trait Loci. Genetics 136:1457-1468

Table 1: Monthly precipitation and GDU accumulation for the period 1987-2007 and for the growing season 2005-2007 at the Agronomy and Agricultural Engineering Research Center (AAERC) near Boone, Iowa (42°1´ N, 93°45´ W)

	GDU accumulation (ºC)[a]				Precipitation (mm)			
	2005	2006	2007	1987-2007	2005	2006	2007	1987-2007
	Monthly accumulation			Average[b]	Monthly accumulation			Average[b]
May	179	222	265	211	46.4	58.2	146.3	113.5
June	363	350	345	336	117.6	13.5	46.2	113.8
July	393	419	407	396	83.7	79.8	66.8	122.0
August	351	376	416	370	72.1	58.7	185.1	97.9

[a] Growing degree units were calculated according to Cross and Zuber (1972).
[b] Average monthly accumulation for the period 1987 to 2007.

Table 2: Mean phenotypic values of C103, C103AP and 218 maize RILs and variance components for ears per plant (EPP), GDU to anthesis and plant height in the combined analysis of the 2005, 2006 and 2007 datasets

	EPP	GDU to anthesis	Plant height
	------ number ------	--------- gdu[a] ----------	------- cm -------
		Mean Phenotypic Values[b]	
C103	0.88 ± 0.43	792.71 ± 34.22	170.66 ± 14.12
C103AP	2.94 ± 0.41***[b]	867.63 ± 33.78	180.61 ± 13.92*[b]
Mean RILs	1.96 ± 0.03*[b]	793.70 ± 2.29	168.11 ± 0.95**[b,c]
		Variance Components	
σ^2_G	0.16 ± 0.02***	1141.93 ± 114.14***	182.31 ±18.60 ***
$\sigma^2_{G\times E}$	0.03 ± 0.007***	52.22 ± 12.01***	21.40 ± 2.59 ***
σ^2	0.12 ± 0.006	197.16 ± 11.44	27.85 ± 1.62
		Broad Sense Heritability	
	0.84 ± 0.02	0.96 ± 0.005	0.93 ± 0.007

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
***Significant at the 0.001 probability level.
[a] Growing degree units according to Cross and Zuber (1972).
[b] Tukey multiple comparison test.
[c] Significant at the 0.001 probability level for the difference between C103AP and the mean of the population of RILs.

Table 3: Phenotypic (above diagonal) and genetic correlations with their standard errors in parenthesis for ears per plant (EPP), GDU to anthesis and plant height in a maize population of 218 RILs derived from inbreds C103 and C103AP in the combined analysis of the 2005, 2006 and 2007 datasets

	EPP	GDU to anthesis	Plant height
EPP	-	0.17 (0.05)	0.21 (0.05)
GDU to anthesis	0.26 (0.07)	-	0.26 (0.05)
Plant height	0.28 (0.07)	0.34 (0.06)	-

[a] Phenotypic and genetic correlation coefficients were calculated using REML from the MIXED procedure in SAS. Approximate sampling variances and standard errors for the estimates were obtained after Holland (2006).

Table 4: Summary of QTL detected from a population of 218 maize RILs derived from inbred lines C103 and C103AP for ears per plant (EPP), GDU to anthesis and plant height (PH)

Trait	Environment	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive Effect[d]	Parental Allele[e]	Partial $R^{2,f}$
							- number -		
EPP	Combined[h]	1	2.04	umc1026	6.86	4.16	0.09	AP	3.9
		2	2.08	umc1798	118.77	5.61	0.11	AP	3.8
		3	3.05	umc1167	0.01	1.75	0.05	AP	1.7
		4	3.10	bnlg1098	3.01	6.65	0.08	C103	6.4
		5	6.01	bnlg107	2.01	2.00	0.06	AP	2.5
		6	6.06	umc1912	2.01	7.72	0.11	C103	9.7
		7	8.04	umc1858	9.84	10.81	0.16	AP	12.2
		8	8.07	bnlg1823	87.70	2.94	0.07	C103	3.6
		9	9.03	mmc0051	25.47	5.03	0.09	AP	5.3
								Total $R^{2,g}$=	49.1
	2005	1	2.04	umc1448	11.14	4.45	0.08	AP	3.1
		2	2.08	bnlg198	90.92	2.74	0.08	AP	4.9
		3	3.10	umc1136	10.63	7.02	0.13	C103	11.4
		4	6.06	umc1859	5.73	4.36	0.08	C103	3.9
		5	8.04	umc1858	10.84	5.97	0.14	AP	8.5
		6	8.06	umc1724	73.22	2.07	0.07	C103	2.5

[a] Chromosomal bin corresponding to the location of the nearest left marker according to MaizeGDB database.
[b] Position in cM from the top of the chromosome calculated by MapMaker v.3
[c] LOD score corresponding to the position of the QTL as calculated by WinQTL cartographer.
[d] Additive effects values calculated as the average from the difference between homozygotes for each parental allele at a locus.
[e] Direction of the additive effect, AP= C103AP.
[f] Proportion of the phenotypic variance explained by each QTL defined in the multiple regression model.
[g] Proportion of the phenotypic variance explained by the sum of the QTL defined in the multiple regression model.
[h] Refers to the combined analysis of the 2005, 2006 and 2007 datasets.
[i] Growing degree units to anthesis calculated according to Cross and Zuber (1972)

Table 4 Cont.

Trait	Environment	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive Effect[d]	Parental Allele[e]	Partial $R^{2,f}$
		7	9.03	mmc0051	25.47	2.87	0.08	AP	3
								Total $R^{2,g}$=	37.3
	2006	1	2.08	bmc1233	106.19	5.20	0.12	AP	8.2
		2	3.05	umc1167	1.01	4.51	0.07	AP	4.1
		3	3.10	bnlg1098	1.01	5.63	0.08	C103	4.6
		4	4.08	bnlg2244	5.01	3.87	0.08	AP	4.6
		5	6.01	bnlg107	0.01	5.59	0.07	AP	3.3
		6	6.06	umc1912	0.01	10.83	0.13	C103	11.4
		7	8.05	umc1309	14.98	11.99	0.14	AP	10.7
		8	9.04	umc1492	64.27	5.46	0.07	AP	5.1
								Total $R^{2,g}$=	52
	2007	1	2.04	bnlg381	2.01	2.75	0.08	AP	3.1
		2	3.10	bnlg1098	5.01	2.36	0.06	C103	3.3
		3	6.06	umc1912	0.01	5.13	0.11	C103	7.6
		4	8.04	umc1858	8.84	2.95	0.10	AP	4.8
		5	8.06	umc1149	59.52	4.37	0.10	C103	4.9
		6	9.02	umc1037	14.44	2.53	0.08	AP	4.1
								Total $R^{2,g}$=	27.8
							-- gdu[i]--		
GDU	Combined[h]	1	1.02	umc1568	0.01	4.06	4.51	AP	1.4
		2	1.07	umc1833	50.97	4.34	5.63	C103	2
		3	4.08	bnlg2244	7.01	3.84	17.14	AP	2.2
		4	4.08	phi092	11.23	4.48	19.64	C103	4.4
		5	6.06	umc1912	2.01	6.94	7.01	C103	5.3
		6	8.05	bnlg1176	13.54	44.89	28.13	AP	58.2
		7	8.05-8.06	umc1665	44.21	3.20	2.28	AP	2.5
								Total $R^{2,g}$=	76
	2005	1	1.07	umc1833	49.97	3.19	5.45	C103	2.4
		2	4.08	bnlg2244	5.01	1.88	10.65	AP	2.0
		3	4.08	phi092	10.23	1.96	12.65	C103	1.8
		4	6.06	umc1912	3.01	7.14	5.94	C103	5.1
		5	8.05	bnlg1176	13.54	35.02	26.52	AP	50.4

Table 4 Cont.

Trait	Environment	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive Effect[d]	Parental Allele[e]	Partial $R^{2,f}$
								Total $R^{2,g}$=	61.7
	2006	1	1.02	umc1568	0.01	4.11	5.70	AP	2.1
		2	2.01	phi96100	0.01	2.58	3.87	AP	0.2
		3	6.06	umc1762	9.92	10.17	8.80	C103	6.8
		4	8.05	umc1309	14.98	45.76	27.8	AP	57.0
		5	9.04	umc1492	67.27	3.07	3.10	AP	0.7
								Total $R^{2,g}$=	66.8
	2007	1	1.02	umc1568	0.01	4.73	7.18	AP	2.4
		2	1.07	umc1833	51.97	2.54	4.82	C103	1.4
		3	4.08	phi092	10.23	2.83	4.94	C103	1.0
		4	6.06	umc1762	9.92	7.92	8.07	C103	5.7
		5	8.05	bnlg1176	13.54	7.69	22.55	AP	43.0
		6	8.05	bnlg666	17.71	3.00	8.46	AP	14.6
								Total $R^{2,g}$=	68.1
							-- cm --		
PH	Combined[h]	1	1.02	umc1976	14.62	4.16	2.38	AP	1.8
		2	1.03	phi001	64.93	6.57	3.78	C103	8.3
		3	2.06	bmc1329	68.18	4.92	4.37	C103	7.9
		4	3.06	umc1644	5.01	5.48	3.71	AP	7.5
		5	3.10	bnlg1098	5.01	7.60	3.52	C103	7.6
		6	4.06	umc1511	11.67	4.86	2.74	C103	4.3
		7	8.05	umc1309	14.98	12.62	5.34	AP	14.3
		8	9.03	mmc0051	25.47	5.6	2.52	AP	3.2
								Total $R^{2,g}$=	54.9
	2005	1	1.03	phi001	65.93	7.40	2.95	C103	7.9
		2	2.06	bmc1329	65.18	6.40	3.59	C103	7.4
		3	3.06	umc1644	8.01	6.17	3.44	AP	5.8
		4	3.10	bnlg1098	3.01	9.27	4.24	C103	10.6
		5	4.06	umc1511	15.67	3.72	2.58	C103	3.8
		6	8.05	umc1309	16.98	5.04	3.08	AP	4.5
		7	9.03	umc1688	4.01	3.98	3.67	AP	5.1
								Total $R^{2,g}$=	45.1

Table 4 Cont.

Trait	Environment	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive Effect[d]	Parental Allele[e]	Partial $R^{2,f}$
	2006	1	1.02	umc1976	19.62	4.55	1.11	AP	0.9
		2	1.03	phi001	64.93	4.03	3.43	C103	6.1
		3	2.06	bmc1329	61.18	7.14	2.52	C103	3.5
		4	3.06	umc1644	5.01	6.89	4.89	AP	9.7
		5	3.10	bnlg1098	4.01	5.48	3.79	C103	6.7
		6	8.05	umc1309	14.98	12.55	6.57	AP	18.3
		7	8.09	dupssr14	124.92	1.13	1.59	AP	1.2
								Total $R^{2,g}$=	46.4
	2007	1	1.03	umc1044	29.58	3.80	3.73	AP	1.3
		2	1.03	phi001	65.93	7.81	5.01	C103	7.8
		3	2.06	bmc1329	67.18	2.60	3.19	C103	4.1
		4	3.06	umc1644	5.01	4.64	3.87	AP	6.1
		5	3.10	bnlg1098	6.01	8.32	4.98	C103	9.5
		6	4.06	umc1511	11.67	5.80	3.26	C103	5.3
		7	8.05	umc1309	14.98	16.03	7.98	AP	20.7
		8	9.04	umc1571	23.74	3.14	2.93	AP	3.3
								Total $R^{2,g}$=	58.1

Table 5: Summary of epistatic interactions detected from a population of 218 maize RILs derived from inbred lines C103 and C103AP for ears per plant (EPP), GDU to anthesis and plant height (PH)

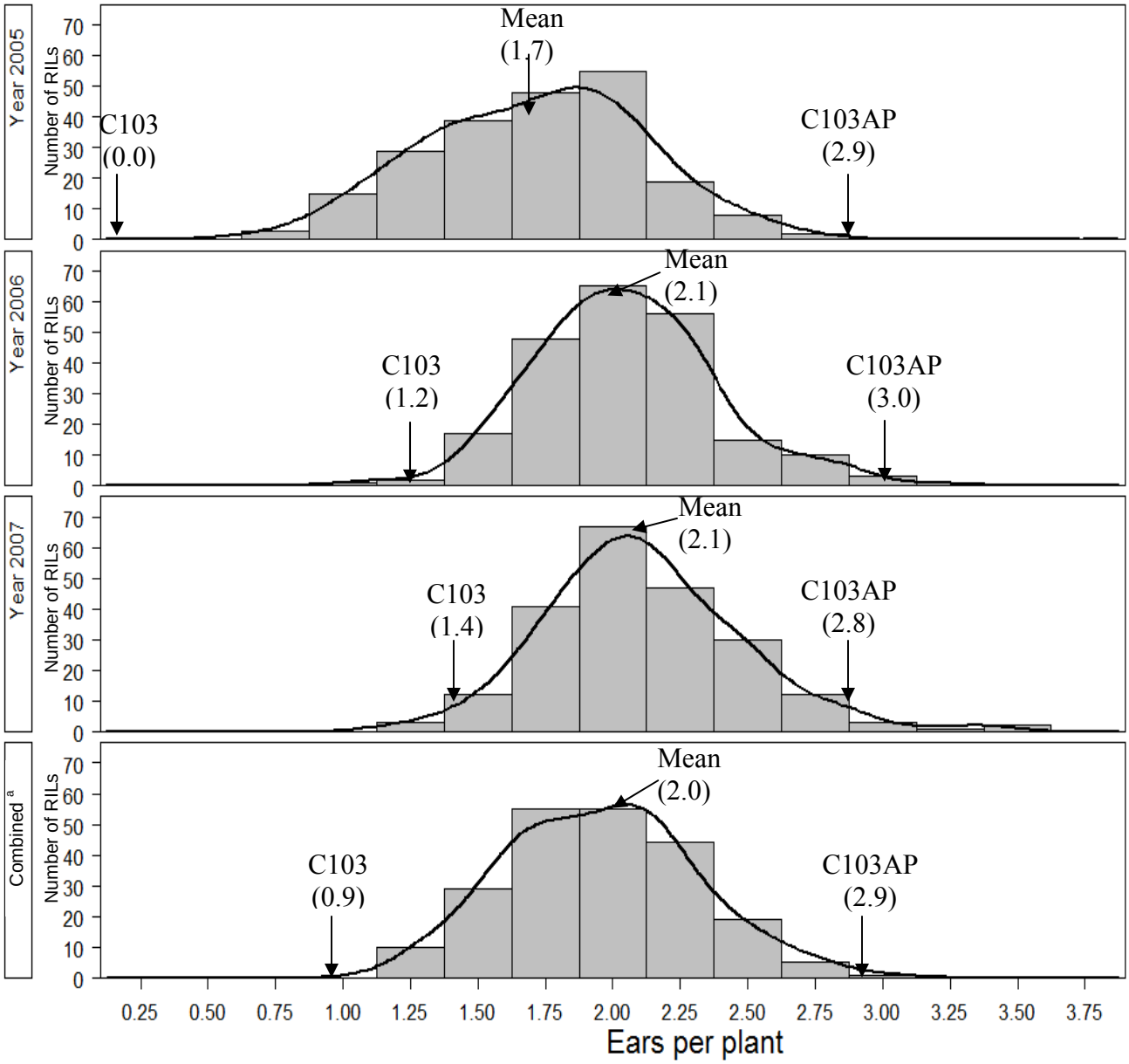
Trait	Environment	Location[a]	Epistatic interactions[b]	Additive effect[c]	Partial $R^{2,d}$
EPP	Combined[f]	2.08 × 3.08	umc1798 × phi046	0.06 (1.89)	2
		2.08 × 8.04	umc1798 × umc1858	0.07 (2.17)	0.6
			Total $R^{2,e}$		2.6
	2005	2.04 × 6.06	umc1448 × umc1859	0.06 (1.61)	1.8
	2006	3.05 × 6.06	umc1167× umc1912	-0.04 (1.17)	1.3
		6.01 × 6.06	bnlg107× umc1912	-0.05 (1.97)	1.3
		3.10 × 9.04	bnlg1098× umc1492	-0.04 (1.21)	1.7
				Total $R^{2,e}$=	4.3
	2007	3.10 × 8.06	bnlg1098 × umc1149	-0.05 (0.87)	0.7
GDU	Combined[f]	1.02 × 8.05-8.06	umc1568 × umc1665	4.35 (2.4)	1.4
	2005	6.06 × 8.05	umc1912× bnlg1176	4.04 (1.39)	3.2
	2006	2.01 × 6.06	phi96100× umc1762	-2.85 (1.08)	0.6
		1.02 × 8.05	umc1568× umc1309	2.65 (0.73)	0.6
		1.02 × 9.04	umc1568× umc1771	2.68 (0.91)	0.1
		6.06 × 9.04	umc1912× umc1771	-3.15 (1.21)	0.2
				Total $R^{2,e}$=	1.5
	2007	1.02 × 8.05	umc1568× bnlg666	4.66 (1.80)	0.6
PH	Combined[f]	1.03 × 4.06	phi001 × umc1511	1.18 (0.71)	0.2
		3.10 × 4.06	bnlg1098 × umc1511	-1.61 (1.12)	1.3
		2.06 × 8.05	bmc1329 × umc1309	-1.96 (0.93)	1.5

[a] Chromosomal location for each locus involved in the epistatic interaction according to maizeGDB
[b] Nearest left loci from the QTL involved in the interaction
[c] Additive effect values and their LOD value in parenthesis
[d] Proportion of the phenotypic variance explained by each epistatic interaction term defined in the multiple regression model
[e] Proportion of the phenotypic variance explained by the sum of the epistatic terms defined in the multiple regression model
[f] Refers to the combined analysis of the 2005, 2006 and 2007 datasets

Table 5 Cont.

Trait	Environment	Location[a]	Epistatic interactions[b]	Additive effect[c]	Partial $R^{2,d}$
		1.02 × 9.03	umc1976 × mmc0051	1.77 (1.42)	3.4
		2.06 × 9.03	bmc1329 × mmc0051	2.46 (1.99)	1.3
				Total $R^{2,e}$=	7.7
	2005	1.03 × 3.06	phi001 × umc1644	2.10 (1.99)	4
		2.06 × 3.06	bmc1329 × umc1644	1.60 (0.86)	1
		3.10 × 4.06	bnlg1098 × umc1511	-2.12 (1.74)	3.2
		2.06 × 9.03	bmc1329 × umc1688	1.88 (1.10)	0.1
		8.05 × 9.03	umc1309 × umc1688	1.79 (1.12)	0.3
				Total $R^{2,e}$=	9.6
	2006	1.02 × 2.06	umc1976× bmc1329	-3.62 (3.31)	7.2
		2.06 × 3.06	bmc1329× umc1644	1.88 (1.07)	1.3
				Total $R^{2,e}$=	8.5
	2007	1.03 × 4.06	phi001× umc1511	2.15 (1.83)	0.2

Figure 1: Histograms showing the phenotypic distribution of the number of ears per plant (EPP) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population



[a] refers to the combined analysis of the 2005, 2006 and 2007 datasets

Figure 2: The genetic map from a maize population of 218 RILs derived from inbred lines C103 and C103AP using 123 SSR loci. The 1-LOD support interval of QTL is shown. Dotted lines represent genomic regions not polymorphic between C103 and C103AP and connect linkage groups of the same chromosome. SSR locus position (in cM) is shown on the left of each chromosome and its name is shown on the right

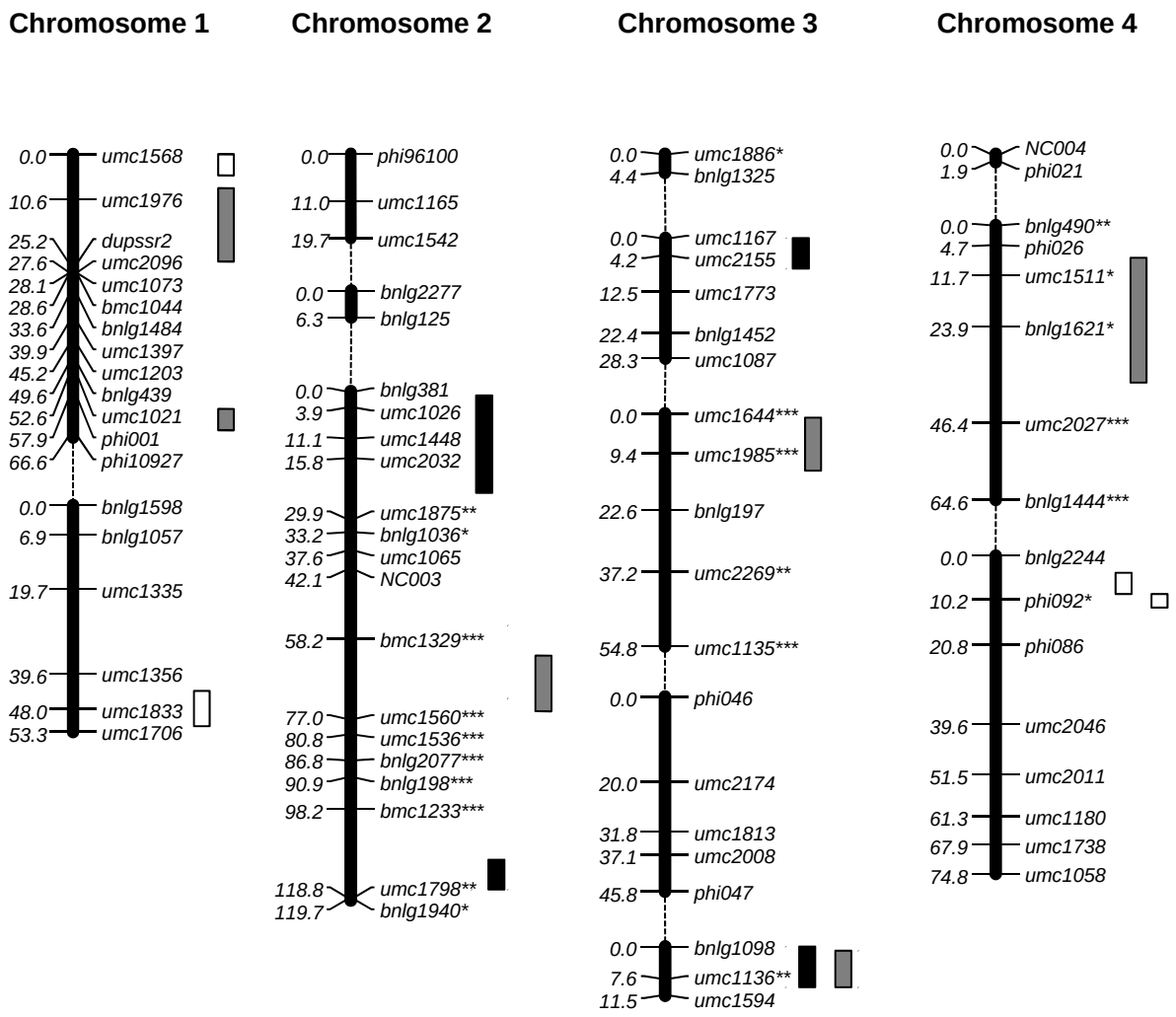
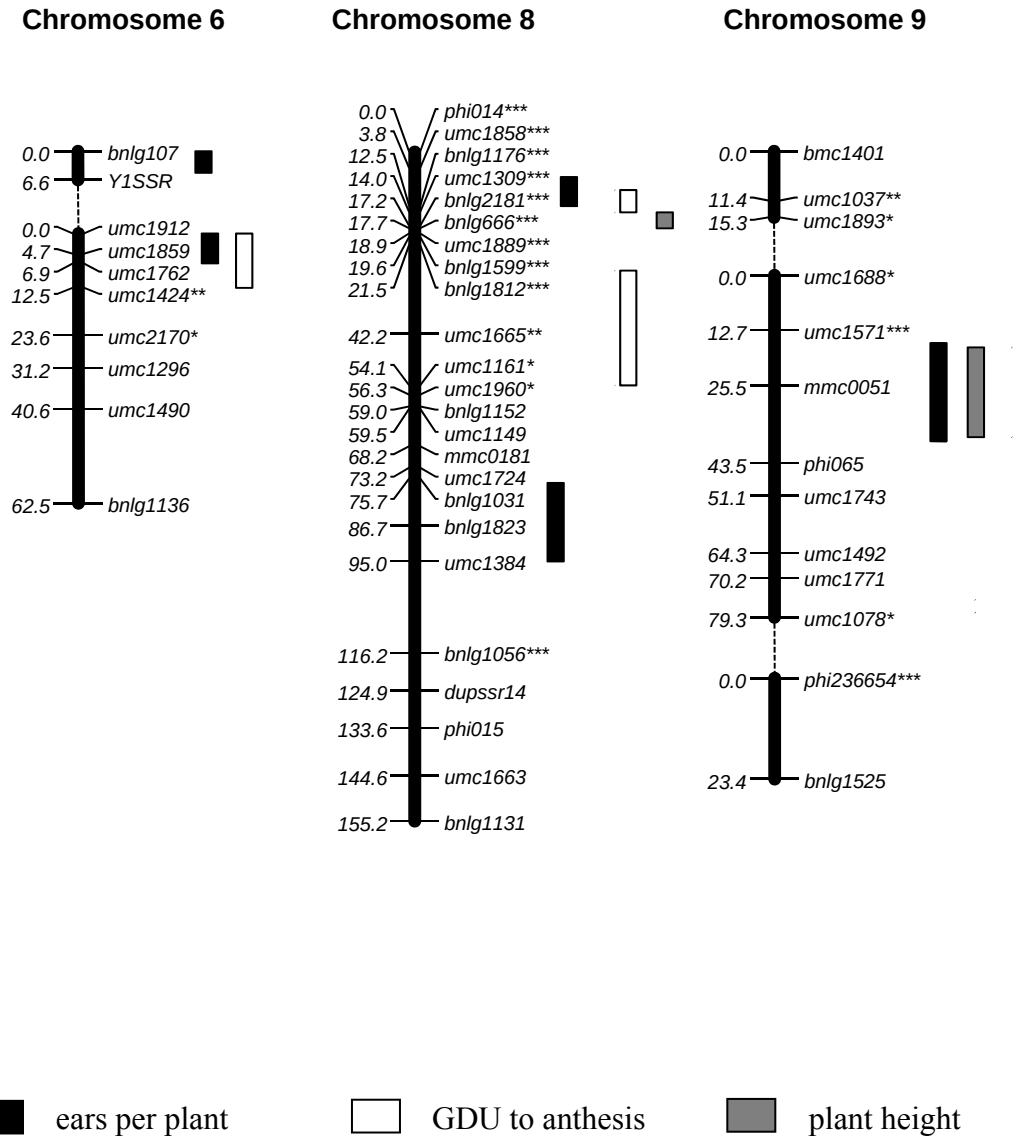


Figure 2 Cont.



*Departure from a genotypic ratio of 1:1 significant at the 0.05 probability level
** Departure from a genotypic ratio of 1:1 significant at the 0.01 probability level
*** Departure from a genotypic ratio of 1:1 significant at the 0.001 probability level

APPENDIX

Table A1: Mean phenotypic values of inbreds C103, C103AP and 218 RILs and variance components for ears per plant (EPP), GDU to anthesis and plant height in years 2005, 2006 and 2007

Trait	Environment	Mean Phenotypic Values[b]					
		C103	C103AP	RILs	σ_G^2	σ^2	H^2
EPP (number)	2005	0.00 ± 0.59	2.92 ± 0.49***	1.71 ± 0.04**	0.21 ± 0.03***	0.12 ± 0.01	0.78 ± 0.03
	2006	1.17 ± 0.48	3.04 ± 0.41**	2.05 ± 0.03*, [c]	0.15 ± 0.02***	0.07 ± 0.007	0.81 ± 0.03
	2007	1.36 ± 0.53	2.83 ± 0.49	2.11 ± 0.04	0.19 ± 0.03***	0.16 ± 0.02	0.71 ± 0.04
GDU[a] to anthesis	2005	817.31 ± 40.31	852.69 ± 36.55	779.75 ± 2.54*, [d]	1258.59 ± 137.04***	305.24 ± 31.26	0.89 ± 0.01
	2006	803.87 ± 34.02	874.11 ± 32.68	795.28 ± 2.24*, [d]	1041.19 ± 105.93***	115.96 ± 12.01	0.95 ± 0.007
	2007	785.62 ± 36.36	876.77 ± 35.82	805.94 ± 2.45*, [d]	1244.20 ± 126.75***	136.80 ± 14.19	0.95 ± 0.007
Plant height (cm)	2005	162.29 ± 14.98	178.11 ± 13.63	162.03 ± 0.95	174.46 ± 19.24***	44.83 ± 4.60	0.89 ± 0.02
	2006	171.67 ± 14.79	186.42 ± 14.24	173.93 ± 0.97	198.72 ± 19.98***	17.85 ± 1.83	0.96 ± 0.006
	2007	177.31 ± 15.96	179.06 ± 15.78	168.34 ± 1.08	241.41 ± 24.26***	20.19 ± 2.11	0.96 ± 0.006

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
***Significant at the 0.001 probability level.
[a] Growing degree units according to Cross and Zuber (1972).
[b] Tukey multiple-comparison test.
[c] Difference between C103 and the mean of the population.
[d] Difference between C103AP and the mean of the population.

Table A2: Phenotypic (above diagonal) and genetic correlations with their standard errors (in parenthesis) in years 2005, 2006 and 2007 for ears per plant (EPP), GDU to anthesis and plant height

	Environment	EPP	GDU to anthesis	Plant height
EPP	2005	-	0.17 (0.06)	0.16 (0.06)
	2006	-	0.30 (0.06)	0.19 (0.06)
	2007	-	0.08 (0.06)	0.27 (0.05)
GDU to anthesis	2005	0.26 (0.08)	-	0.03 (0.06)
	2006	0.37 (0.07)	-	0.39 (0.05)
	2007	0.14 (0.08)	-	0.36 (0.06)
Plant height	2005	0.21 (0.08)	0.08 (0.08)	-
	2006	0.22 (0.07)	0.45 (0.06)	-
	2007	0.35 (0.07)	0.42 (0.06)	-

[a] Phenotypic and genetic correlation coefficients were calculated using REML from the MIXED procedure in SAS. Approximate sampling variances and standard errors for the estimates were obtained after Holland (2006).

Table A3: Regression coefficients from single-marker analysis at loci of chromosomes 2, 5 and 10 from a maize population of 218 RILs derived from inbreds C103 and C103AP.

	Chromosome location[a]						
	2.08	5.00	5.01	5.07		10.06	
	---------- SSR Loci ----------						
Environment	mmc0381	umc1325	umc1766	bnlg1346	phi085	bnlg1028	umc1477
	---------- Ears per plant ----------						
2005	0.004	0.04	0.07**	-0.004	-0.03	0.05	0.02
2006	-0.02	0.03	0.05*	0.02	-0.03	0.05*	0.03
2007	-0.03	0.06*	0.07**	-0.01	-0.03	0.03	-0.01
Combined[b]	-0.02	0.04	0.07**	0.005	-0.03	0.05*	0.02
	----------GDU to anthesis----------						
2005	-2.27	1.46	6.40**	-1.32	-1.29	2.61	-1.76
2006	-2.18	1.63	5.07*	-0.92	-0.03	2.42	-0.21
2007	-2.06	1.62	3.80	-0.02	-1.86	2.46	-0.91
Combined[b]	-2.31	1.70	5.07*	-0.68	-1.03	2.58	-0.79
	--------- Plant Height----------						
2005	0.62	1.02	0.66	1.28	-1.13	-2.39**	1.12
2006	-0.02	0.87	0.93	0.52	-0.59	-1.67	0.30
2007	0.01	1.02	0.56	0.32	-0.63	-1.66	0.07
Combined[b]	0.212	0.99	0.67	0.75	-0.80	-1.94*	0.55

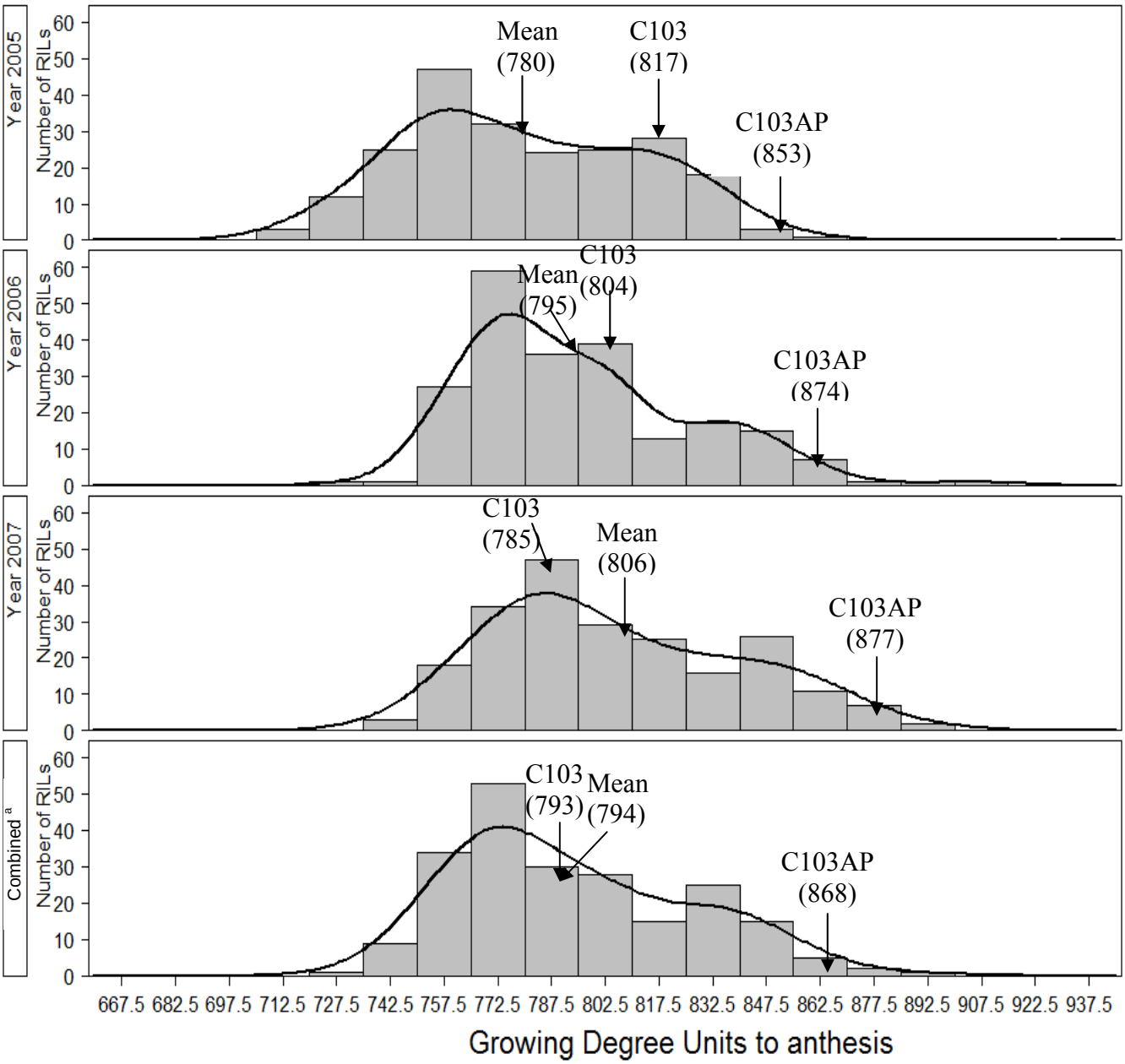
* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
[a] Chromosomal bin corresponding to the location of the SSR locus according to MaizeGDB database.
[b] Refers to the combined analysis of the 2005, 2006 and 2007 datasets.

Table A4: Summary of QTL detected by Multiple-trait Composite Interval Mapping in the three years of evaluation (2005, 2006 and 2007) for ears per plant (EPP), GDU to anthesis and plant height from a maize population of 218 RILs derived from the inbreds C103 and C103AP

Trait	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive Effect	Parental Allele[d]
EPP						- number -	
	1	6.06	umc1912	0.00	5.12	0.09	C103
	2	8.05	umc1309	14.00	4.69	0.09	AP
GDU to anthesis						-- gdu[e]--	
	1	6.06	umc1912	2.00	6.68	7.82	C103
	2	6.06	umc1762	8.90	6.77	7.84	C103
	3	8.04	umc1858	11.80	36.93	26.18	AP
	4	8.05	bnlg1812	25.50	24.61	25.00	AP
Plant Height						-- cm--	
	1	1.03	phi001	61.9	5.68	3.70	C103
	2	3.06	umc1644	6	5.43	3.83	AP
	3	3.06	umc1985	13.4	4.91	3.39	AP
	4	3.10	bnlg1098	6	5.92	3.41	C103
	5	4.06	umc1511	13.7	4.59	1.78	C103
	6	8.05	umc1309	14	14.06	3.37	AP
	7	8.05	bnlg1812	23.5	8.82	2.51	AP
	8	8.05	umc1665	46.2	4.61	0.83	AP

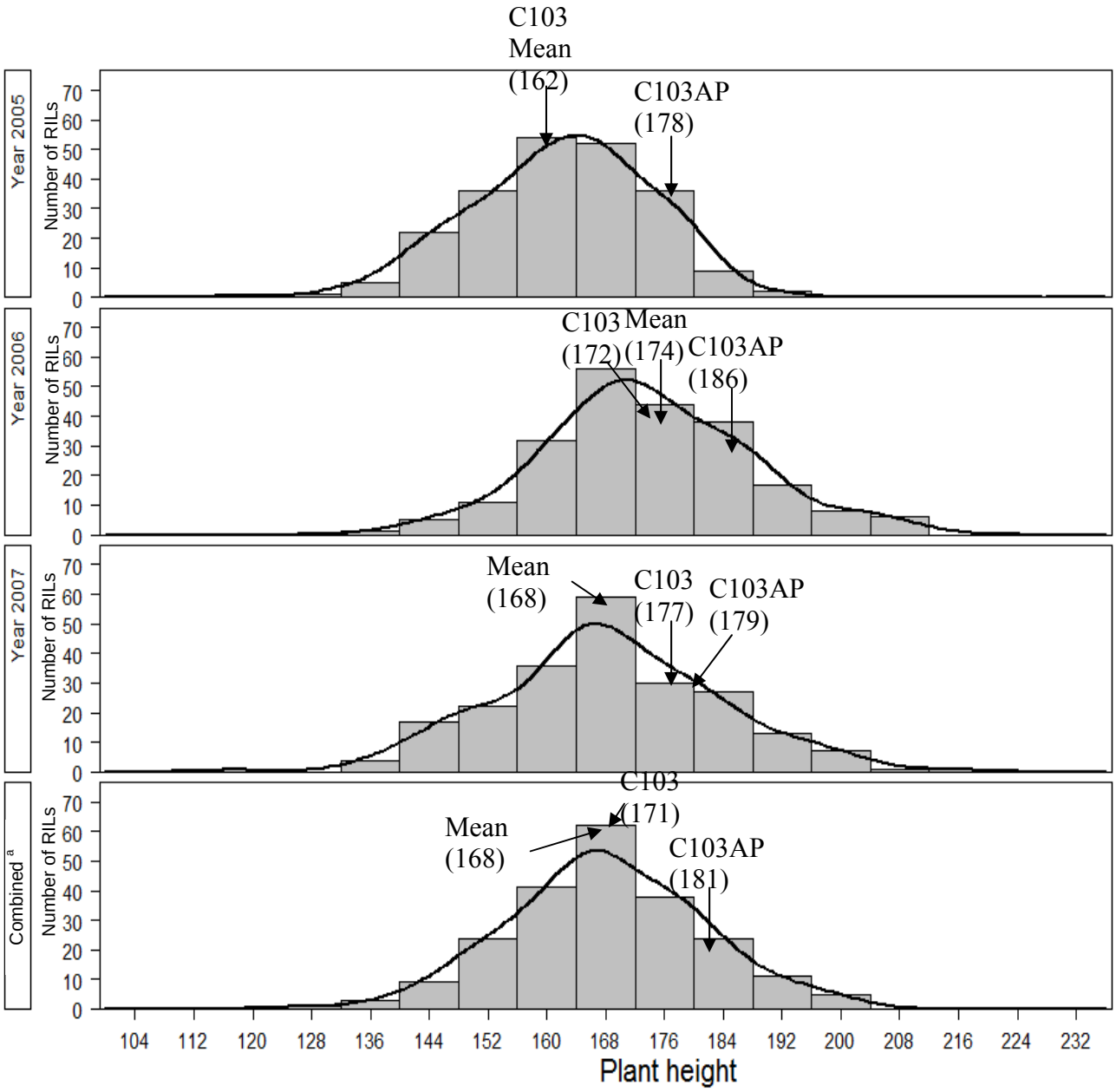
[a] Chromosomal bin corresponding to the location of the nearest left marker according to MaizeGDB database.
[b] Position in cM from the top of the chromosome calculated by MapMaker v.3
[c] LOD score corresponding to the position of the QTL as calculated by WinQTL cartographer
[d] Direction of the additive effect, AP= C103AP
[e] Growing degree units according to Cross and Zuber (1972)

Figure A1: Histograms showing the distribution of GDU to anthesis from a population of 218 RILs derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population



[a] refers to the combined analysis of the 2005, 2006 and 2007 datasets.

Figure A2: Histograms showing the distribution of plant height from a population of 218 RILs derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population



[a] refers to the combined analysis of the 2005, 2006 and 2007 datasets.

CHAPTER 3: GENETIC ANALYSIS OF TASSEL ARCHITECTURE IN A NEAR-ISOGENIC POPULATION OF MAIZE RECOMBINANT INBRED LINES

A paper to be submitted to Theoretical and Applied Genetics

M. L. Mayor, M. Lee and A. A. Mahama

ABSTRACT

In U.S. maize breeding programs tassel size has decreased along with an increase in grain yield. Other morphological and physiological traits that promote efficiency in growth and partitioning of dry matter also changed. The first objective of this study was to detect genomic regions associated with tassel morphology in a population of recombinant inbred lines (RILs) of the biparental cross of inbreds C103 and C103AP, near-isogenic lines. The second objective was to search for interactions between these regions as well as with the environment. The C103 tassel has fewer primary branches and a longer central spike and branching zones than C103AP. The parental inbreds also differ for prolificacy; thus providing an opportunity to simultaneously map genetic factors for tassel and ear development. Evaluation of both inbreds at 890 Simple Sequence Repeat (SSR) loci determined that C103 and C103AP were polymorphic at 174 loci. The genetic map was constructed with 123 SSR loci spanning 894 cM. Both parents and the 218 RILs were evaluated during two consecutive years in replicated experiments. A total of 32 QTL were detected for tassel branch number (TBN), tassel length (TL), central spike (CSL) and

branching zone length (BZL). The majority of these QTL were located on chromosomes 1, 2, 3, 4 and 8. The QTL for TBN, TL and CSL with the strongest association to the phenotypic variance were located in bins 2.01, 2.06, 2.08 and 9.03. These regions were previously not associated with tassel architecture on basis of candidate genes and QTL mapping. Nine of the QTL with the strongest association to variation in tassel morphology were located in the same bins as candidate genes involved in the inflorescence pathway.

INTRODUCTION

In the U.S maize grain yields have increased several folds during the previous one hundred years. Morphological and physiological traits that promote efficiency in growth, development and partitioning such as tassel size have also changed during this period (Duvick 2005). Tassels can affect grain yield by reducing light interception into the canopy as well as by utilizing carbohydrate resources (Duncan et al. 1967). In addition to its relevance for seed production, variation in tassel morphology could be an indicator of genetic diversity in both wild and cultivated maize (Orr and Sundberg 1994).

The maize inflorescences develop from the shoot apical meristem (SAM). The SAM gives rise to the terminal tassel, which has long branches and develops the male flowers. The ears are derived from axillary shoot meristems, have a prominent axis with no long branches, and develop the female flowers. The inflorescence meristems (IMs) of the tassel and the ear each produce spikelet pair meristems (SPMs). Each SPM forms a short branch, bearing two Spikelet Meristems (SMs), which in turn produce a pair of floral meristems (FMs). In the ear, only one of the FMs develops into a fertile flower (Bommert et al. 2005b). Genes related to the initiation, maintenance and development of inflorescence meristem have been identified

and assembled in a pathway of maize floral development (Malcomber et al. 2006). The gene loci *Suppressor of sessile spikelet 1 (Sos1)*, *Barren inflorescence1 (Bif1)* and *2* (*bif2*); and *barren sterile* (*bs)* control inflorescence branches by affecting steps in the development of the inflorescence meristem (Evans and Barton 1997). Other genes affecting inflorescence development are *fasciated ear2* (*fea2*), ramosas (*ra1* and *ra2*), *thick tassel dwarf* (*td1*) and *branched silkless* (*bd1*) (Bommert et al. 2005a; Bommert et al. 2005b; Bortiri et al. 2006; Vollbrecht et al. 2005).

Quantitative Trait Loci (QTL) studies for tassel morphology (i.e. tassel branch number, length of the branching zone, central spike and tassel length) reported that QTL for the individual variation of these traits were most frequently located on chromosomes 2, 4, 5 and 7 (Table 1). A common aspect of four of these QTL studies was that inbred B73 was the parent in the genetic populations used to assess tassel morphology, suggesting a narrow genetic background (Mickelson et al. 2002; www.maizegdb.org/MIP; verified 12/6/07; Table 1). Only one of these studies simultaneously considered variation in ear morphology (i.e. kernel row number, number of kernels per row and kernel number density) along with variation in tassel morphology (Upadyayula et al. 2006a). However, it is plausible that genetic factors influencing the expression of tassel architecture might affect ear development, since they have a common developmental origin (Bommert et al. 2005b).

In this study, four traits defining tassel architecture were investigated in a population of 218 recombinant inbred lines (RILs) derived from a cross between two near-isogenic maize inbred lines (NILs). The description of QTL for ear development (i.e. prolificacy) in this population is the focus of another manuscript. Populations derived from NILs have been used to simultaneously map, verify and incorporate QTL into adapted, elite genetic

backgrounds as well as to verify and fine-map QTL previously detected in other mapping populations (Graham et al. 1997; Salvi et al. 2002; Sanchez et al. 2000; Szalma et al. 2007). Given the close genetic relationship between the parents of this population, fewer and smaller regions in the genome carrying the QTL of interest would be segregating. In this context, the proportion of the phenotypic variability explained by a QTL should be increased along with an increase in the detection of minor QTL by elimination of QTL for other, unrelated traits (Eshed and Zamir 1995; Zamir and Eshed 1998).

The objectives of this study were: to detect genomic regions associated with four tassel morphological traits in the C103/C103AP population, and to define interactions between these regions as well as with the environment.

MATERIALS AND METHODS

Plant material

C103 is a single-eared inbred that tends to abort ear development and produce barren stalks. This inbred belongs to the Lancaster Sure Crop Group and was developed at the Connecticut Experimental Station (Jones and Everett 1949). Inbred C103 was mated to an accession of a prolific Argentinean popcorn population (AP) to improve prolificacy. Inbred line C103AP was obtained by four backcrosses with continuous selection for resemblance to C103 and for development of secondary ear-shoots under populations of 40000 plants ha^{-1} (Duvick 1974). The converted inbred (C103AP) was compared to the original (C103) in hybrid combinations at three plant populations: 24700, 49400 and 74100 plants ha^{-1}. At all densities, testcross hybrids from both inbreds differed in the number of ears per plant with C103AP hybrids being more prolific than C103 hybrids (Duvick 1974). Our observations

established that the inbreds per se also differed for number of ears per plant, plant height, growing degree units to anthesis and tassel morphology.

Population development and phenotypic evaluation

A population of 218 recombinant inbred lines (RILs) from the cross between C103 and C103AP was created by self-pollinating unselected F_2 plants. Each RIL was derived from a single F_2 plant by self-pollinating one plant per line until the $F_{7:8}$ generation. For the field experiments, seed for each RIL was increased by selfing three plants in the $F_{7:8}$ generation and harvesting them in bulk.

During 2006 and 2007, the 218 RILs, two entries for C103AP and one entry for C103 were evaluated in a 13×17 lattice with 2 replications at the Agronomy and Agricultural Engineering Research Center (AAERC) near Boone, Iowa (42°1′ N, 93°45′ W). Nitrogen (urea) fertilizer was applied pre-plant at 175 kg N ha^{-1}. Weed control was accomplished with 1.75 L ha^{-1} of Dual II Magnum (S-metolachlor/benoxacor) and 2.22 L ha^{-1} of atrazine [2-chloro-4-(ethylamine)-6-(isopropylamine)-s-triazine] and manual cultivation. Annual and historic weather data were collected from an automated weather station located less than one kilometer from the field plots (Table 1).

Experiments were hand-planted on 10 May 2006 and 6 May 2007 to achieve a final plant density of 24700 plants ha^{-1}. Each plot consisted of a row of 5.5 m long. The seeding rate was two kernels per hill. The distance between hills in a row was 0.56 m with 0.75 m between rows. At growth stage V5 (Ritchie et al. 1996), each hill was thinned to one plant.

After flowering, tassel architecture was defined by four features of tassel morphology assessed on the tassel of each plant in a plot. The traits were defined based on studies from Upadyayula et al. (2006a) and shown in Figure 1. The tassel branch number (TBN) was

determined by counting the number of primary lateral tassel branches from each tassel. In a plot, TBN was measured per plant and averaged by the number of plants in a plot. The total length of the tassels was measured from the non branching node present below the first primary branch to the tip of central spike of the tassel. The tassel length (TL) was obtained by averaging individual tassel measurements by the number of plants in a plot. The length of the central spike (CSL) was measured from the top primary branch to the tip of the central spike. The central spike length was calculated by averaging these measurements by the number of plants in a plot. The length of the branching zone (BZL) of the tassel was defined as the distance from the top branch nearest to the tip of the central spike to the non branching node present below the lowermost primary branch. The BZL was calculated as the difference between CSL and TL.

Statistical analyses of phenotypic data

For each trait, a combined analysis was completed for the 2006 and 2007 datasets. Homogeneity of variance across years was evaluated using the Levene's test on absolute residuals (Levene 1960). Best linear unbiased predictors (BLUPs) estimated the phenotypic value of each RIL (Bernardo, 2002). A mixed-effects linear model with the REML method was implemented to estimate the phenotypic values described above following to the model: $y_1 = \mu_1 + \tau_{1i} + \alpha(\tau)_{1ij} + \beta(\alpha\tau)_{1ijk} + \gamma_{1l} + (\gamma\tau)_{1il} + \varepsilon_{1ijkl}$; where $\mu_1 =$ overall mean for trait *1*; $\tau_{1i} =$ effect of the *i* environment (*i*=1, …, 3); $\alpha(\tau)_{1ij}$ =effect of the *j* replications (*j*= 1, 2); $\beta(\alpha\tau)_{1ijk} =$ effect of the *k* blocks within each replication (*k*=1, 2… 17); $\gamma_{1l} =$ effect of the *l* RILs (*l*=1, 2…, 218); $(\gamma\tau)_{1il} =$ term to account for the genotype by environment interaction and $\varepsilon_{ijkl} =$ error term. Environments, replications, and block effects were considered as fixed

whereas the effects of the RILs were specified as random. Best linear unbiased estimators (BLUE) of C103, C103AP and the mean of the RILs phenotypic values were obtained by Restricted Maximum Likelihood (REML). Degrees of freedom were estimated using the Kenward–Rogers version of the Satterthwaite approximation. The Tukey multiple-comparison test was used to assess differences in TBN, TL, CSL and BZL between both inbreds and between each inbred and the mean of the population (SAS 2007; version 9.1.3).

The estimates of variance components and their standard errors were obtained by the REML method and used to calculate the broad-sense heritabilities on a progeny-mean basis ($H^2 = \frac{\hat{\sigma}_{\lambda}^2}{\left(\frac{\hat{\sigma}_{\tau}^2}{e}\right) + \hat{\sigma}_{\lambda}^2 + \left(\frac{\hat{\sigma}_{\varepsilon}^2}{re}\right)}$ where: $\hat{\sigma}_{\lambda}^2$ = genetic variance of the trait; $\hat{\sigma}_{\tau}^2$ = variance due to environmental factors; $\hat{\sigma}_{\varepsilon}^2$ = error variance; *e*= number of environments and *r* = number of replications) and their standard errors (Holland et al. 2003). The estimates of the phenotypic and genetic correlations were obtained via REML method from the MIXED procedure in SAS. Approximate sampling variances and standard errors for the phenotypic and genotypic correlation estimates were obtained with the delta method (Holland 2006; Lynch and Walsh 1998).

Genotypic evaluation

Before anthesis, leaf tissue from 218 RILs, C103 and C103AP was collected and lyophilized. Each inbred was represented by bulked leaf tissue of 10 plants. Genomic DNA was extracted via the modified-CTAB protocol (Saghai-Maroof et al. 1984). The original donor parent, Argentinean popcorn (AP), was not available for evaluation. C103 and C103AP were evaluated at 890 Simple Sequence Repeat (SSR) loci distributed throughout

the maize genome. These loci were selected from the maize database based on polymorphism among a standard set of inbred lines (www.maizegdb.org; verified 12/07/07). A high frequency of monomorphic loci between C103 and C103AP was expected since the latter was developed by four backcrosses and phenotypic selection for resemblance to C103. Polymorphism between C103 and C103AP was detected at 174 SSR loci. A subset of 123 SSR loci, with clear and repeatable PCR products, and maximum coverage of the genome, was used to construct the linkage map.

The PCR reagents were: 10 X buffer (Tris-HCl 100 mM, pH 8.3, 500 mM KCl, 1mg mL^{-1} gelatin); 2.5 mM MgCl2; 0.1 mM dNTPs; 0.17 pmol μL^{-1} forward- and reverse-primers, 0.2 mg mL^{-1} BSA, 1X betaine, 0.5 units μL^{-1} *Taq* polymerase and 50 ng genomic DNA. A final reaction volume of 15 µL was achieved by adding ddH_2O. The touchdown PCR protocol was used and the cycling was as follows: activation of the polymerase for 1 min at 95 °C followed by twenty cycles of 94 °C for 1 min, annealing at 65 °C for 1 min, and extension at 72 °C for 2 min (the annealing temperature was decreased by 1 °C every 2 cycles down to 56 °C), then 21 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and extension at 72 °C for 2 min. PCR products were stored at 4 °C until electrophoresis in 4% MetaPhor® agarose gel. Gels were stained with ethidium bromide, photographed and scored visually.

Genotypic analysis and genetic map construction

The population structures most commonly used for QTL mapping possess sufficient polymorphism to generate linkage maps suitable to scan the whole genome. In some cases, the polymorphism available between the parents is not sufficient to generate linkage maps that evaluate the entire genome. For example, a genetic map with more than one linkage

group per chromosome has been used for the study of prolificacy in the Golden Glow maize population (de Leon et al. 2005). In the present study, several genomic regions were monomorphic between C103 and C103AP, leading to a genetic map with more than one linkage group per chromosome (Figure 2). These monomorphic loci could arise from regions that are identical by descent (IBD) between C103 and C103AP or regions that are alike in state between these inbreds. Since the AP parent was not available for evaluation, the monomorphic regions could not be resolved to be IBD or alike in state.

At each SSR locus, a chi-square goodness-of-fit test was performed to detect deviations from the expected 1:1 genotypic segregation ratio. The genetic map was constructed using MAPMAKER Version 3.0 (Lander et al. 1987). The SSR loci were placed in linkage groups based on a LOD score of 3 and a maximum Haldane distance of 50 cM. For each linkage group, the 'three point' command was used to pre-compute the likelihoods of all three-point combinations. The LOD threshold to exclude a three-point order from a linkage group was set to 4. The 'order' command was used to find the best linear order of loci in a given linkage group based on the informativeness criteria (2 cM minimum distance between any pair of loci, the 218 individuals being informative for each locus and loci being codominant). Since the 'order' command adds loci sequentially, once the linkage map was obtained, it was tested using the 'ripple' command and the best loci order(s) was defined based on a LOD threshold of 2. Unplaced markers were located using the 'build' command. The 'triple error detection on' command was used to detect mistyping or misclassification in the genotyping.

QTL detection

Genomic regions associated with TBN, TL, CSL and BZL, were first defined using Composite Interval Mapping (CIM) from WinQTL Cartographer (Wang et al. 2007; Zeng 1994). The empirical LOD threshold for CIM was determined by 1000 permutations for an overall significance level of 5% (Churchill and Doerge 1994). The empirical LOD thresholds for the combined analysis were as follows: 2.7 for TBN and TL; and 2.7 for CSL and BZL. The forward and backward stepwise selection criterion was used in the regression model to select cofactors accounting for genetic background variation when mapping QTL by interval test ($p_{entry/remove} \leq 0.1$). The window size on either side of the loci flanking the interval was defined as 10 cM.

Results obtained from CIM were used to explore the genetic regression models and search for epistatic interactions with Multiple Interval Mapping (MIM). The underlying theoretical bases to define the best genetic model in MIM were described by Kao and Zeng (1997), Kao et al. (1999) and Zeng et al. (1999). The Hannan and Quinn (1979) Bayesian Information Criteria (BIC) with the penalty function $c(n)= 2\ln(\ln(n))$, with n=218 was used to define the critical value for testing QTL effects and epistatic interactions to be included in the multiple regression genetic model.

The QTL detected by CIM were included in the initial MIM model if the LOD threshold was >2.4 and the minimum distance between QTL was 5 cM defining an initial genetic model for MIM. This initial model was tested for significance in MIM based on the critical value (Hannan and Quinn 1979) and the significant effects were retained. A new search for QTL was performed and the genetic map was scanned to determine the best position of the QTL. At that position, a likelihood ratio test was performed to retain or delete

this QTL from the model based on the critical value. A search for epistatic effects was conducted among those pair-wise interaction terms not yet included in the model. The process was repeated until no more significant epistatic effects were found. Each QTL in the MIM genetic model was re-evaluated for its significance. If the LOD for a QTL (marginal or epistatic) effect was less than the significance threshold conditioned on the other effects, the effect was removed from the model. However, if the marginal effect of a QTL, with a significant epistatic effect with other QTL, fell below the threshold, this marginal effect was retained. Finally, the position of each QTL was optimized in a sequential order by evaluating the model likelihood for possible positions conditional on the positions of other QTL in the model. The position that maximized the model likelihood is chosen as the new position of the QTL (Tao et al. 2003).

Once the final genetic model was defined, the genetic variance was estimated by the sum of squares of the final QTL model, and the phenotypic variance was estimated by the total sum of squares. The estimate of the broad-sense heritability was approximated by the coefficient of determination (R^2) of the MIM model (Kao et al. 1999). The proportion of total genetic variance explained by the final QTL model was obtained by dividing R^2 by the progeny mean broad-sense heritability of the trait (i.e. R^2/H^2) (Dudley 1993).

Several regions in the genetic map were monomorphic such as chromosomes 5, 7 and 10. Chromosome 5 was evaluated at 90 SSR loci, but only 4 of them being polymorphic (bins 5.00, 5.01 and 5.07). On chromosome 10, only two loci, in bin 10.06, out of 54 SSR loci screened were polymorphic. Since these loci could not be assembled into linkage groups, a reasonable alternative for simple detection of QTL linked to a locus was single-marker analysis. The distribution of the phenotypic values for each trait was examined separately for

each locus and likelihood ratio tests were performed to detect association between the phenotypic variation and genotypic classes at each such locus (Wang et al. 2007).

Multiple Trait Composite Interval Mapping (Mt-CIM) was used to test QTL × environment interaction for each trait (Jiang and Zeng 1995). The empirical LOD threshold for Mt-CIM was determined by 1000 permutations for an overall significance level of 5% (Churchill and Doerge 1994). Hence, the LOD thresholds were: 3.6 for TBN, 3.6 for TL, 3.7 for CSL and BZL. The forward and backward stepwise selection criterion was used to select cofactors accounting for genetic background variation when mapping QTL by the interval test ($p_{entry/remove} \leq 0.1$). The window size on either side of the flanking loci was defined as 10 cM. By testing the QTL × E interaction, the null hypothesis of $b_1 = b_2 = \ldots = b_j$ at the chromosome regions where QTL have been identified by the joint mapping is being tested; where b_j corresponds to the additive effect of a QTL in the jth environment (Jiang and Zeng 1995).

Assessment of pleiotropy and linkage for QTL

Pleiotropy, defined as a single gene (or DNA sequence) influencing multiple traits (Wright 1968), was assessed by first comparison of confidence intervals of the QTL and then by testing the position of these QTL. For each QTL, a one-LOD support interval was constructed as described by Lander and Botstein (1989). QTL with non-overlapping one-LOD support intervals were considered different QTL (Conneally et al. 1985, Lander and Bolstein 1989). Multiple-trait MIM (Mt-MIM) was used to perform a likelihood ratio test for the hypothesis of pleiotropy versus close linkage between two QTL with overlapping supporting intervals (Jiang and Zeng 1995).

Each QTL with an effect on one trait only, have positions symbolically specified by p_1 for the QTL having an effect on trait 1 and p_2 for the QTL having an effect on trait 2. If the two positions are in the same marker interval, p_1 and p_2 are then defined as the ratios of the recombination frequencies between a locus and the two positions, respectively, and between the two flanking SSR loci. The hypotheses were formulated as H_0: $p_1 = p_2$, H_1: $p_1 \neq p_2$. The test statistic is the likelihood ratio LR = -2 ln(L_0/L_1), where L_1 is the maximum of the likelihood for the alternative model (given by H_1) and L_0 is the maximum of the likelihood that corresponds to the null hypothesis of $p_1 = p_2$. The LR test statistic for testing position is follows an approximate chi-square distribution with 1 degree of freedom (Jiang and Zeng 1995).

RESULTS

Phenotypic analysis

C103AP had smaller tassels with reduced branching zone and central spike length when compare to C103. The mean of RILs was similar to that of C103AP for CSL and BZL, while it differed from C103 with respect to TL. C103 had fewer tassel primary branches than C103AP. The population of RILs had a lower number of branches than both parents (Table 3). Although no significant differences were observed between C103 and C103AP for TBN, TL, CSL and BZL, the population of RILs showed a wide range of variation for these traits (Table 3). Transgressive segregation was observed for all traits during both years. This could be caused either by non-additive gene action such as epistasis or by the parental lines being homozygous for sets of alleles at different loci having opposite effects (Lynch and Walsh 1998; Rieseberg et al. 2003).

Genetic and genotype by environment variance components were significant for all traits. The contribution of the G×E variance component to the total phenotypic variance was less than 10%, CSL having the highest estimate. Broad-sense heritability values ranged from 0.90 to 0.95 (Table 3). Consequently, most of the phenotypic variance had a relatively large genetic component that could facilitate the detection of QTL with large effect on the genetic variation observed for the traits (Kearsey 1998; Lande and Thompson 1990).

Genotypic analysis and genetic map construction

Considering that C103AP was developed after 4 backcrosses with selection for the C103 phenotype, theoretically C103 and C103AP were expected to be highly related (F=0.97). This degree of relatedness was reflected in the low level of polymorphic SSR loci. When both inbreds were evaluated at 890 SSR loci, 174 loci were polymorphic; thus the genetic map was constructed with information from a smaller portion of the genome. As a consequence of this highly isogenic background, more than one linkage group was defined per chromosome. A distance of 23.5 cM between two loci was used as the minimum threshold to define a linkage group to avoid the detection of false QTL positions in regions of the genetic map that were not sufficiently sampled by marker loci. The genetic map covered a total of 894 cM of the maize genome with an average distance of 7.3 cM between two loci (Figure 2). Linkage maps for chromosomes 5, 7 and 10 were not presented because very few polymorphic loci (4 loci in chromosome 5 and 2 loci in chromosome 10) were detected for those regions of the genome.

Forty-seven loci showed a significant departure from the expected 1:1 genotypic ratio. Loci with segregation departure from the 1:1 genotypic ratio on chromosomes 2 and 8 had an excess of C103 alleles. On chromosome 4, the 6 loci with significant

departure from the 1:1 genotypic ratio and a region in chromosome 9, from locus *umc1037* to locus *umc1571*, had an excess of C103AP alleles (Figure 2). Single-locus segregation distortion skews the genotypic frequencies from the expected mendelian segregation ratio and introduce a bias in the estimation of recombination fractions (Bailey 1949; Lorieux et al. 1995). Therefore, the segregation distortion observed in these regions could affect the position of the QTL detected.

QTL detection

For tassel branch number (TBN), eight QTL were found that explained 56% of the phenotypic variance and 59% of the total genetic variance (Table 5). Two QTL accounted for 24.8% of the phenotypic variance were located in bins 2.01 and 2.06. C103AP alleles increased the TBN in these locations. From single-marker analysis, QTL were detected linked to locus *bnlg1346* (bin 5.07) in 2006 and the combined analysis; locus *phi085* (bin 5.07) in 2006 and at locus *umc1477* (bin 10.06) in all environments ($p<0.05$).

With respect to tassel length (TL), eight QTL were detected, which contributed 46.4% of the phenotypic variance and 49.9% of the total genetic variance. Three QTL summarized 24.3% of the phenotypic variance and were located in bins 1.06, 2.08 and 8.04. In bin 1.06 and 2.08 C103AP alleles increased TL; while in bin 8.04, C103 alleles increased TL (Table 5). For those loci analyzed by single-marker analysis, differences between genotypic classes were detected at loci *mmc0181* (bin 2.08; $p<0.01$) and *bnlg1028* (bin 10.06; $p<0.05$) suggesting QTL linked to those loci in all environments.

For central spike length (CSL), eight QTL accounted for 47% of the phenotypic variance and 52.2% of the total genetic variance. Three QTL, in bins 2.01, 2.06 and 9.03, were responsible for 23.7% of the phenotypic variance detected in the combined analysis. In

bins 2.01 and 9.03, C103 alleles increased CSL; whereas in bin 2.08, C103AP alleles increased the value of this trait (Table 5). A QTL linked to locus *mmc0381* (bin 2.08) was detected in 2006 ($p<0.05$).

Eight QTL were detected for the BZL summarizing 52.3% of the phenotypic variance. Three QTL, in bins 1.06, 2.01 and 4.08, mostly contributed to the phenotypic variance (35.7%). Alleles from C103AP increase BZL in bin 1.06 and 2.01; whereas alleles from C103 increase the value of this trait in bin 4.08 (Table 5). Also, from differences between genotypic classes, QTL were associated to loci *mmc0381* (bin 2.08), *bnlg1346* and *phi085* (bin 5.07; $p<0.05$), and *bnlg1028* and *umc1477* (bin 10.06; $p<0.01$).

Epistatic interactions were detected for all traits, with the exception of BZL. The contribution of these interactions to the genetic variance of the QTL model ranged from 6.3%, for TL, to 7.3% for CSL (Table 5). For TBN, locus *umc1167* (3.05) interacted with locus *umc1165* (bin 2.01) in both years. Two epistatic interactions, *bnlg2277* (bin 2.02) × *bmc1233* (bin 2.08); and *bmc1233* × *umc1889* (bin 8.05), were detected in 2007 for CSL. One epistatic interaction was detected during each year of evaluation for BZL, although different loci were interacting for each year (*umc1165* was interacting with *bnlg1823* (bin 8.07) in 2006, while *umc1568* (bin 1.02) was interacting with *phi086* (bin 4.08) in 2007.

Environmental conditions recorded for the period May-August varied during both years of evaluation. The amount of precipitation was lower in 2006 compared to 2007, whereas the GDUs were similar for both seasons (Table 2). CSL had the highest contribution of the G × E variance component to the phenotypic variance (Table 3). These environmental effects along with the G ×E interaction were reflected in the detection of QTL for both years. Most of the QTL for TBN, TL and BZL were detected in all environments. For TBN five

QTL were detected in both years while only two QTL were detected either in 2006 (QTL 3) or 2007 (QTL 6). Also, a QTL in bin 9.04 was detected only for the combined analysis. For TL, QTL 1 was only detected in 2006, and QTL 4 and 8 were detected only in 2007. Three QTL associated with phenotypic variation in CSL, (QTL 2, 3 and 7) were detected in or adjacent to those bins and with the same parental allele in 2006 and 2007. For BZL, two QTL, in bin 2.08 and 3.10 were detected in the combined analysis but not in individual years, whereas, QTL 8 from the combined analysis was detected in 2006.

Assessment of pleiotropy and linkage for QTL

Phenotypic correlations between two traits can arise from environmental, developmental or genetic factors. Likewise, genetic correlations can arise either by close linkage or pleiotropy between loci affecting both traits. Phenotypic and genetic correlations between the four tassel morphological traits indicate that TBN and CSL were negatively correlated. Positive correlations were detected between TBN and BZL; and CSL and TL (Table 4). Most of the QTL for these traits were detected in nine regions on chromosomes 1, 2, 3, 4 and 8 (Figure 2).

The positive correlations found between TBN and BZL, and CSL and TL were supported by the QTL results and statistical test on the position of these QTL (Table 4 and Table 5). In bins 1.06, 4.08, and 8.06 the statistical test suggested pleiotropy between TBN and BZL ($p>0.1$). Linkage between these traits was found in bin 2.01 ($p<0.05$). Statistical evidence of pleiotropy between TL and CSL ($p=0.6$) was detected in several genomic regions: bins 1.03, 2.08, 3.06 and 9.03 ($p>0.4$).

For TBN and TL, linked QTL were detected in bins 1.06 and 4.08 ($p<0.05$). Furthermore, QTL for TBN and CSL were linked in bin 2.01 ($p<0.05$). Evidence of

pleiotropy between TL and BZL was detected in bins 1.06, 2.08 and 8.04 ($p>0.05$) based on 1-LOD support interval and on the statistical test. Whereas in bin 4.08 the statistical test detected evidence of linkage between these trait ($p=0.01$).

Although BZL and CSL were complementary traits, both components of TL, almost no linear relationship was detected between them ($r_P=-0.05$, $r_G = 0.01$). This result suggests that few loci may be common between those traits. As shown in Figure 2, and from the statistical analysis, evidence of pleiotropy between these traits was detected only in bins 2.01 and 2.08 ($p>0.3$).

DISCUSSION

The genetic analysis of tassel architecture was addressed by studying the number of primary branches (TBN) and the length of the tassel (TL), the central spike (CSL) and the branching zone (BZL) in a population of RILs derived from two near-isogenic inbreds, C103 and C103AP. Given the nature of the population, much of the maize genome was either identical by descent or alike in state; thus allelic segregation was detected at a reduced genomic region (Figure 2). Although the search for QTL was conducted at fewer genomic regions, a total of 32 QTL for all traits were detected (Table 5).

Novel genetic factors having strong association with TBN were identified. The two QTL with the strongest association to the phenotypic variance of TBN, QTL 2 and 3, were located in regions where candidate genes for maize inflorescence have not been detected (bins 2.01 and 2.06). However, QTL for TBN have been identified close to these regions (bins 2.02 and 2.05 and 2.07) (Berke and Rocheford 1999, Mickelson et al. 2002; Table 1). Similarly, QTL 2 (bin 2.01) and QTL 3 (bin 2.08) for CSL; and QTL 7 for TL (bin 8.04) are

also in locations where candidate genes and QTL for these traits have not been described (Table 1).

Previous study on QTL associated with tassel morphology considered the simultaneous analysis of several tassel traits (i.e. TBN, TL, CSL, tassel branch angle, tassel weight, central spike density and primary branch spikelet density) by principal component analysis. In those studies several QTL were identified for these linear combinations of tassel traits in some of the bins mentioned here (i.e. bin 1.03, 1.06, 2.01, 2.08, 3.06, 4.06, 4.08 and 9.03; Table 1). However, the QTL for individual traits described here and those derived from that study could not be directly compared since the principal components were constructed using other tassel traits not considered in this study.

The development of tassel and ear rely in the differentiation of the shoot apical meristem. Reports on QTL detection for prolificacy (i.e. the increase in the number of ears per plant) have been described in another manuscript. The QTL related for the variation in the number of ears per plant (EPP) were detected in chromosomes 2, 3, 6, 8 and 9. Five regions detected for EPP were also detected here for tassel traits (bins 2.08, 3.05, 3.10, 6.01, 8.04 and 9.03). These results suggest that some genetic factors in these locations might be affecting both traits in this genetic context.

In this study QTL for tassel morphology were located in the same bins (1.03, 1.06, 3.06 and 4.08) where gene loci affecting inflorescence development have been described. Numerous genetic factors involved in the inflorescence pathway have been described in bin 3.06, at which QTL for TL and CSL were found. Candidate loci affecting inflorescence development in this bin include: *barren stalk1 (ba1)*, *ERF gene in tassels/ears* (*erf4a*), *liguleless2 (lg2)* and *unbranched1 (ub1)* (www.maizegdb.org/MIP; verified 12/06/2007). The

locus *ba1* affects lateral branching in the inflorescence and is required for initiation of axillary meristems (Malcomber et al. 2006). The gene locus *liguleless2* (*lg2*) is also required for an accurate vegetative to reproductive phase transition of the shoot apex (Walsh and Freeling 1999). Similarly, two candidate genes, *barren inflorescence2 (bif2)* and *CFRME53 monopterous-like gene* (*mpA*) from the inflorescence pathway are located in the same bins as QTL for BZL, TBN and TL (bins 1.06 and 4.08). The locus *bif2* affects the formation of tassel branches, spikelet pairs in ears and tassels; and is involved in the formation of paired spikelets. Mutations in the *monopterous* (*MP*) gene affect both root and shoot apical meristem formation during embryogenesis (Evans and Barton 1997). The QTL found in this study for TBN, TL, CSL and BZL provide evidence on the putative influence of these candidate loci in the expression of tassel morphological traits. More detailed studies are needed to corroborate or reject their actual functions and effects on phenotype.

In another region, bin 1.03, where QTL for TL and CSL were located, at least two candidate genes affecting floral meristem identity and /or determinacy were identified: *Zea mays MADS14 (zmm14)* and *clavata-1 like gene (clv28)*(http://www.maizegdb.org/MIP; verified 12/6/2007). Most of these loci are directly or indirectly related to inflorescence development (Laudencia-Chingcuanco and Hake 2002; Malcomber et al. 2006; McSteen et al. 2000). Given that QTL for TL and CSL were detected in this location, the loci mentioned above could be further studied to assess their involvement in changes affecting tassel architecture.

The importance of epistasis in the genetic variance of quantitative traits among progeny of line crosses has been demonstrated (Carlborg et al. 2006; Malmberg et al. 2005; Peripato et al. 2002). The epistatic interactions detected herein did not greatly contribute to

the genotypic variance of the four tassel traits. Also, no significant additive by additive epistatic interactions have been detected for TBN in a population of RILs from inbreds B73 and Mo17 (Mickelson et al. 2002). Epistasis was either not measured or reported in other QTL studies of tassel morphology (Berke and Rocheford 1999; Upadyayula et al. 2006a; Upadyayula et al. 2006b).

Correlations among traits can arise from several sources including environmental conditions, shared developmental processes, pleiotropy, close linkage and alternative splicing between those traits. The statistical tests performed here attempted to resolve whether pleiotropy or close linkage was the genetic cause for the correlated response between several of the tassel architectural traits considered. However, to fully distinguish among these hypotheses better mapping resolution would be needed. Such resolution could be achieved by increasing the number of individuals in the population to help identify informative recombinants. The identification of recombinant progeny for the associated traits would define that they were in close linkage. The final proof that pleiotropy is the cause of the association between two traits can be obtained by identification and manipulation of the DNA sequence and derived gene products associated with the variation for those traits. In this study, the evidence of pleiotropic regions for tassel architectural traits should facilitate the introgression of AP alleles that reduce the size of tassels and yet increase the number of primary branches into elite maize germplasm.

The QTL detected in this study for the four traits defining tassel morphology were mainly located on chromosomes 1, 2, 3, 4 and 8. QTL for TBN, TL and CSL were detected in bins 2.01, 2.06 and 2.08. Similar observations were made for CSL and TL in bin 9.03. The QTL with the strongest association to variation in TBN, TL and BZL were detected in the

same or close location during both years. Nine of the QTL for tassel morphology have been related to candidate genes involved in the inflorescence pathway, such as *zmm14, clv28, bif2, mpA*, *ba1, ERF4a, lg2* and *ub1,* however*,* further studies on these genes are needed to validate their involvement in tassel architecture.

REFERENCES

Bailey N (1949) The estimation of linkage with differential viability, II and III. Heredity 3:220-228

Berke TG, Rocheford TR (1999) Quantitative Trait Loci for tassel traits in maize. Crop Sci 39:1439-1443

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

Burr B, Burr FA (1991) Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. Trends Genet 7:55-60

Bommert P, Lunde C, Nardmann J, Vollbrecht E, Running M, Jackson D, Hake S, Werr W (2005a) *thick tassel dwarf1* encodes a putative maize ortholog of the Arabidopsis CLAVATA1 leucine-rich repeat receptor-like kinase. Development 132:1235-1245

Bommert P, Satoh-Nagasawa N, Jackson D, Hirano H-Y (2005b) Genetics and evolution of inflorescence and flower development in grasses. Plant Cell Physiol 46:69-78

Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S (2006) *ramosa2* encodes a LATERAL ORGAN BOUNDARY domain protein that determines the fate of stem cells in branch meristems of maize. Plant Cell 18:574-585

Carlborg O, Jacobsson L, Ahgren P, Siegel P, Andersson L (2006) Epistasis and the release of genetic variation during long-term selection. Nat Genet 38:418-420

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971

Conneally PM, Edwards JH, Kidd KK, Lalouel J-M, Morton NE, Ott J, White R (1985) Report of the committee on methods of linkage analysis and reporting. Cytogen. Cell Genet. 40:356-359

de Leon N, Coors JG, Kaeppler SM (2005) Genetic control of prolificacy and related traits in the Golden Glow Maize Population: II. Genotypic analysis. Crop Sci 45:1370-1378

Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci 33:660-668

Duncan WG, Williams WA, Loomis RS (1967) Tassels and the productivity of maize. Crop Sci 7:37-39

Duvick DN (1974) Continuous backcrossing to transfer prolificacy to a singled-eared inbred line of maize. Crop Sci 14:69-71

Duvick DN (2005) The contribution of breeding to yield advances in maize (*Zea Mays* L.). Adv Agron 86:83-145

Eshed Y, Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141:1147-1162

Evans MMS, Barton MK (1997) Genetics of angiosperm shoot apical meristem development. Annu Rev Plant Physiol Plant Mol Biol 48:673-701

Graham GI, Wolff DW, Stuber CW (1997) Characterization of a yield Quantitative Trait Locus on chromosome five of maize by fine mapping. Crop Sci 37:1601-1610

Hannan EJ, Quinn BG (1979) The determination of the order of an autoregression. J R Soc Stat 41:190-195

Holland JB, Nyquist WE, Cervantez-Martinez CT (2003) Estimating and interpreting heritability for plant breeding: an update. Plant Breeding Rev 22:9-112

Holland JB (2006) Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. Crop Sci 46:642-654

Jiang C, Zeng ZB (1995) Multiple trait analysis of genetic mapping for Quantitative Trait Loci. Genetics 140:1111-1127

Jones DF, Everett HL (1949) Hybrid field corn. Conn Agric Exp Stn Bull, New Haven, CT

Kao C-H, Zeng Z-B (1997) General formulas for obtaining the MLEs and the asymptotic variance- covariance matrix in mapping quantitative trait loci when using the EM algorithm. Biometrics 53:653-665

Kao C-H, Zeng Z-B, Teasdale RD (1999) Multiple Interval Mapping for Quantitative Trait Loci. Genetics 152:1203-1216

Kearsey M (1998) The principles of QTL analysis (a minimal mathematics approach). J Exp Bot 49:1619-1623

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

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits Using RFLP Linkage Maps. Genetics 121:185-199

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181

Laudencia-Chingcuanco D, Hake S (2002) The *indeterminate floral apex1* gene regulates meristem determinacy and identity in the maize inflorescence. Development 129:2629-2638

Levene H (1960) Contributions to probability and statistics: essays in honor of Harold Hotelling. In: Olkin Iea (ed). Stanford University Press, pp 278-292.

Lorieux M, Goffinet B, Perrier X, León DG, Lanaud C (1995) Maximum-likelihood models for mapping genetic markers showing segregation distortion. 1. Backcross populations.

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

Malcomber ST, Preston JC, Reinheimer R, Kossuth JE, Kellogg EA (2006) Developmental gene evolution and the origin of grass inflorescence diversity. Adv Bot Res 44:426-481

Malmberg RL, Held S, Waits A, Mauricio R (2005) Epistasis for fitness-related quantitative traits in *Arabidopsis thaliana* grown in the field and in the greenhouse. Genetics 171:2013-2027

McSteen P, Laudencia-Chingcuanco D, Colasanti J (2000) A floret by any other name: control of meristem identity in maize. Trends Plant Sci 5:61-66

Mickelson SM, Stuber CS, Senior L, Kaeppler SM (2002) Quantitative trait loci controlling leaf and tassel traits in a B73 x Mo17 population of maize. Crop Sci 42:1902-1909

Orr AR, Sundberg MD (1994) Inflorescence development in a perennial teosinte: *Zea perenni*s (Poaceae). Am J Bot 81:598-608

Peripato AC, de Brito RA, Vaughn TT, Pletscher LS, Matioli SR, Cheverud JM (2002) Quantitative Trait Loci for maternal performance for offspring survival in mice. Genetics 162:1341-1353

Rieseberg LH, Widmer A, Arntz AM, Burke JM (2003) The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. Philos T R Soc B 358:1141-1147

Ritchie SW, Hanway JJ, Benson GO (1996) How a corn plants develops. In: Serv ISUCE (ed) Spec Rep. Iowa State University, Ames, IA.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci U S A 81:8014-8018

Salvi S, Tuberosa R, Chiapparino E, Maccaferri M, Veillet S, van Beuningen L, Isaac P, Edwards K, Phillips RL (2002) Toward positional cloning of *Vgt1*, a QTL controlling the transition from the vegetative to the reproductive phase in maize. Plant Mol Biol 48:601-613

Sanchez AC, Brar DS, Huang N, Li Z, Khush GS (2000) Sequence Tagged Site marker-assisted selection for three bacterial blight resistance genes in rice. Crop Sci 40:792-797

SAS (2007) SAS/STAT Software. 9.1.3 edn. SAS Institute Inc. , Cary, NC.

Szalma S, Hostert B, LeDeaux J, Stuber C, Holland J (2007) QTL mapping with Near-Isogenic Lines in maize. Theor Appl Genet 114:1211-1228

Tao Y, Zeng Z-B, Li J, Hartl DL, Laurie CC (2003) Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. II. Mapping hybrid male sterility loci on the third chromosome. Genetics 164:1399-1418

Upadyayula N, da Silva H, Bohn M, Rocheford T (2006a) Genetic and QTL analysis of maize tassel and ear inflorescence architecture. Theor Appl Genet 112:592-606

Upadyayula N, Wassom J, Bohn M, Rocheford T (2006b) Quantitative trait loci analysis of phenotypic traits and principal components of maize tassel inflorescence architecture. Theor Appl Genet 113:1395-1407

Vollbrecht E, Springer PS, Goh L, Buckler Iv ES, Martienssen R (2005) Architecture of floral branch systems in maize and related grasses. Nature 436:1119-1126

Walsh J, Freeling M (1999) The *liguleless2* gene of maize functions during the transition from the vegetative to the reproductive shoot apex. Plant J 19:489-495

Wang S, Basten CJ, Zeng Z-B (2007) Windows QTL Cartographer 2.5 edn. Department of Statistics, North Carolina State University, Raleigh, NC.

Wright S (1968) Evolution and the genetics of populations. I. Genetic and biometric foundations. Univ. Chicago Press., Chicago

Zamir D, Eshed D (1998) Case history in germplasm introgression: tomato genetics and breeding using nearly isogenic introgression lines derived from wild species. In: Paterson A (ed) Molecular dissection of complex traits. CrC Press LLC, New York, pp 207-217

Zeng Z-B. KC-H, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. Genet Res Camb 74:11

Zeng ZB (1994) Precision mapping of Quantitative Trait Loci. Genetics 136:1457-1468

Table 1: Summary of QTL for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) reported in previous and in the current study

Genetic Map Location[g]	TBN	TL	CSL	BZL
	- number of times reported in the literature [h] -			
1.01	1[b*], 1[f]	1[b*]	1[b*]	1[b*]
1.02	6[b*]	6[b*]	6[b*]	1[a], 6[b*]
1.03	2[b*]	1[a], 2[b*]	1[a], 2[b*]	2[b*]
1.03-1.06	1[c*]	1[c*]	1[c*]	1[c*]
1.04	1[b*]	1[b*]	1[b*]	1[b*]
1.05	2[b*]	2[b*]	2[b*]	2[b*]
1.06	1[a], 4[b*]	1[a], 4[b*]	4[b*]	1[a], 4[b*]
1.06/1.07	1[c*]	1[c*]	1[c*]	1[c*]
1.07	1[b*]	1[b*]	1[b*]	1[b*]
1.09	4[b*]	4[b*]	4[b*]	4[b*]
1.09/1.10	1[c*]	1[c*]	1[c*]	1[c*]
1.10	3[b*], 3[d*]	3[b*], 3[d*],3[d]	3[b*], 3[d*]	3[b*], 3[d], 3[d*]

[a] Reported by Mayor et al. herein on C103×C103AP RILs population
[b] Reported by the Maize Inflorescence Portal on the Intermated B73×Mo17 (IBM) population (www.maizegdb.org/MIP/; verified 7/24/08).
[c] Reported by the Maize Inflorescence Portal and Upadyayula et al. (2006b) on the (ILP×B73) B73 BC1S1 population (www.maizegdb.org/MIP/; verified 7/24/08)
[d] Reported by the Maize Inflorescence Portal and Upadyayula et al. (2006b) on the (IHO×B73) B73 BC1S1 population (www.maizegdb.org/MIP/; verified 7/24/08)
[e] Reported by Berke and Rocheford (1999) on the IHO×ILO(EM) population
[f] Reported by Mickelson et al. (2002) on the B73×Mo17 RILs population
[*] QTL detected by principal component analysis on: TL, CSL, TBN, Branch Length, Tassel branch angle (TBA), Tassel weight (TW), Central spike density (CSD) and Primary branch spikelet density (PSD)
[g] Bin location according to MaizeGDB. A bin is an arbitrary subdivision of the maize genome that serves as a collection point for mapped genetic loci. The segments are designated with the chromosome number followed by a two-digit decimal (i.e. 1.01 corresponds to bin 1 from chromosome 1).
[h] The numeral corresponds to the number of times a QTL for a given trait was detected in that bin (i.e. For bin 1.01, 1 QTL for PC was reported in population “b” and 1 QTL for TBN was described in population “f”).

Table 1 Cont.

Genetic Map Location[g]	TBN	TL	CSL	BZL
1.12	1^{b*}	1^{b*}	1^{b*}	1^{b*}
2.00	1^{b*}	1^{b*}	1^{b*}	1^{b*}
2.01	1^{a}, 1^{b*}	1^{b*}	1^{a}, 1^{b*}	1^{a}, 1^{b*}
2.02	1^{b*}, 1^{f}	1^{b*}	1^{b*}	1^{b*}
2.03	2^{b*}	2^{b*}	2^{b*}	2^{b*}
2.04	10^{b*}, 5^{d*}	10^{b*}, 5^{d*}	10^{b*}, 5^{d*}	10^{b*}, 5^{d*}
2.05	1^{f}			
2.06	1^{a}			
2.07	1^{b*}, 1^{e}, 1^{f}	1^{b*}	1^{b*}	1^{b*}
2.08	5^{b*}, 1^{c*}, 1^{d*}	1^{a}, 5^{b*}, 1^{c*}, 1^{d*}	1^{a}, 5^{b*}, 1^{c*}, 1^{d*}	1^{a}, 5^{b*}, 1^{c*}, 1^{d*}
2.09	4^{b*}	4^{b*}	4^{b*}	4^{b*}
3.01	1^{b*}	1^{b*}	1^{b*}	1^{b*}
3.02	2^{b*}	2^{b*}	2^{b*}	2^{b*}
3.04	8^{b*}	8^{b*}	8^{b*}	8^{b*}
3.05	1^{a}, 2^{b*}, 4^{c*}, 1^{f}	$2^{b*}4^{c*}$	2^{b}, 2^{b*}, 4^{c*}	$2^{b*}4^{c*}$
3.05/3.06			1^{d}	
3.07	1^{b*}	1^{b*}	1^{b*}	1^{b*}
3.06	4^{b*}	4^{b*}	4^{b*}	4^{b*}
3.08	1^{b*}	1^{b*}	1^{b*}, 1^{a}	1^{b*}
3.09	1^{b*}, 1^{c*}	1^{a}, 1^{b*}, 1^{c*}	1^{b*}, 1^{c*}	1^{b*}, 1^{c*}
3.10				1^{a}
4.01	1^{b*}	1^{b*}	1^{b*}	1^{b*}
4.02	2^{b*}	2^{b*}	2^{b*}	2^{b*}
4.03	2^{b*}	2^{b*}	2^{b*}	2^{b*}
4.05	1^{b*}, 1^{c}	1^{b*}	1^{b}, 1^{b*}, 1^{c}	1^{b*}
4.05/4.06	1^{c*}	1^{c*}, 1^{d}	1^{c*}	1^{c*}
4.06	1^{b*}	1^{b}, 1^{b*}	1^{a}, 1^{b*}	1^{b*}
4.07	1^{b*}	1^{b*}	1^{b*}	1^{b*}
4.08	1^{a}, 4^{b*}, 1^{c*}, 1^{e} 1^{f}	1^{a}, 4^{b*}, 1^{c*}	1^{b}, 1^{c}, 4^{b*}, 1^{c*}	1^{a}, 4^{b*}, 1^{c*}
4.09	2^{b*}	2^{b*}	2^{b*}	2^{b*}, 1^{d}
4.11	1^{b*}	1^{b*}	1^{b*}	1^{b*}

Table 1 Cont.

Genetic Map Location[g]	TBN	TL	CSL	BZL
5.00	1[b*]	1[b*]	1[b*]	1[b*]
5.01	2[b*]	2[b*]	2[b*]	2[b*]
5.02	1[b*]	1[b*]	1[b*]	1[b*]
5.02/5.03		1[d]		
5.03	2[b*]	2[b], 2[b*]	2[b*]	2[b*]
5.04	1[b*], 2[c*]	1[b*], 2[c*]	1[b*], 2[c*]	1[b*], 2[c*]
5.05	1[b*]	1[b*]	1[b*]	1[b*]
5.06	2[b*]	2[b*]	2[b*], 2[c]	2[b*]
5.06/5.07	1[d], 1[d*]	1[d*]	1[d*]	1[d*]
5.08	1[b*]	1[b*]	1[b*]	1[b*]
6.00	1[b*]	1[b*]	1[b*]	1[b*]
6.01	1[a]			
6.02	1[b*]	1[b*]	1[b*]	1[b*]
6.05	6[b*]	6[b*]	6[b*]	6[b*]
6.06/6.07	1[d*]	1[d], 1[d*]	1[d*]	1[d*]
6.07	1[b*], 1[c*]	1[b*], 1[c*]	1[b*], 1[c*], 1[c]	1[b*], 1[c*]
6.08	2[b*]	2[b*]	2[b*]	2[b*]
7.00	6[b*]	6[b*]	6[b*]	6[b*]
7.00/7.02	1[c], 1[c*]	1[c*]	1[c*]	1[c*]
7.01	2[b*]	2[b*]	2[b*]	2[b*]
7.02	1[b], 2[b*], 1[d], 1[d*]	1[b], 2[b*], 1[d], 1[d*]	2[b*], 1[d*]	1[b], 2[b*], 1[d], 1[d*]
7.03	1[e], 2[b*]	2[b*]	2[b*]	2[b*]
7.04	4[b*]	4[b*]	4[b*]	4[b*]
7.06	1[b*]	1[b*]	1[b*]	1[b*]
8.01	1[b*]	1[b*]	1[b*]	1[b*]
8.02	1[c*]	1[c*]	1[c*]	1[c*]
8.02/8.03				1[d]
8.03	3[b*]	3[b*]	3[b*]	1[b], 3[b*]
8.04		1[a]		1[a]
8.05	10[b*]	10[b*]	10[b*]	10[b*]
8.06	1[a], 1[b*]	1[b*]	1[b*]	1[a], 1[b*]
8.07	1[b*]	1[b*]	1[b*]	1[b*]

Table 1 Cont.

Genetic Map Location[g]	TBN	TL	CSL	BZL
8.08	3[b*]	3[b*]	3[b*]	3[b*]
8.09	1[b*]	1[b*]	1[b*]	1[b*]
9.00	1[b*]	1[b*]	1[b*]	1[b*]
9.01	3[b*]	3[b*]	3[b*]	3[b*]
9.02	5[b*], 1[d*]	5[b*], 1[d*]	5[b*], 1[d*]	5[b*], 1[d*]
9.02/9.03	1[d*]	1[d*]	1[d*]	1[d*]
9.03	2[b*]	1[a], 2[b*]	1[a], 2[b*]	2[b*]
9.04	1 [a], 1[b*]	1[b*]	1[b*]	1[b*]
9.05	2[b*]	2[b*]	2[b*]	2[b*]
9.06	1[b*]	1[b*]	1[b*]	1[b*]
9.07	1[b*]	1[b*]	1[b*]	1[b*]
10.01	1[b*]	1[b*]	1[b*]	1[b*]
10.02			2[b], 2[d]	
10.02/10.04	1[d*]	1[d*]	1[d*]	1[d*]
10.03	1[c*]	1[c*]	1[c*]	1[c*]
10.04	3[b*]	3[b*]	3[b*]	3[b*]
10.06	2[b*]	2[b*]	2[b*]	2[b*]
10.07	5[b*]	5[b*]	5[b*]	5[b*]

Table 2: Monthly precipitation and GDU accumulation for the period 1987-2007 and for the growing season 2006-2007 at the Agronomy and Agricultural Engineering Research Center (AAERC) near Boone, Iowa (42°1´ N, 93°45´ W)

	GDU accumulation (ºC)[a]			Precipitation (mm)		
	2006	2007	1987-2007	2006	2007	1987-2007
	Monthly accumulation		Average[b]	Monthly accumulation		Average[b]
May	222	265	211	58.2	146.3	113.5
June	350	345	336	13.5	46.2	113.8
July	419	407	396	79.8	66.8	122.0
August	376	416	370	58.7	185.1	97.9

[a] Growing degree units were calculated according to Cross and Zuber (1972).
[b] Average monthly accumulation for the period 1987-2007.

Table 3: Mean phenotypic values of maize inbreds C103, C103AP and 218 RILs, variance components and broad-sense heritability for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL)

	TBN	TL	CSL	BZL
	--- number ---	---------- cm ----------		
		Mean Phenotypic Values[a]		
C103	11.14 ± 2.27	42.80 ± 2.65	31.09 ± 2.18	11.70 ± 1.54
C103AP	11.87 ± 2.25	36.20 ± 2.63	26.97 ± 2.16	9.23 ± 1.52
Mean RILs	8.57 ± 0.15	35.97 ± 0.18 *, [b]	26.88 ± 0.15	9.09 ± 0.10
		Variance Components		
σ^2_G	4.81 ± 0.49***	6.55 ± 0.68***	4.27 ± 0.46***	2.20 ± 0.23***
$\sigma^2_{G\times E}$	0.25 ± 0.05***	0.54 ± 0.11***	0.58 ± 0.10***	0.10 ± 0.04***
σ^2	0.53 ± 0.04	0.99 ± 0.07	0.81 ± 0.06	0.47 ± 0.03
		Broad-Sense Heritability		
	0.95 ± 0.007	0.93 ± 0.01	0.90 ± 0.01	0.93 ± 0.01

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
***Significant at the 0.001 probability level.
[a] Tukey multiple comparison test.
[b] Significant at the 0.01 probability level for the difference between C103 and the mean of the RILs.
[c] Refers to the combined analysis of the 2006 and 2007 datasets.

Table 4: Phenotypic (above diagonal) and genetic correlations (below diagonal) with their standard errors (in parenthesis) for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) in a maize population of 218 RILs derived from inbreds C103 and C103AP in the combined analysis of the 2005, 2006 and 2007 datasets

Trait	TBN	TL	CSL	BZL
TBN	-	0.17 (0.06)	-0.28 (0.05)	0.70 (0.03)
TL	0.21 (0.07)	-	0.81 (0.02)	0.55 (0.04)
CSL	-0.28 (0.07)	0.81 (0.03)	-	-0.05 (0.06)
BZL	0.75 (0.03)	0.59 (0.05)	0.01 (0.08)	-

[a] Phenotypic and genetic correlation coefficients were calculated using REML from the MIXED procedure in SAS. Approximate sampling variances and standard errors for the estimates were obtained after Holland (2006).

Table 5: Summary of QTL detected from a population of 218 maize RILs derived from inbred lines C103 and C103AP for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL)

Trait	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive effect[d]	Parental allele[e]	Partial $R^{2,f}$
						- number -		
TBN	1	1.06	bnlg1057	6.95	4.45	0.47	AP	6.1
	2	2.01	umc1165	17.03	11.73	0.80	AP	13.1
	3	206	bmc1329	67.18	9.02	0.77	AP	11.8
	4	3.05	umc1167	0.01	4.40	0.38	AP	3.7
	5	4.08	phi092	19.23	4.46	0.51	C103	8.2
	6	6.01	bnlg107	3.01	3.80	0.46	C103	3.5
	7	8.06	bnlg1031	84.68	6.80	0.63	AP	8.1
	8	9.04	umc1771	72.22	1.06	0.24	AP	1.5
				Epistatic interactions				
	umc1165 × umc1167				1.56	0.29		2.5
	bmc1329 × bnlg107				1.71	-0.35		1.4
							Total $R^{2,g}$=	59.9
						-- cm --		
TL	1	1.03	umc1021	54.61	1.37	0.35	AP	3.3
	2	1.06	bnlg1598	3.01	3.09	0.55	AP	7.6
	3	2.08	umc1798	118.77	3.82	0.56	AP	6.4

[a] Chromosomal bin corresponding to the location of the nearest left marker according to MaizeGDB.
[b] Position in cM from the top of the chromosome calculated by MapMaker v.3
[c] LOD score corresponding to the position of the QTL as calculated by WinQTL cartographer. For epistatic interactions, the LOD score corresponds to the effect of the interaction.
[d] Additive effects values calculated as the average from the difference between homozygotes for each parental allele at a locus.
[e] Direction of the additive effect, AP= C103AP.
[f] Proportion of the phenotypic variance explained by each QTL defined in the multiple regression model.
[g] Proportion of the phenotypic variance explained by the sum of the QTL defined in the multiple regression model.
[h] Refers to the combined analysis of the 2006 and 2007 datasets.

Table 5 Cont.

Trait	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive effect[d]	Parental allele[e]	Partial $R^{2,f}$
	4	3.06	umc1985	9.45	3.06	0.39	C103	2.7
	5	3.09	umc2008	37.16	2.20	0.41	C103	4.2
	6	4.08	phi092	11.23	3.31	0.51	C103	6
	7	8.04	umc1858	8.84	5.27	0.79	C103	10.3
	8	9.03	mmc0051	31.47	5.28	0.51	C103	5.9
				Epistatic interactions				
	umc1985 × mmc0051				1.36	0.35		3.1
							Total $R^{2,g}$ =	49.5
						-- cm--		
CSL	1	1.03	phi001	61.93	3.13	0.42	AP	4.9
	2	2.01	umc1165	17.03	6.75	0.63	C103	7.5
	3	2.08	bmc1233	116.19	4.33	0.62	AP	7.4
	4	3.06	umc1644	5.01	3.90	0.29	C103	3.9
	5	3.08	phi046	8.01	2.19	0.43	C103	7.2
	6	4.06	umc2027	46.45	2.62	0.44	C103	2
	7	8.05	umc1889	18.90	6.67	0.48	C103	5.3
	8	9.03	mmc0051	30.47	4.79	0.53	C103	8.8
				Epistatic interactions				
	umc1644× umc2027				0.93	0.29		2.7
	bmc1233× umc1889				1.37	0.34		1
							Total $R^{2,g}$ =	50.7
						-- cm --		
BZL	1	1.02	umc1568	1.01	3.22	0.29	AP	2.8
	2	1.06	bnlg1057	6.95	8.85	0.48	AP	13.3
	3	2.01	umc1165	17.03	7.52	0.46	AP	11.1
	4	2.08	umc1798	118.77	1.45	0.15	AP	2.1
	5	3.10	umc1136	7.63	3.40	0.30	C103	4.9
	6	4.08	phi086	23.84	7.20	0.46	C103	11.3
	7	8.04	phi014	0.01	3.18	0.30	C103	4.6
	8	8.06	bnlg1031	84.68	2.75	0.27	AP	2.2
							Total $R^{2,g}$ =	52.3

Figure 1: Graphical representation of the measurements assessed on the tassels adapted from Upadyayula et al (2006a). TL tassel length, BZL branching zone length, CSL central spike length

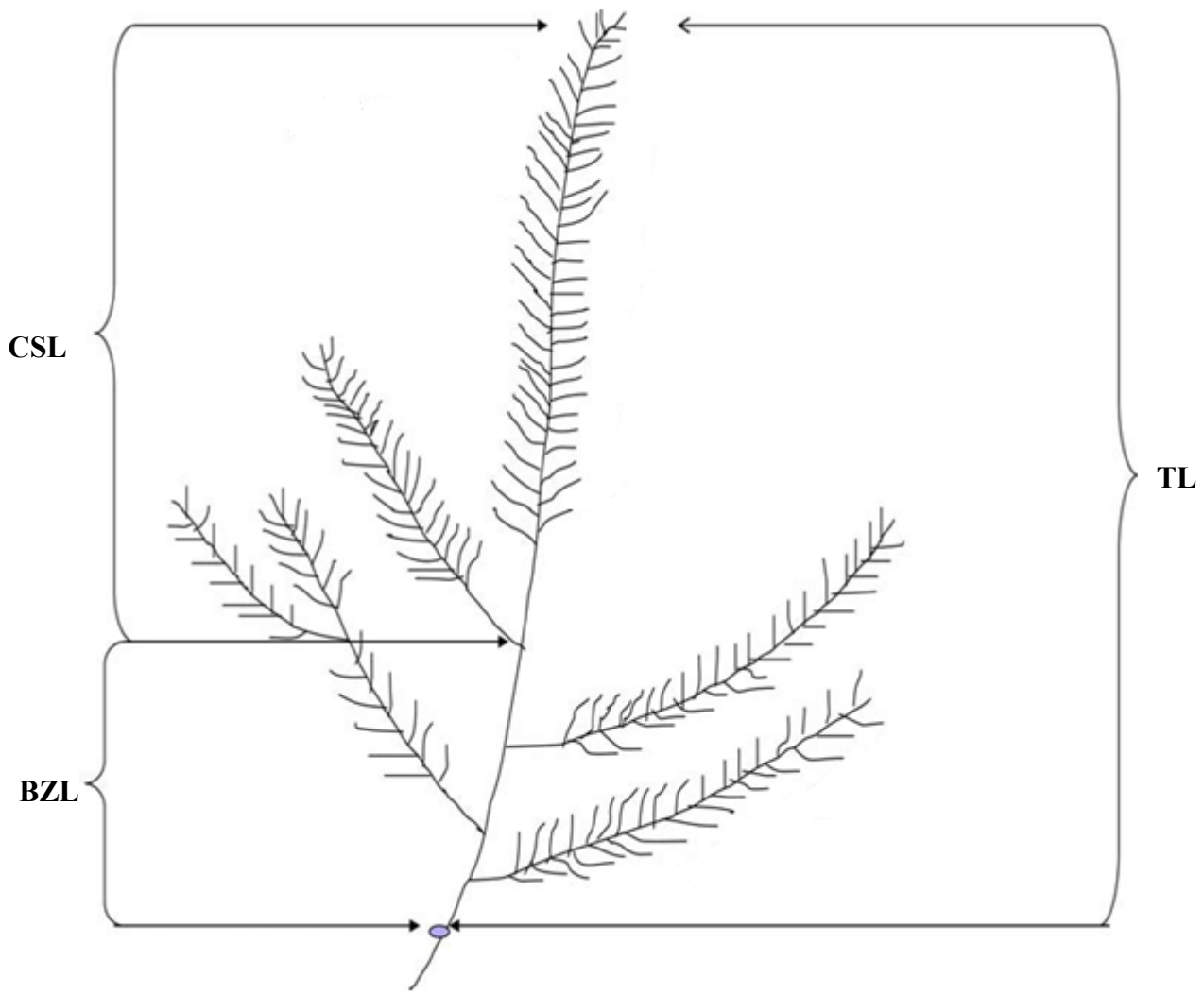


Figure 2: Linkage map of the C103/C103AP population constructed with information from 218 RILs. Relative positions of loci are given in centimorgans at the left side of the linkage groups. Loci names are given at the right side of the chromosomes. QTLs with 1-LOD support intervals for the combined analysis are indicated. Dotted lines represent genomic regions not polymorphic between C103 and C103AP and connect linkage groups of the same chromosome

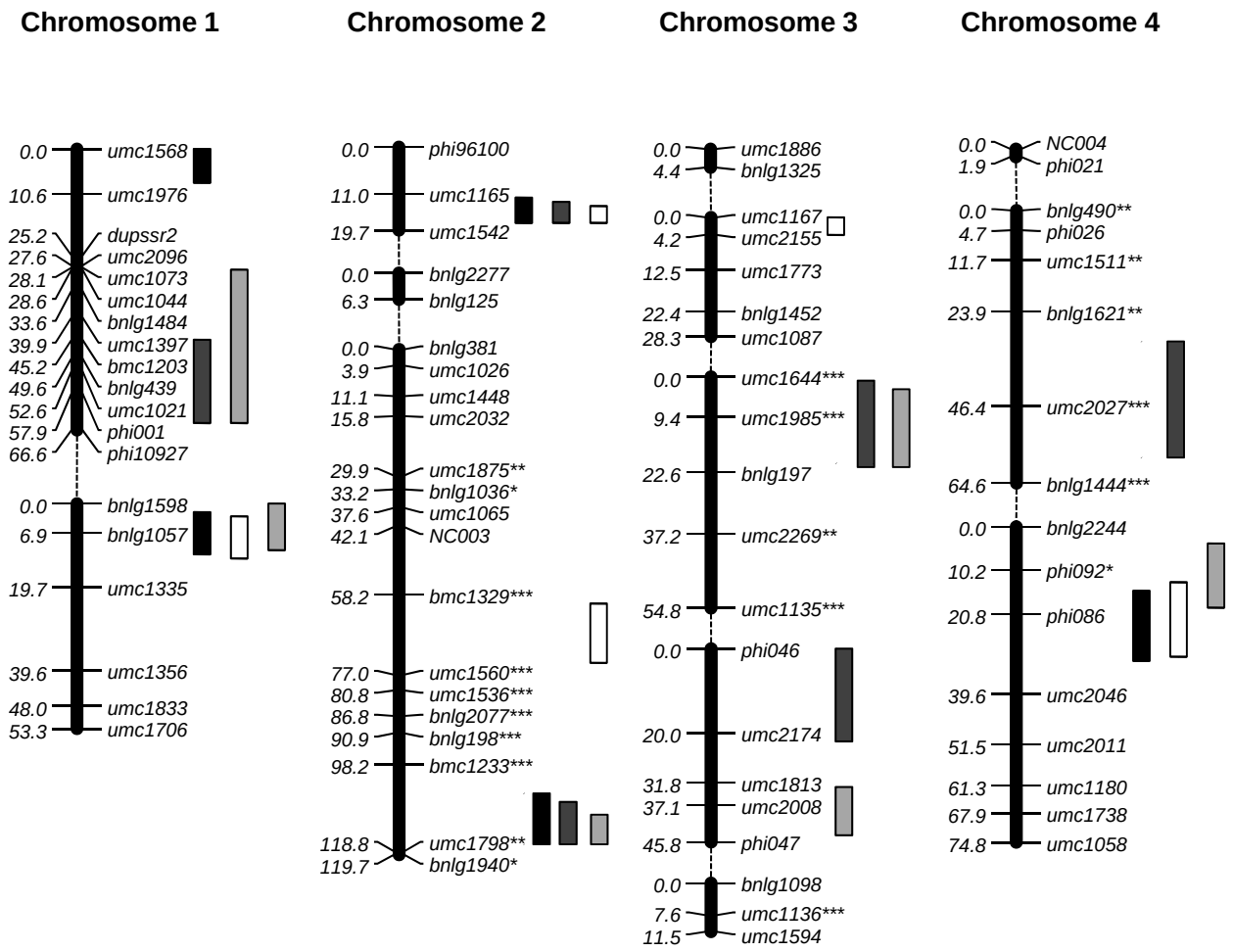
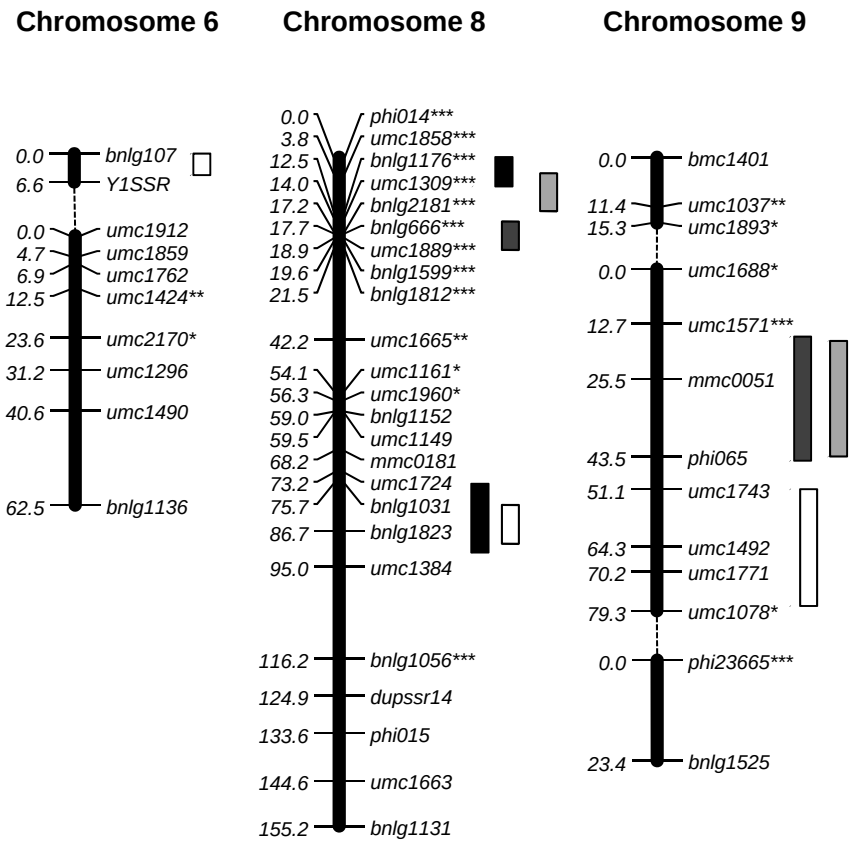


Figure 2 Cont.



Tassel Branch Number | Tassel Length | Central Spike Length | Branching zone length

*Departure from a genotypic ratio of 1:1 significant at the 0.05 probability level
** Departure from a genotypic ratio of 1:1 significant at the 0.01 probability level
*** Departure from a genotypic ratio of 1:1 significant at the 0.001 probability level

APPENDIX

Table A1: Mean phenotypic values of inbreds C103, C103AP and 218 RILs, variance components and broad-sense heritability for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) in years 2006 and 2007

		Mean Phenotypic Values[a]					
Trait	Environment	C103	C103AP	RILs	σ_G^2	σ^2	H^2
TBN (number)	2006	11.21 ± 2.35	11.53 ± 2.29	8.78 ± 0.16	5.00 ± 0.51***	0.67 ± 0.07	0.94 ±0.01
	2007	11.13 ± 2.33	12.13 ± 2.31	8.36 ± 0.16	5.15 ± 0.51***	0.37 ± 0.04	0.97 ± 0.05
TL (cm)	2006	43.33 ± 2.76	35.70 ± 2.71	35.93 ± 0.19*, [b]	7.17 ± 0.74***	1.00 ± 0.10	0.93 ± 0.01
	2007	42.14 ± 2.76	36.57 ± 2.70	36.02 ± 0.19	7.08 ± 0.73***	1.00 ± 0.10	0.93 ± 0.01
CSL (cm)	2006	31.98 ± 2.29	26.39 ± 2.24	26.69 ± 0.15	4.81 ± 0.51***	0.82 ± 0.08	0.92 ± 0.01
	2007	30.08 ± 2.31	27.48 ± 2.25	27.08 ± 0.16	4.87 ± 0.51***	0.81 ± 0.08	0.92 ± 0.01
BZL (cm)	2006	11.35 ± 1.63	9.31 ± 1.59	9.25 ± 0.11	2.47 ± 0.26***	0.36 ± 0.04	0.93 ± 0.01
	2007	12.05 ± 1.57	9.10 ± 1.52	8.93 ± 0.11	2.12 ± 0.23***	0.58 ± 0.06	0.88 ± 0.02

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
***Significant at the 0.001 probability level.
[a] Tukey multiple-comparison test.
[b] Difference between C103 and the mean of the RILs.

Table A2: Phenotypic (above diagonal) and genetic correlations with their standard errors (in parenthesis) for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) in years 2006 and 2007

	Environment	TBN	TL	CSL	BZL
TBN	2006	-	0.15 (0.06)	-0.30 (0.06)	0.68 (0.03)
	2007	-	0.19 (0.06)	-0.26 (0.06)	0.70 (0.03)
TL	2006	0.16 (0.07)	-	0.81 (0.02)	0.55 (0.04)
	2007	0.20 (0.07)	-	0.82 (0.02)	0.55 (0.04)
CSL	2006	-0.32 (0.07)	0.81 (0.03)	-	-0.05 (0.06)
	2007	-0.27 (0.07)	0.83 (0.02)	-	-0.04 (0.06)
BZL	2006	0.73 (0.04)	0.57 (0.05)	-0.02 (0.07)	-
	2007	0.77 (0.03)	0.56 (0.05)	0.01 (0.07)	-

[a] Phenotypic and genetic correlation coefficients were calculated using REML from the MIXED procedure in SAS. Approximate sampling variances and standard errors for the estimates were obtained after Holland (2006).

Table A3: Summary of QTL detected from a population of 218 maize RILs derived from inbred lines C103 and C103AP for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) from a population of 218 RILs in years 2006 and 2007

Trait	Year	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive effect[d]	Parental allele[e]	Partial $R^{2, f}$
							- number -		
TBN	2006	1	1.06	bnlg1057	6.95	4.67	0.48	AP	6
		2	2.01	umc1165	17.03	9.57	0.73	AP	10.5
		3	2.06	NC003	52.07	2.00	0.42	AP	6.4
		4	2.08	bnlg198	94.92	4.56	0.62	AP	9.2
		5	3.05	umc1167	2.01	5.29	0.38	AP	3.4
		6	4.08	phi086	20.84	4.25	0.47	C103	7.3
		7	6.06-6.07	umc1296	31.20	3.82	0.34	AP	1.5
		8	8.05	umc1889	18.90	1.89	0.41	AP	2.6
		9	8.06	bnlg1031	83.68	5.29	0.58	AP	7.2
		Epistatic interactions							
		umc1165× umc1167				0.92	0.23		1.8
		umc1167× umc1296				1.91	-0.32		1.3
								Total $R^{2,g}$ =	57.2
	2007	1	1.06	bnlg1057	6.95	4.56	0.50	AP	6.1
		2	2.01	umc1165	17.03	10.27	0.77	AP	12.5
		3	2.06	bnlg1036	35.23	4.34	0.45	AP	6.2
		4	2.08	bnlg198	94.92	3.36	0.44	AP	5

[a] Chromosomal bin corresponding to the location of the nearest left marker according to MaizeGDB.
[b] Position in cM from the top of the chromosome calculated by MapMaker v.3
[c] LOD score corresponding to the position of the QTL as calculated by WinQTL cartographer
[d] Additive effects values calculated as the average from the difference between homozygotes for each parental allele at a locus.
[e] AP= C103AP
[f] Proportion of the phenotypic variance explained by each QTL defined in the multiple regression model.
[g] Proportion of the phenotypic variance explained by the sum of the QTL defined in the multiple regression model.

Table A3 Cont.

Trait	Year	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive effect[d]	Parental allele[e]	Partial $R^{2,f}$
		5	3.05	umc1167	0.01	5.21	0.48	AP	4.9
		6	4.08	phi086	20.84	4.92	0.53	C103	7.6
		7	6.01	bnlg107	6.01	3.53	0.38	C103	2.9
		8	8.06	bnlg1031	84.68	5.64	0.59	AP	6.4
		Epistatic interactions							
		umc1165× umc1167				1.39	0.28		2.1
		bnlg198× umc1167				1.16	0.29		0.7
		umc1165× bnlg107				1.41	-0.29		1.4
								Total $R^{2,g}$ =	55.8
TL							--- cm ---		
	2006	1	1.06	bnlg1598	2.01	4.63	0.69	AP	8.4
		2	2.08	umc1798	118.77	3.59	0.60	AP	6.7
		3	3.08	phi046	9.01	2.97	0.64	C103	7.2
		4	4.08	phi092	13.23	3.08	0.58	C103	6.4
		5	8.04	umc1858	9.84	3.46	0.69	C103	7.9
		6	9.02	umc1037	11.44	2.21	0.45	C103	3.7
								Total $R^{2,g}$ =	40.3
	2007	1	1.03	phi001	58.93	2.59	0.51	AP	6.5
		2	1.06	bnlg1598	5.01	2.93	0.57	AP	7.7
		3	3.06	umc1985	18.45	2.67	0.55	C103	5.5
		4	3.09	umc2008	37.16	1.16	0.32	C103	3
		5	4.08	phi092	14.23	2.35	0.48	C103	4.8
		6	8.04	phi014	0.01	1.42	0.42	C103	4
		7	8.05	bnlg1176	12.54	2.42	0.59	C103	7.1
		8	9.03	mmc0051	33.47	5.64	0.75	C103	9.4
								Total $R^{2,g}$ =	48
CSL							--- cm ---		
	2006	1	2.01	umc1165	19.03	3.45	0.52	C103	4.7
		2	2.08	umc1798	118.77	1.84	0.38	AP	3.7
		3	3.08	phi046	5.01	3.39	0.58	C103	8.5
		4	8.05	umc1889	18.90	2.21	0.47	C103	4

Table A3 Cont.

Trait	Year	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive effect[d]	Parental allele[e]	Partial $R^{2,f}$
		5	9.02	umc1037	14.44	2.84	0.48	C103	5.9
								Total $R^{2,g}$ =	26.8
	2007	1	1.03	phi001	59.93	5.93	0.59	AP	7.8
		2	2.01	umc1165	17.03	7.24	0.74	C103	8.6
		3	2.02	bnlg2277	0.01	2.88	0.37	AP	1
		4	2.08	bmc1233	116.19	4.55	0.53	AP	5.3
		5	3.06	umc1985	17.45	8.15	0.72	C103	12.5
		6	4.06	umc2027	53.45	1.30	0.52	C103	2.5
		7	8.05	umc1889	18.90	7.62	0.53	C103	5.8
		8	9.03	mmc0051	29.47	9.02	0.74	C103	13.5
		Epistatic interactions							
		bnlg2277× bmc1233				1.24	0.27		1.2
		bmc1233× umc1889				1.58	0.37		0.8
								Total $R^{2,g}$ =	59
BZL							--- cm ---		
	2006	1	1.02	umc1568	2.01	2.73	0.29	AP	2.8
		2	1.06	bnlg1057	6.95	9.01	0.52	AP	13.8
		3	2.01	umc1165	19.03	7.79	0.46	AP	10.5
		4	4.08	phi086	25.84	8.12	0.56	C103	13.8
		5	8.05	bnlg1176	12.54	3.17	0.34	C103	5.5
		6	8.07	bnlg1823	86.70	3.03	0.24	AP	1.9
		Epistatic interactions							
		umc1165 × bnlg1823				1.08	-0.17		1.1
								Total $R^{2,g}$ =	49.4
	2007	1	1.02	umc1568	0.01	3.42	0.25	AP	2.2
		2	1.06	bnlg1057	6.01	7.31	0.46	AP	11.7
		3	2.01	umc1165	19.03	5.29	0.38	AP	8.9
		4	3.05	umc1167	0.01	2.13	0.23	AP	3.9
		5	4.08	phi086	24.84	7.67	0.48	C103	12.6
		6	8.04	phi014	0.01	1.68	0.22	C103	3.2
		Epistatic interactions							

Table A3 Cont.

umc1568 × phi086	1.23	0.19		0.2
			Total $R^{2,g}$ =	42.7

Table A4: Regression coefficients from single-marker analysis at loci of chromosomes 2, 5 and 10 from a maize population of 218 RILs derived from inbreds C103 and C103AP for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL)

	Chromosome location[b]						
	2.08	5.00	5.01	5.07		10.06	
	---------- SSR Loci ----------						
Environment	mmc0381	umc1325	umc1766	bnlg1346	phi085	bnlg1028	umc1477
	---------- TBN ----------						
2006	-0.07	0.07	-0.13	0.35*	-0.34*	0.17	-0.32*
2007	-0.26	0.01	-0.20	0.30	-0.22	0.26	-0.30*
Combined[a]	-0.16	0.04	-0.16	0.33*	-0.28	0.21	-0.31*
	---------- TL ---------						
2006	-0.71***	0.24	0.06	0.12	-0.01	0.48**	-0.20
2007	-0.42*	0.09	0.05	0.29	-0.31	0.37*	-0.17
Combined[a]	-0.56**	0.16	0.06	0.20	-0.16	0.43*	-0.18
	---------- CSL ----------						
2006	-0.41*	0.09	0.07	-0.12	0.24	0.17	0.05
2007	-0.14	-0.07	0.04	0.05	-0.11	0.05	0.01
Combined[a]	-0.27	0.02	0.06	-0.03	0.06	0.11	0.03
	---------- BZL ----------						
2006	-0.29*	0.14	-0.01	0.24*	-0.25*	0.31**	-0.24*
2007	-0.27*	0.14	0.01	0.22*	-0.18*	0.30**	-0.16
Combined[a]	-0.29**	0.14	-0.01	0.24*	-0.22*	0.31**	-0.21*

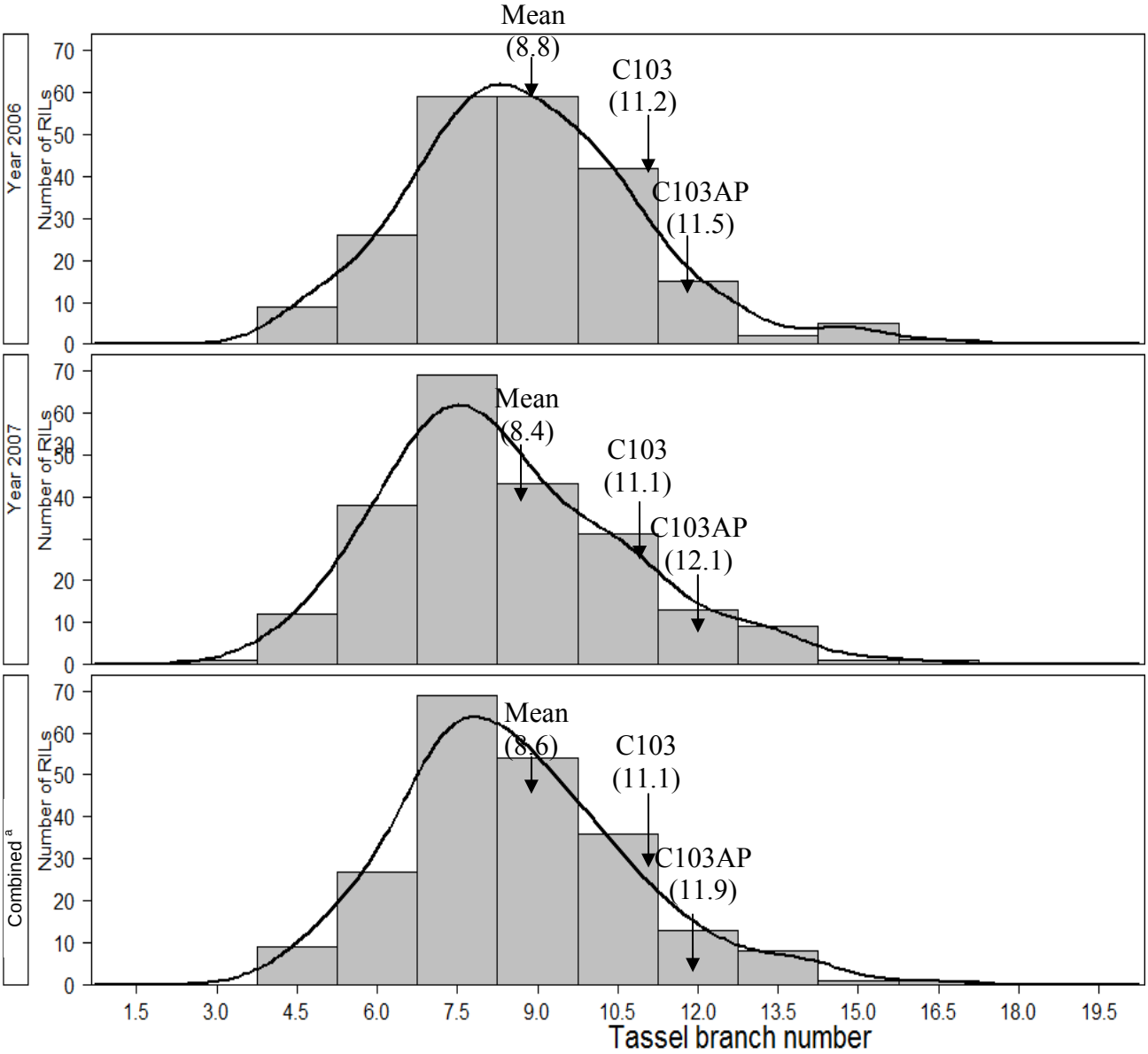
* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
*** Significant at the 0.001 probability level.
[a] Refers to the combined analysis across years 2005, 2006 and 2007.
[b] Chromosomal bin corresponding to the location of the SSR locus according to MaizeGDB database.

Table A5: Summary of QTL detected by Multiple-trait Composite Interval Mapping in the three years of evaluation (2005, 2006 and 2007) for tassel branch number (TBN), tassel length (TL), central spike length (CSL) and branching zone length (BZL) from a population of 218 RILs derived from inbreds C103 and C103AP

Trait	QTL	Location[a]	Nearest locus	Position[b]	LOD[c]	Additive effect	Parental allele[d]
						- number -	
TBN	1	2.01	phi96100	8.00	5.61	0.69	AP
	2	2.01	umc1165	17.00	8.01	0.79	AP
	3	2.06	NC003	50.10	3.79	0.69	AP
	4	4.08	phi086	24.80	5.68	0.80	C103
						-- cm --	
TL	1	1.06	bnlg1598	2.00	5.25	0.75	AP
	2	8.04	phi014	2.00	3.69	0.79	C103
	3	8.04	umc1858	11.80	6.00	0.98	C103
	4	8.05	bnlg666	17.70	5.15	0.89	C103
	5	9.03	mmc0051	27.50	4.43	0.73	C103
						-- cm--	
CSL	-	-	-	-	-	-	
						-- cm --	
BZL	-	-	-	-	-	-	

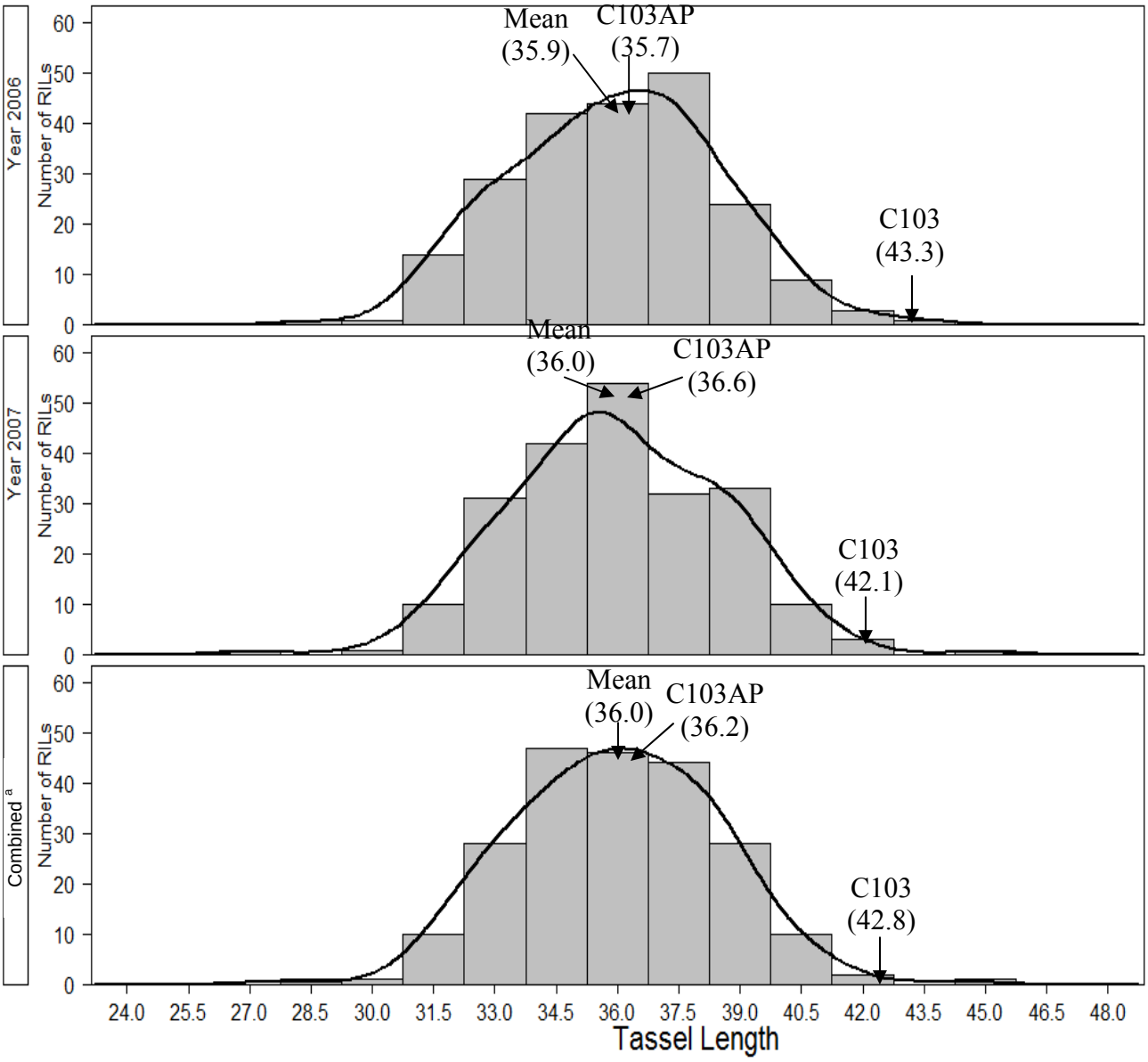
[a] Chromosomal bin corresponding to the location of the nearest left marker according to MaizeGDB database.
[b] Position in cM from the top of the chromosome calculated by MapMaker v.3
[c] LOD score corresponding to the effect of the QTL as calculated by WinQTL cartographer
[d] Direction of the additive effect, AP= C103AP

Figure A1: Histograms showing the phenotypic distribution of the tassel branch number (TBN) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population



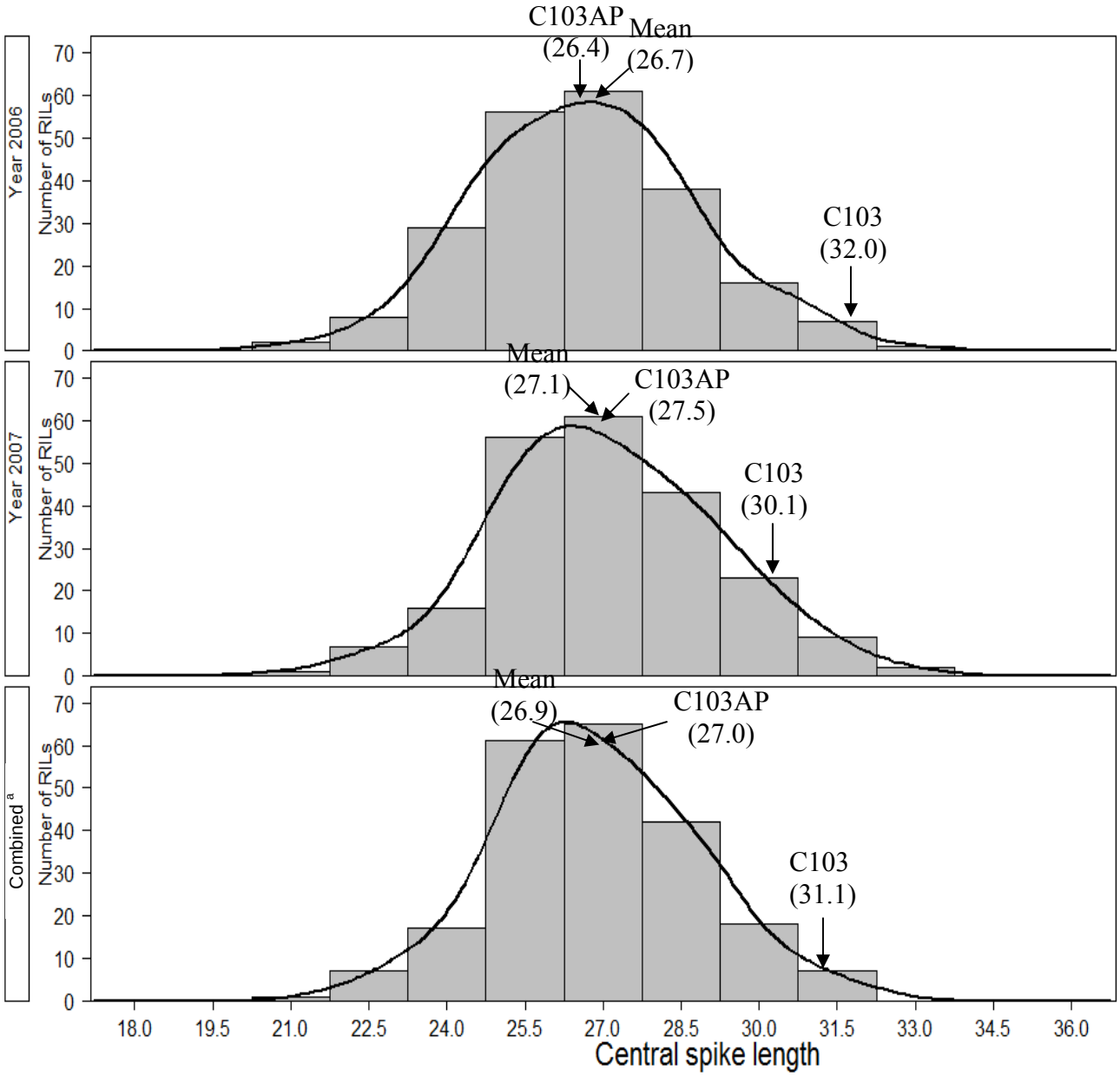
[a] refers to the combined analysis of the 2005, 2006 and 2007 datasets.

Figure A2: Histograms showing the phenotypic distribution of the tassel length (TL) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population



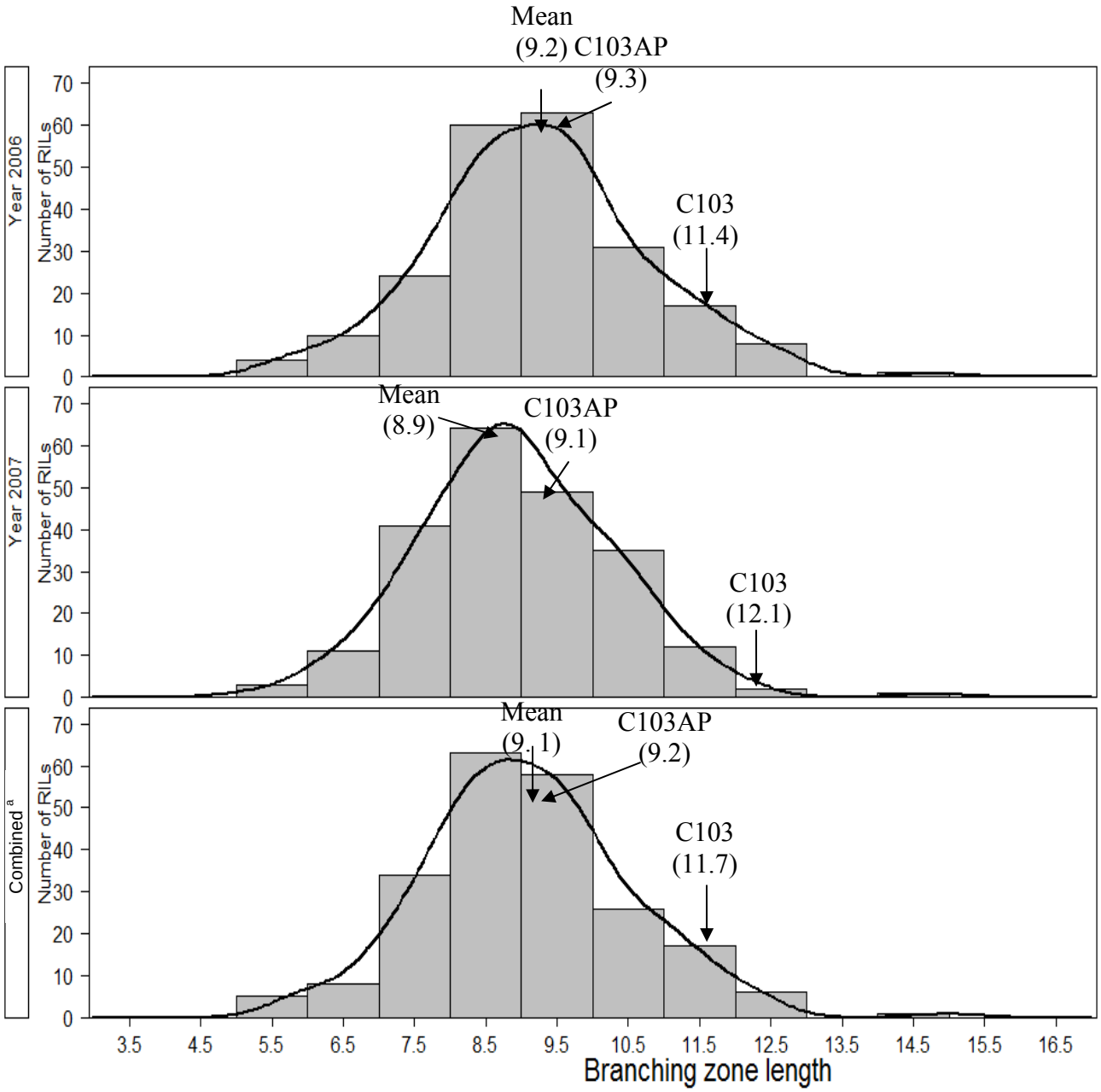
[a] refers to the combined analysis of the 2005, 2006 and 2007 datasets.

Figure A3: Histograms showing the phenotypic distribution of the central spike length (CSL) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population



[a] refers to the combined analysis of the 2005, 2006 and 2007 datasets.

Figure A4: Histograms showing the phenotypic distribution of the branching zone length (BZL) from a population of 218 RILs of maize derived from inbreds C103 and C103AP. Values in parenthesis under C103 and C103AP are the mean of the respective parental inbred line over replications. The value in parenthesis under "Mean" is the mean of the population



[a] refers to the combined analysis of the 2005, 2006 and 2007 datasets.

CHAPTER 4: GENERAL CONCLUSIONS

This study was intended to increase our knowledge and understanding of the genetic factors related to meristem initiation, maintenance and development in maize inflorescence. The investigation was facilitated by the use of a population of recombinant inbred lines derived from the mating of near-isogenic inbred lines, C103 and C103AP.

Six novel regions for prolificacy were detected on chromosomes 3, 6, 8 and 9. On chromosome 8 (bins 8.04-8.05), a region flanked by loci *umc1858* and *umc1309* had a major influence in prolificacy, GDU to anthesis and plant height. In bin 6.06, statistical evidence supported pleiotropy for QTL associated with GDU to anthesis and number of ears per plant. The statistical evidence of pleiotropic gene action concurs with a previous hypothesis that genetic factors in C103AP increased the number of ears per plant, delayed flowering and increased plant height relative to C103 (Duvick 1974).

When tassel architecture was evaluated in the population of RILs, 32 QTL were detected for tassel branch number (TBN), tassel length (TL), central spike (CSL) and branching zone lengths (BZL). The QTL were mainly located on chromosomes 1, 2, 4, 8 and 9. Novel QTL for TBN, TL and CSL were detected in bins 2.01, 2.06 and 2.08 in which QTL have not been described for these traits individually. Similar observations were made for CSL and TL in bin 9.03.

Since the development of the tassel and ears are related processes it can be inferred that genetic factors influencing ear development may also affect some aspects of tassel architecture. Five regions, on chromosomes 2, 3, 6, 8 and 9 (bins 2.08, 3.05, 3.10, 6.01, 8.04

and 9.03) associated with variation in EPP were also detected for tassel traits suggesting common genetic factors for these aspects of ear and tassel development.

Several QTL were detected in the combined analysis but not in each year of evaluation, reflecting a QTL by environment interaction that could eventually contribute to the G x E variance component. Since environmental cues affect the initiation, maintenance and growth of the apical meristem, inconsistent QTL should be studied further since they may contribute to the understanding and the regulation of this developmental pathway.

The statistical tests performed attempted to resolve whether pleiotropy or close linkage was a genetic cause for the correlated response between the traits. Improvements in the mapping resolution, by increasing the number of individuals in the population and more recombinants and molecular analysis, would be needed to distinguish linkage from pleiotropy or other mechanisms. The identification of recombinant progeny for the two associated traits would define that they were in close linkage. Proof that pleiotropy is the cause of the association between two traits can be obtained if DNA sequences and their functions associated with the variation for those traits were identified.

Results presented herein contribute to the understanding of the genetic factors affecting ear development and tassel morphology. However, the approach used only contributes as a first step towards positional cloning of the DNA sequences responsible for the variation in prolificacy and tassel morphology. Fine mapping of these regions (i.e. bins 8.04-8.05) would be necessary in order to define the specific DNA sequence variants underlying QTL and subsequently locate them in the physical map of maize. Ultimately

validation of these candidate sequences can be achieved by complementation tests or other forward genetic approaches.

LITERATURE CITED

Agrama HAS, Moussa ME (1996) Mapping QTLs in breeding for drought tolerance in maize (Zea mays L.). Euphytica 91:89-97

Austin DF, Lee M (1996) Comparative mapping in $F_{2:3}$ and $F_{6:7}$ generations of Quantitative Trait Loci for grain yield and yield components in maize. Theor Appl Genet 92:817-826

Austin DF, Lee M (1998) Detection of Quantitative Trait Loci for grain yield and yield components in maize across generations in stress and nonstress environments. Crop Sci 38:1296-1308

Barazesh S, McSteen P (2008) *barren inflorescence1* functions in organogenesis during vegetative and inflorescence development in maize. Genetics 179:389-401

Bechoux N, Bernier G, Lejeune P (2000) Environmental effects on the early stages of tassel morphogenesis in maize (*Zea mays* L.). Plant, Cell & Environment 23:91-98

Bommert P, Lunde C, Nardmann J, Vollbrecht E, Running M, Jackson D, Hake S, Werr W (2005a) *thick tassel dwarf1* encodes a putative maize ortholog of the Arabidopsis CLAVATA1 leucine-rich repeat receptor-like kinase. Development 132:1235-1245

Bommert P, Satoh-Nagasawa N, Jackson D, Hirano H-Y (2005b) Genetics and evolution of inflorescence and flower development in grasses. Plant Cell Physiol 46:69-78

Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S (2006) *ramosa2* encodes a LATERAL ORGAN BOUNDARY domain protein that determines the fate of stem cells in branch meristems of maize. Plant Cell 18:574-585

Burr B, Burr FA (1991) Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. Trends Genet 7:55-60

Burr B, Burr FA, Thompson KH, Albertson MC, Stuber CW (1988) Gene mapping with Recombinant Inbreds in maize. Genetics 118:519-526

de Leon N, Coors JG (2002) Twenty-four cycles of mass selection for prolificacy in the Golden Glow maize population. Crop Sci 42:325-333

de Leon N, Coors JG, Kaeppler SM (2005) Genetic control of prolificacy and related traits in the Golden Glow maize population: II. Genotypic Analysis. Crop Sci 45:1370-1378

Duncan WG, Williams WA, Loomis RS (1967) Tassels and the productivity of maize. Crop Sci 7:37-39

Duvick DN (1974) Continuous backcrossing to transfer prolificacy to a singled-eared inbred line of maize. Crop Sci 14:69-71

Duvick DN (2005) The contribution of breeding to yield advances in maize (*Zea Mays* L.). Adv Agron 86:83-145

Earley EB, Lyons JC, Inselberg E, Maier RH, Leng ER (1974) Earshoot development of midwest dent corn (*Zea mays* L.) Illinois Exp Stat Bull 747

Eshed Y, Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141:1147-1162

Evans MMS, Barton MK (1997) Genetics of angiosperm shoot apical meristem development. Annu Rev Plant Physiol Plant Mol Biol 48:673-701

Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Enrico Pe M, Schmidt RJ (2004) The role of *barren stalk1* in the architecture of maize. Nature 432:630-635

Graham GI, Wolff DW, Stuber CW (1997) Characterization of a yield Quantitative Trait Locus on chromosome five of maize by fine mapping. Crop Sci 37:1601-1610

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

Jampatong S, Darrah LL, Krause GF, Barry BD (2000) Effect of one- and two-eared selection on stalk strength and other characters in maize. Crop Sci 40:605-611

Lauter N, Doebley J (2002) Genetic Variation for phenotypically invariant traits detected in teosinte: implications for the evolution of novel forms. Genetics 160:333-342

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

Mackay TFC (2004) The genetic architecture of quantitative traits: lessons from Drosophila. Curr Opin Genetics Dev 14:253-257

Malcomber ST, Preston JC, Reinheimer R, Kossuth JE, Kellogg EA (2006) Developmental gene evolution and the origin of grass inflorescence diversity. Adv Bot Res 44:426-481

McSteen P, Hake S (2001) *barren inflorescence2* regulates axillary meristem development in the maize inflorescence. Development 128:2881-2891

McSteen P, Leyser O (2005) Shoot branching. Annu Rev Plant Biol 56:353-374

Mickelson SM, Stuber CS, Senior L, Kaeppler SM (2002) Quantitative Trait Loci controlling leaf and tassel traits in a B73 x Mo17 population of maize. Crop Sci 42:1902-1909

Motto M, Moll RH (1982) Prolificacy in maize: a review. Maydica 28:53-76

Moulia B, Loup C, Chartier M, Allirand JM, Edelin C (1999) Dynamics of architectural development of isolated plants of maize (*Zea mays* L.), in a non-limiting environment: the branching potential of modern maize. Ann Bot 84:645-656

Muehlbauer GJ, Specht JE, Thomas-Compton MA, Staswick PE, Bernard RL (1988) Near-Isogenic Lines a potential resource in the integration of conventional and molecular marker linkage maps. Crop Sci 28:729-735

Orr AR, Sundberg MD (1994) Inflorescence development in a perennial teosinte: *Zea perennis* (Poaceae). Am J Bot 81:598-608

Ritter MK, Padilla CM, Schmidt RJ (2002) The maize mutant *barren stalk1* is defective in axillary meristem development. Am J Bot 89:203-210

Russell WA (1968) Testcrosses of one- and two-ear types of corn belt maize inbreds. i. performance at four plant stand densities. Crop Sci 8:244-247

Russell WA, Eberhart SA (1968) Testcrosses of one- and two-ear types of corn belt maize inbreds. ii. stability of performance in different environments. Crop Sci 8:248-251

Salvi S, Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. Trends Plant Sci 10:297-304

Salvi S, Tuberosa R, Chiapparino E, Maccaferri M, Veillet S, van Beuningen L, Isaac P, Edwards K, Phillips RL (2002) Toward positional cloning of *Vgt1*, a QTL controlling the transition from the vegetative to the reproductive phase in maize. Plant Mol Biol 48:601-613

Sanchez AC, Brar DS, Huang N, Li Z, Khush GS (2000) Sequence Tagged Site Marker-Assisted Selection for three bacterial blight resistance genes in rice. Crop Sci 40:792-797

Sass JE, Loeffel FA (1959) Development of axillary buds in maize in relation to barrenness. Agron J 51:484-486

Schmitz G, Theres K (2005) Shoot and inflorescence branching. Curr Opin Plant Biol 8:506-511

Sorrells ME, Lonnquist JH, Harris RE (1979) Inheritance of prolificacy in maize. Crop Sci 19:301-306

Szalma S, Hostert B, LeDeaux J, Stuber C, Holland J (2007) QTL mapping with Near-Isogenic Lines in maize. Theor Appl Genet 114:1211-1228

Tamas IA (1995) Hormonal regulation of apical dominance. In: Davies PJ (ed) Plant Hormones: Physiology, Biochemistry and Molecular Biology, Second edition edn. Kluwer Academic Publishers.The Netherlands, pp 572-597

Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27:205-233

Upadyayula N, da Silva H, Bohn M, Rocheford T (2006a) Genetic and QTL analysis of maize tassel and ear inflorescence architecture. Theor Appl Genet 112:592-606

Veldboom LR, Lee M (1994) Molecular-marker-facilitated studies of morphological traits in maize. II: Determination of QTLs for grain yield and yield components. Theor Appl Genet 89:451-458

Veldboom LR, Lee M (1996) Genetic mapping of Quantitative Trait Loci in maize in stress and nonstress environments: I. Grain yield and yield components. Crop Sci 36:1310-1319

Vollbrecht E, Springer PS, Goh L, Buckler Iv ES, Martienssen R (2005) Architecture of floral branch systems in maize and related grasses. Nature 436:1119-1126

Wang Y, Li J (2008) Molecular Basis of Plant Architecture. Annu Rev Plant Biol 59

Westerbergh A, Doebley J (2002) Morphological traits defining species differences in wild relatives of maize are controlled by multiple Quantitative Trait Loci. Evolution 56:273-283

Wu X, McSteen P (2007) The role of auxin transport during inflorescence development in maize (*Zea mays*, Poaceae). Am J Bot 94:1745-1755

Xu N, York K, Miller P, Cheikh N (2004) Co-regulation of ear growth and internode elongation in corn. Plant Growth Regul 44:231-241

Zamir D, Eshed D (1998) Case history in germplasm introgression: tomato genetics and breeding using nearly isogenic introgression lines derived from wild species. In: Paterson A (ed) Molecular dissection of complex traits. CrC Press LLC, New York, pp 207-217

APPENDIX 1: GENETIC MAP CONSTRUCTION FOR THE C103/C103AP POPULATION OF RILS

Table A1.1: List of loci that were not polymorphic between C103 and C103AP

Location[a]	Non-polymorphic loci
1.00	bnlg149, umc1106 , umc1353, umc1613
1.01	umc1547, umc1727, umc1948, umc1977, bnlg1112, bnlg1124, umc1566, umc1160, umc1269, umc2012, phi056, phi427913, umc1071, phi097, umc2157, umc1222
1.02	bnlg1007, bnlg109, bnlg1178, bnlg1429, bnlg147, umc1166, umc1070, umc2224
1,03	bnlg1458, bnlg176, bnlg182, bnlg1953, umc1701, umc2097, umc2185, bmc2180, umc1880
1.04	bnlg1016, bnlg1811, bnlg2295, umc1144, umc1254, dupssr26, umc1169, umc1849
1.05	bnlg652, umc1076, umc1124, umc1395, umc1689, umc1703
1.06	bnlg1273, bnlg421, umc2237, umc2238, umc2239
1.07	bnlg1025, bnlg1564, bnlg257, bnlg615, umc1128f, umc1147 , umc1278, umc1661
1.08	dupssr12, PHI002, PHI037, PHI038, PHI039, phi423298 , umc1446, bnlg1629
1.09	bnlg1268 , bnlg400, phi011, phi055, phi094, TBAJR, umc1082, umc1184, umc1306 , umc1512 , umc2047, bnlg100, bnlg1597
1.10	bnlg1347 , bnlg1671, phi265454, phi308707, umc1726 , umc2149 , umc2189
1.11	bnlg1055, bnlg131, bnlg504, phi064, PHI120, umc1064, umc1421, umc1630
1.12	umc1605, umc1739, umc1797, umc1819, umc2100
2.00	phi402893 , umc1622, umc2186
2.01	bnlg1338, umc1227, umc1552, umc2245 , umc2094
2.02	bnlg1297, bnlg469, PHI098, umc1265, umc1823, umc1961, umc1824
2.03	bnlg2248, dupssr27, mmc0231, phi109642, umc1185, umc1845, bnlg1537, umc1555
2.04	bmc1175, bnlg1018, bnlg108, bnlg121, bnlg166, PHI083, umc1007, umc1024, umc1259, umc1326, umc1454, umc2007, umc2030
2.05	bnlg1047, bnlg1063, bnlg180, bnlg1893, bnlg2328, NC131, NC132, bnlg371, dupssr21
2.06	bnlg1887, umc1004, umc1079, umc1108, umc1946, umc2023, umc2178
2.07	bnlg1045, mmc0143, umc1285, umc1560, umc2129 , bmc1335, bnlg1633, umc1890

[a] Chromosomal location according to maize database (www.maizegdb.org; verified 12/06/08).
[b] SSR not polymorphic not mapped: dupssr4, dupssr20, dupssr32, dupssr22
[c] 6 SSR with multiple locations were placed at the lowest chromosomal position defined in MaizeGDB.

Table A1.1 Cont.

Location[a]	Non-polymorphic loci
2.08	bmc1140, bmc1662, bmc1746, bnlg1258, dupssr24, dupssr25, mmc0191, mmc0381, PHI090, PHI127, umc1126, umc1140, umc1992, umc2085, umc1230, umc1798
2.09	bnlg1520, umc1256, umc1736, umc2184, umc1252
2.10	umc2214, umc1696
3.00	umc1746, umc1931f, umc2105 , umc2109 , umc2118 , umc2106
3.01	PHI049, phi104127, umc1394 , umc1892 , umc1970 , umc2049, umc1780
3.02	phi374118, umc1057, umc1458, umc1183
3.03	bnlg1523
3.04	bnlg1019, bnlg602, NC030, PHI029, PHI036, PHI099, umc1223, umc1425, umc1717, umc1835, umc2000, umc2117, bmc1638, bmc2136, umc1012, umc1392, umc1750, umc1965, umc2119
3.05	bnlg420, phi053, phi073, umc1102, umc1307, umc1400, umc2020, bmc1035, bnlg1505, umc1174, umc1539, umc1973, umc2166, umc1593
3.06	bnlg1160, bnlg1449, bnlg1605, bnlg1796, bnlg2241, dupssr17, dupssr23, dupssr5, umc1674, umc1985, bmc1063, umc2169
3.07	umc1148, umc1404, umc1489, umc1528, umc1825, umc2050, umc1767
3.08	bnlg1350, PHI088, umc1273, umc1915, umc1320
3.09	dupssr33, bnlg1536, bnlg1754, mmc0001, umc1639
3.10	umc2048
4.00	phi072, umc2278 , bmc1434f, bnlg1370, bnlg372
4.01	NC135, phi213984, umc1017, umc1022, umc1164 , umc1276, umc1365, umc1669, umc1682, umc1757, umc1759, bnlg1318, umc1228, umc1561
4.02	umc1288, umc1294, umc1509, umc1943
4,03	umc2039, umc2082, umc1902, umc1926
4.04	PHI074, PHI096, umc1652, umc1821, umc1963, umc1382
4.05	bnlg1265, PHI079, umc1031, umc1303, umc1702, umc1791, umc1851, umc1895, umc1896
4.06	bnlg1137 , bnlg252, dupssr16, umc1142
4.07	dupssr34, umc1194, mmc0341
4.08	PHI066, dupssr28
4.09	umc1101, umc1328, umc1573, umc1803, umc1834, umc1940 , bnlg1565, umc1574, umc1631, umc1643, umc1854, umc1999, bnlg292
4.10	bnlg1917 , bnlg589, umc1109, umc1532 , umc1699, umc2044
4.11	bnlg1890 , PHI006, PHI019, PHI076, umc1707 , umc1716, bnlg1337, bnlg2186
5.00	bnlg1006, mmc0151, NC130, umc1097, umc1240, umc1253, umc1260, umc1308, umc1423, umc1496, umc1901

Table A1.1 Cont.

Location[a]	Non-polymorphic loci
5.01	bnlg1382, bnlg143, bnlg1836, NC007, PHI024, umc1781, umc2036, umc1523, umc2179
5.02	bnlg105, bnlg565, dupssr1, phi330507, phi333597, phi396160, umc1587 , umc1761, umc2115 , bmc1660
5.03	bmc1046, bnlg1902, bnlg557, dupssr7, phi008, umc1056, umc1274, umc1447, umc1692, PHI113, umc1557, umc1705, umc1852, umc2060, umc2063, umc2113, umc2159, umc2161
5.04	bnlg1208 , bnlg150, bnlg603, bnlg653, dupssr10, umc1860 , umc1990
5.05	bnlg278, mmc0081 , umc1687, umc1722, umc1800, umc1822 , bnlg1237, umc1019
5.06	bnlg609, PHI087, phi100, PHI101, umc1524, umc1680, umc1941, mmc0481, umc2072
5.07	bnlg118, bnlg1306, PHI058, PHI048, umc1537
5.08	umc1225, umc1792, umc2143, umc1829
5.09	bnlg386, bnlg389, umc1153, umc2209
6.00	bnlg1043, bnlg1600, bnlg161, bnlg238, PHI075, PHI126
6.01	bnlg1422, bnlg1432, bnlg1867, bnlg2191, bnlg249, bnlg426, dupssr18, PHI077, umc1133, umc1753, bmc1433, bmc1538, bnlg391, umc1229, umc1625
6.02	bnlg1371, umc1083, umc1257, umc1628, bmc1628, bnlg2151
6.03	umc1887
6.04	mmc0523 , umc1014 , umc1979, umc2006 , bnlg480, NC009, NC010, PHI031
6.05	bnlg1154, bnlg1732, NC013, PHI078, PHI081, PHI102, PHI129, umc1805f, bnlg1702, NC012, ph445613, PHI025, umc1020, umc1187
6.06	bnlg345, dupssr15, umc1897
6.07	bnlg1740, bnlg1759 , phi070, phi123
6.08	PHI089, umc1127 , umc2059, umc2324
7.00	bnlg1642, bnlg1367, umc1241, umc1546 , umc1695 , umc2177, bnlg2132, mmc0171, bnlg1686, umc1840
7.01	bnlg1292, PHI057, PHI112, umc1066 , umc1409, umc1577, umc2160 , bnlg2160
7.02	bnlg1792, bnlg2203, bnlg398, bnlg657, dupssr11, phi034, PHI114, umc1016, dupssr9
7.03	bnlg339, bnlg434, bnlg572, phi091, umc1015 , umc1936
7.04	bnlg155, bnlg1666, dupssr13, phi328175, umc1029, umc1295, umc1944, umc1412, umc1103, umc1342, umc1710, umc1768, umc1782, umc2133
7.05	phi082, umc1154 , umc1407, umc1760, phi069
7.06	phi045, PHI051, PHI116, bmc1161
8.00	phi420701 , umc1359
8.01	umc1075, umc1139, umc1304, umc1483, umc1786 , umc1817, umc2042, bnlg1194, bnlg1252, umc1327, umc1414 , umc1592
8.02	bnlg2235, PHI119, umc1974, umc2004 , umc1360

Table A1.1 Cont.

Location[a]	Non-polymorphic loci
8.03	bnlg1863, bnlg669, dupssr3, PHI115, PHI125, umc1735, umc1807, umc1904, bmc1067, bnlg2082 , umc1415, umc1778, umc1868, umc1910, umc1984
8.04	bnlg119, PHI121, PHI060, umc1460, umc1765
8.05	bnlg1176, bnlg1599, bnlg162, umc1340, umc1562, umc1651, umc1670, umc1846
8.06	umc1997
8.07	bnlg1828, umc2014
8.08	umc1069, umc1933, PHI080
8.09	umc1638
9.00	bnlg1272 , bnlg1724, umc1279 , umc1647, umc1957
9.01	bnlg1810, PHI028, PHI033, PHI044, PHI067, PHI068, umc1040, umc2078, bnlg1288
9.02	bnlg244, dupssr6, phi017, umc1033, umc1958, dupssr19
9.03	bmc1730, bnlg430, NC133, phi022, PHI061, umc1120, umc1691, umc1267, umc1657, umc1921
9.04	bnlg1714, PHI016, PHI032, PHI042, umc1522, umc1570, umc2121
9.05	bnlg1884, PHI040, phi108411, umc1231, bnlg1129, umc1270, umc1357
9.06	umc1789
9.07	bmc1375f, bnlg128, bnlg279, umc1675 , umc1804 , umc2099 , bnlg1506, bnlg619, umc1505, dupssr29
9.08	umc1137 , umc1982
10.00	dupssr8, PHI041, PHI117, PHI118, umc1293 , umc1380, umc2399
10.01	umc1152, umc1291 , umc1318 , umc1319 , umc2053
10.02	mmc0501, PHI052, phi059, PHI063, umc1337 , umc2018
10.03	bnlg1037 , bnlg210, bnlg640, PHI050, umc1367, umc1863 , phi054, umc1785
10.04	PHI062, phi071, PHI084, umc1077, umc1272, dupssr31
10.05	bnlg1185, bnlg1250 , umc1506, umc1677 , bnlg137
10.06	bnlg153, bnlg236, PHI035, umc1045, umc1993 , umc2122, bnlg1677, bnlg2190, bnlg594, dupssr30
10.07	bnlg1360, bnlg1450, phi323152, umc1084, umc1344

Table A1.2: Polymorphic loci not included in the construction of the C103/C103AP genetic map. Genomic location based on the Maize database, number of alleles observed for C103 and C103AP, chi-square statistic and p-value for testing the 1:1 segregation ratio and genetic map annotation are shown

SSR loci	Bin	Observed			Chi-sq statistic	p-value	Genetic map annotation
		C103	C103AP	Total			
bnlg1014	1.01	2	216	218	210.07	< 0.001	Nearest locus >70 cM
umc1177	1.01	106	112	218	0.17	0.6845	Unlinked
phi095	1.03	109	109	218	0.00	1.0000	Unlinked
umc1403	1.03	2	216	218	210.07	< 0.001	WBA [a]
umc1917	1.04	34	184	218	103.21	< 0.001	WBA
umc2112	1.04	209	9	218	183.49	< 0.001	WBA
umc1035	1.06	130	88	218	8.09	0.0044	Scoring problems
umc1419	2.00-2.01	124	93	217	4.43	0.0353	Assign to LG3 [b]
umc1465	2.04	128	90	218	6.62	0.0101	WBA
umc1637	2.06	1	217	218	214.02	< 0.001	WBA
umc1749	2.06	1	217	218	214.02	< 0.001	WBA
umc1042	2.07	1	217	218	214.02	< 0.001	WBA
bnlg1316	2.08	4	213	217	201.29	< 0.001	Extreme deviation from the 1:1 genotypic ratio
bnlg2144	2.08	1	217	218	214.02	< 0.001	Extreme deviation from the 1:1 genotypic ratio
umc2071	3.01	62	156	218	40.53	< 0.001	WBA
phi193225	3.02	143	75	218	21.21	< 0.001	WBA
bnlg1447	3.03	3	215	218	206.17	< 0.001	WBA
bnlg1647	3.04	2	216	218	210.07	< 0.001	WBA
bnlg1246	3.05/8.05/6.01/5.05	118	100	218	1.49	0.2228	WBA
umc1027	3.06	110	108	218	0.02	0.8923	WBA
bnlg1779	3.07	89	129	218	7.34	0.0067	WBA

[a] WBA= wrong bin assignment within the same linkage group
[b] LG= Linkage Group

Table A1.2 Cont.

SSR loci	Bin	Observed			Chi-sq statistic	p-value	Genetic map annotation
		C103	C103AP	Total			
bnlg1496	3.09	102	116	218	0.90	0.3430	Assigned to LG5
umc1641	3.1	70	148	218	27.91	< 0.001	WBA
umc1117	4.04	121	97	218	2.64	0.1041	WBA
bnlg667	4.05/1.11	127	91	218	5.94	0.0148	Assigned to LG1
NC005	4.05	8	210	218	187.17	< 0.001	WBA
bnlg2291	4.07	3	215	218	206.17	< 0.001	WBA
bnlg1927	4.07	1	217	218	214.02	< 0.001	Extreme deviation from the 1:1 genotypic ratio
umc2009	4.07-4.08	4	214	218	202.29	< 0.001	Extreme deviation from the 1:1 genotypic ratio
umc1775	4.07-4.08	2	216	218	210.07	< 0.001	Extreme deviation from the 1:1 genotypic ratio
mmc0321	4.08	5	213	218	198.46	< 0.001	Extreme deviation from the 1:1 genotypic ratio
umc1132	4.08	1	217	218	214.02	< 0.001	Extreme deviation from the 1:1 genotypic ratio
phi093	4.08	107	111	218	0.07	0.7865	WBA
umc1150	NA	130	88	218	8.09	0.0044	Position unknown
umc1006	6.02	106	112	218	0.17	0.6845	Unassigned
umc1413	6.05	2	216	218	210.07	< 0.001	WBA
bnlg1136	6.07	116	102	218	0.90	0.3430	WBA
phi100175	8.03	216	2	218	210.07	< 0.001	WBA
bnlg240	8.06	56	162	218	51.54	< 0.001	WBA
umc1170	9.02	3	215	218	206.17	< 0.001	WBA
bnlg127	9.03	27	191	218	123.38	< 0.001	WBA
phi027	9.03	2	216	218	210.07	< 0.001	WBA
umc1519	9.04	1	217	218	214.02	< 0.001	WBA
umc1366	9.06	92	126	218	5.30	0.0213	Assigned to LG1

Table A1.3: Genomic location; observed and expected number of C103 and C103AP alleles for a 1:1 segregation ratio, chi-square statistics and its correspondent probability for SSR loci from the genetic map of the C103/C103AP population of RILs with significant departure from 1:1 segregation ratio

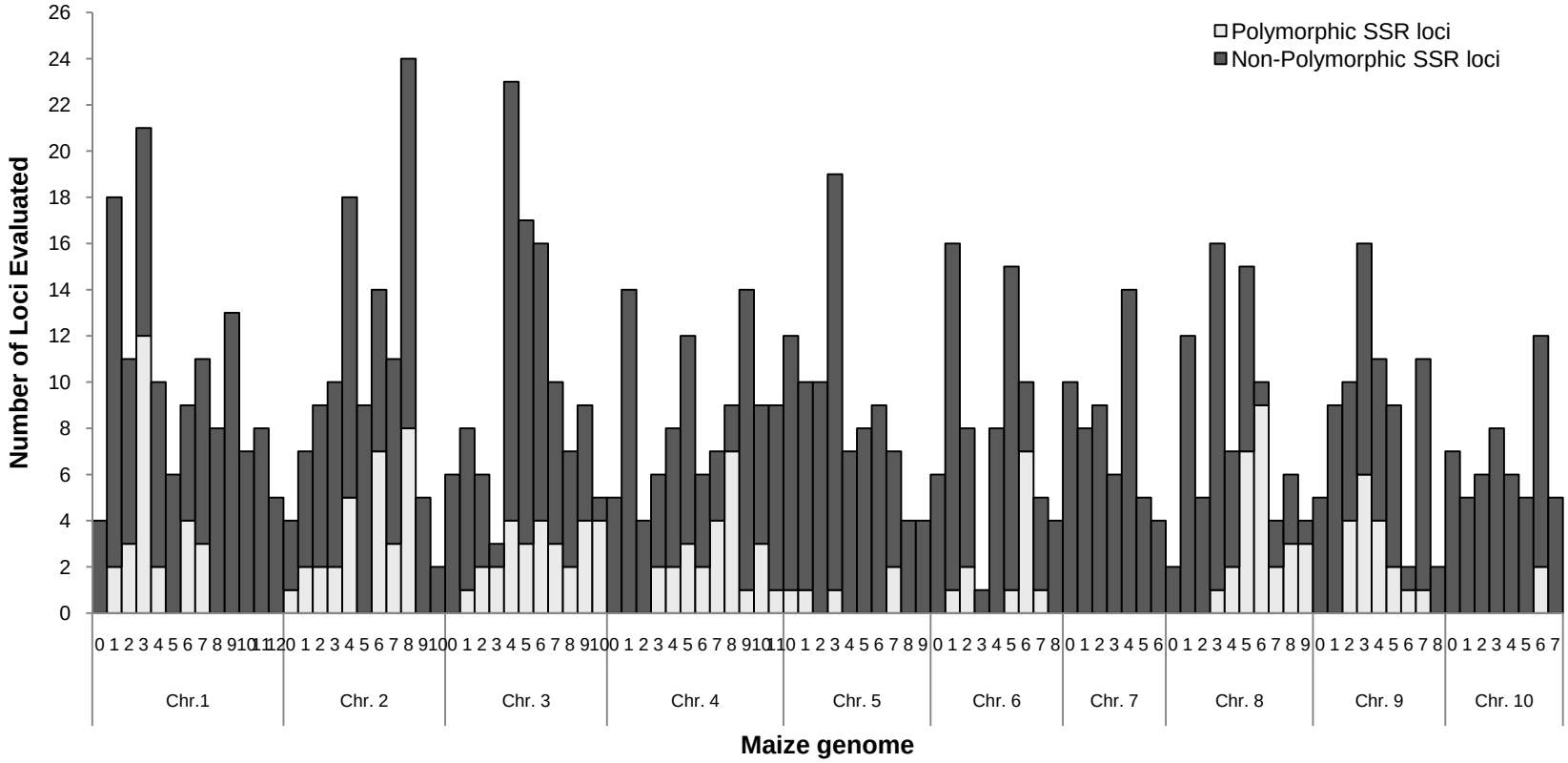
		Observed			Expected (1:1)			
SSR loci	Location	C103	C103AP	Total	C103	C103AP	χ^2 statistic	Probability value
bmc1329	2.06	136	81	217	108.5	108.5	13.94	<0.001
bnlg1036	2.06	125	93	218	109	109	4.70	0.030
umc1875	2.06	130	88	218	109	109	8.09	0.004
umc1536	2.07	149	69	218	109	109	29.36	< 0.001
umc1560	2.07	156	62	218	109	109	40.53	< 0.001
bmc1233	2.08	143	75	218	109	109	21.21	< 0.001
bnlg1940	2.08	126	92	218	109	109	5.30	0.021
bnlg198	2.08	145	73	218	109	109	23.78	< 0.001
bnlg2077	2.08	137	81	218	109	109	14.39	< 0.001
mmc0381	2.08	54	163	217	108.5	108.5	54.75	< 0.001
umc1798	2.08	128	90	218	109	109	6.62	0.010
umc1886	3.02	93	125	218	109	109	4.70	0.030
umc1644	3.06	133	85	218	109	109	10.57	0.001
umc1985	3.06	135	83	218	109	109	12.40	<0.001
umc2269	3.06	90	128	218	109	109	6.62	0.010
umc1135	3.07	72	146	218	109	109	25.12	< 0.001
umc1136	3.1	131	87	218	109	109	8.88	0.003
bnlg490	4.04	90	128	218	109	109	6.62	0.010
umc1511	4.06	93	125	218	109	109	4.70	0.030
bnlg1621	4.06	94	124	218	109	109	4.13	0.042
umc2027	4.06	44	174	218	109	109	77.52	< 0.001
bnlg1444	4.08	13	205	218	109	109	169.10	< 0.001
phi092	4.08	93	125	218	109	109	4.70	0.030

[a]Code used in mapmaker: “A” homozygous for C103; “B” homozygous for C103AP

Table A1.3 Cont.

SSR loci	Location	Observed C103	Observed C103AP	Observed Total	Expected (1:1) C103	Expected (1:1) C103AP	χ^2 statistic	Probability value
umc1325	5.00	92	126	218	109	109	5.30	0.021
bnlg1346	5.07	80	138	218	109	109	15.43	<0.001
umc1424	6.06	86	132	218	109	109	9.71	0.002
umc2170	6.06	92	126	218	109	109	5.30	0.021
phi014	8.04	149	69	218	109	109	29.36	< 0.001
umc1858	8.04	154	64	218	109	109	37.16	< 0.001
bnlg1176	8.05	160	58	218	109	109	47.72	< 0.001
bnlg1599	8.05	157	61	218	109	109	42.28	< 0.001
bnlg1812	8.05	157	61	218	109	109	42.28	< 0.001
bnlg2181	8.05	165	53	218	109	109	57.54	< 0.001
bnlg666	8.05	163	55	218	109	109	53.50	< 0.001
umc1309	8.05	159	57	216	108	108	48.17	< 0.001
umc1889	8.05	158	60	218	109	109	44.06	< 0.001
umc1161	8.06	124	94	218	109	109	4.13	0.042
umc1665	8.06	128	88	216	108	108	7.41	0.007
umc1960	8.06	125	93	218	109	109	4.70	0.030
bnlg1056	8.08	138	80	218	109	109	15.43	< 0.001
umc1037	9.02	90	128	218	109	109	6.62	0.010
umc1893	9.02	91	125	216	108	108	5.35	0.021
umc1688	9.03	93	125	218	109	109	4.70	0.030
umc1571	9.04	81	137	218	109	109	14.39	<0.001
phi236654	9.05	148	69	217	108.5	108.5	28.76	< 0.001
umc1078	9.05	124	94	218	109	109	4.13	0.042
bnlg1028	10.06	90	128	218	109	109	6.62	0.010

Figure A1.1: Total number of polymorphic and non-polymorphic SSR loci evaluated on each chromosome from the screening of C103 and C103AP inbreds**.**



APPENDIX 2: FORMULAE USED IN THE STATISTICAL ANALYSIS

PHENOTYPIC ANALYSIS

General Linear Models

Homogeneous variance models

The mixed linear models used here are an extension of the general linear model by allowing a more flexible specification of the covariance matrix of $\boldsymbol{\varepsilon}$ and can be written as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\varepsilon} \quad [1]$$

where $\boldsymbol{y}$ is the $n \times 1$ response vector; $\boldsymbol{X}$ is the incidence matrix for the fixed effects of dimensions $n \times b$; $\boldsymbol{\beta}$ is the $b \times 1$ fixed effects vector that contains the overall mean ($\boldsymbol{\mu}$); a classification variable used to obtain the parental and the population mean effects; replication within years effect; blocks within replication effect and year effects (when combined across environments, otherwise only blocks within replication effect); $\boldsymbol{Z}$ is the incidence matrix for the random effects of dimensions $n \times r$; $\boldsymbol{\gamma}$ is the $r \times 1$ vector of random effects for the estimation of genotype and the 'genotype × environment' interaction effects (when data was combined across environments); $\boldsymbol{\varepsilon}$ is the $n \times 1$ vector of random errors.

The assumptions of this model are that $\begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim \boldsymbol{N}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{D} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} \end{pmatrix} \right)$; thus, the variance-covariance matrices of γ and ε are given by: $D(\boldsymbol{\gamma}) = var(\boldsymbol{\gamma}) = \boldsymbol{D}$ and $D(\boldsymbol{\varepsilon}) = var(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma} = \boldsymbol{I\sigma^2}$. These assumptions also imply independence between random effects and random errors, hence, the variance of $\boldsymbol{y}$ is $\boldsymbol{V = ZDZ' + \Sigma}$.

Heterogeneous variance models

The model was specified as:

$$\boldsymbol{y_i} = \boldsymbol{X_i\beta} + \boldsymbol{Z_i\gamma} + \boldsymbol{\varepsilon_i} \quad [2]$$

where i corresponds to environments (years of evaluation: 2005-2007 in chapter II and 2006-2007 in chapter III). The difference of this model with [1] is that there is a residual variance for each year of evaluation. The same specifications of [1] are applied here with a modification on the assumptions of the model to account for the heterogeneity of variance, hence $\begin{pmatrix} Y \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} D & & & \\ & I\sigma_1^2 & & \\ & & I\sigma_2^2 & \\ & & & I\sigma_3^2 \end{pmatrix}\right)$. The ddfm=SATTERWHITE option was used in the model statement of model [2] for EPP and GDU to anthesis (Littel et al. 2006, SAS 2008b).

Estimation of covariance parameters

Estimation is done by ***generalized least squares (GLS) by*** minimizing

. However, it requires knowledge of $\mathbf{V}$ and, subsequently, knowledge of $\boldsymbol{D}$ and $\boldsymbol{\Sigma}$. To find a reasonable estimate of $\boldsymbol{D}$ and $\boldsymbol{\Sigma}$, the approach used was ***restricted/residual maximum likelihood*** (REML) method, exploiting the assumption that $\boldsymbol{\gamma}$ and ε are normally distributed.

The MIXED procedure from SAS constructs an objective function associated with REML and maximizes it over all unknown parameters. Maximizing the likelihood over $\boldsymbol{D}$ and $\boldsymbol{\Sigma}$ leads to the following log-likelihood function –

– – —, where ; and $\boldsymbol{\rho}$ is the rank of $\boldsymbol{X}$.

To obtain the REML estimates, this function is minimized -2 times by using a ridge-stabilized Newton-Raphson algorithm (SAS 2008b).

Best linear unbiased estimators (BLUE) and predictors (BLUP)

Estimates of $\boldsymbol{\beta}$ for the fixed effects: year, replication (year), block (rep × year), parents, population mean as well as estimates of $\boldsymbol{\gamma}$ for the genotype and genotype × environment random effects were obtained by solving the mixed model equations:

$$\begin{bmatrix} \boldsymbol{X'\widehat{\Sigma}^{-1}X} & \boldsymbol{X'\widehat{\Sigma}^{-1}Z} \\ \boldsymbol{Z'\widehat{\Sigma}^{-1}X} & \boldsymbol{Z'\widehat{\Sigma}^{-1}Z + \widehat{D}^{-1}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\widehat{\beta}} \\ \boldsymbol{\widehat{\gamma}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'\widehat{\Sigma}^{-1}y} \\ \boldsymbol{Z'\widehat{\Sigma}^{-1}y} \end{bmatrix} \quad [4]$$

Thus, the solution for the *BLUE* and *BLUP* vectors is as follows:

$$\boldsymbol{\widehat{\beta}} = \left(\boldsymbol{X'\widehat{V}^{-1}X}\right)^{-}\boldsymbol{X'\widehat{V}^{-1}y} \quad [5]$$

$$\boldsymbol{\widehat{\gamma}} = \boldsymbol{\widehat{D}Z'\widehat{V}^{-1}}\left(\boldsymbol{y} - \boldsymbol{X\widehat{\beta}}\right) \quad [6]$$

Since $\boldsymbol{D}$ and $\boldsymbol{\Sigma}$ are usually unknown and are estimated by using REML. These estimates are substituted into expression [3] to obtain the approximate variance-covariance matrix of $\left(\boldsymbol{\widehat{\beta}} - \boldsymbol{\beta}, \boldsymbol{\widehat{\gamma}} - \boldsymbol{\gamma}\right)$. In this case, ***empirical estimators and predictors are obtained.*** The approximate variance-covariance matrix obtained tends to underestimate the true sampling variability of $\left(\boldsymbol{\widehat{\beta}}\ \boldsymbol{\widehat{\gamma}}\right)$ because no account is made for the uncertainty in estimating $\boldsymbol{D}$ and $\boldsymbol{\Sigma}$. To account for this under estimation, an inflation factor was used by the KENWARDROGER option.

If $\boldsymbol{\widehat{D}}$ is singular, then the mixed model equations are modified as follows:

$$\begin{bmatrix} \boldsymbol{X'\widehat{\Sigma}^{-1}X} & \boldsymbol{X'\widehat{\Sigma}^{-1}Z\widehat{L}} \\ \boldsymbol{\widehat{L}'Z'\widehat{\Sigma}^{-1}X} & \boldsymbol{\widehat{L}'Z'\widehat{\Sigma}^{-1}Z\widehat{L} + I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\widehat{\beta}} \\ \boldsymbol{\widehat{\tau}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'\widehat{\Sigma}^{-1}y} \\ \boldsymbol{\widehat{L}'Z'\widehat{\Sigma}^{-1}y} \end{bmatrix} \quad [7]$$

where $\boldsymbol{\widehat{L}}$ is the lower-triangular Cholesky root of $\boldsymbol{\widehat{D}}$, satisfying $\boldsymbol{\widehat{D}} = \boldsymbol{\widehat{L}\widehat{L}'}$. Both $\boldsymbol{\widehat{\tau}}$ and a generalized inverse of the left-hand-side coefficient matrix are then transformed by using $\boldsymbol{\widehat{L}}$ to determine $\boldsymbol{\widehat{\gamma}}$.

Least Square Means (LS-means)

LS-mean for the parents and the population of RILs were computed as $\boldsymbol{L}\widehat{\boldsymbol{\beta}}$, where $\boldsymbol{L}$ is the coefficient matrix associated with a given fixed effect estimate and $\widehat{\boldsymbol{\beta}}$ is the BLUE from [4]. The approximate standard error for the LS-mean is computed as the square root of $\boldsymbol{L}\left(\boldsymbol{X}'\widehat{\boldsymbol{V}}^{-1}\boldsymbol{X}\right)^{-}\boldsymbol{L}'$.

Phenotypic and Genetic correlation coefficients and standard errors

Phenotypic and genetic correlation coefficients were calculated per year as well as for the combined analysis as follows:

$$r_{P_1,P_2} = \frac{Cov(P_1,P_2)}{\sqrt{Var(P_1)*Var(P_2)}} \qquad [9]$$

$$r_{G_1,G_2} = \frac{Cov(G_1,G_2)}{\sqrt{Var(G_1)*Var(G_2)}} \qquad [10]$$

where: P_1= Phenotypic value of the ith trait; P_2= Phenotypic value of the jth trait; G_1= Genotypic value of the ith trait and G_2= Genotypic value of the jth trait.

Variance and covariance components and approximate standard errors were obtained using REML following Holland (2006). Given the following linear model for a trait (Y_1):

$$y_1 = \mu_1 + \tau_{1i} + \alpha(\tau)_{1ij} + \beta(\alpha\tau)_{1ijk} + \gamma_{1l} + (\gamma\tau)_{1il} + \varepsilon_{1ijkl} \quad [11]$$

where μ_1 = overall mean for trait *1*; τ_{1i} = effect of the *i* environment (*i*=1, …, 3); $\alpha(\tau)_{1ij}$ =effect of the *j* replications (*j*= 1, 2); $\beta(\alpha\tau)_{1ijk}$ = effect of the *k* blocks within each replication (*k*=1, 2… 17); γ_{1l} = effect of the *l* RILs (*l*=1, 2…, 218); $(\gamma\tau)_{1il}$ = term to account for the genotype by environment interaction and ε_{ijkl} = error term.

The joint model for two traits Y_1 and Y_2 will be:

$$\begin{bmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \end{bmatrix} =$$

$$\begin{bmatrix} \boldsymbol{\mu_1} \\ \boldsymbol{\mu_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{T_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{E_1} \\ \boldsymbol{E_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{A_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{B_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{b_1} \\ \boldsymbol{b_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{G_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{g_1} \\ \boldsymbol{g_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{ge_1} \\ \boldsymbol{ge_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \end{bmatrix}$$

[12]

where $\boldsymbol{y_1}$ and $\boldsymbol{y_2}$ are $n \times 1$ vectors of phenotypic observations of the traits 1 and 2, respectively; $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$ are $n \times 1$ vectors of trait mean effects; $\boldsymbol{E_1}$ and $\boldsymbol{E_2}$ are vectors of macro-environmental effects for the two traits, corresponding to *e* environments; $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ are vectors of replication effects in each of *e* environments; $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$ are vectors of block effects within each replication in each of *e* environments; $\boldsymbol{g_1}$ and $\boldsymbol{g_2}$ are vectors of genotype or family effects, corresponding to *g* genotypes; $\boldsymbol{ge_1}$ and $\boldsymbol{ge_2}$ are vectors of genotype × environment interaction effects; $\boldsymbol{\varepsilon_i}$ and $\boldsymbol{\varepsilon_j}$ are vectors of *n* experimental error effects for traits 1 and 2, respectively; and $\boldsymbol{T_1}$, $\boldsymbol{T_2}$, $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, $\boldsymbol{B_1}$, $\boldsymbol{B_2}$, $\boldsymbol{G_1}$, $\boldsymbol{G_2}$, $\boldsymbol{Z_1}$, and $\boldsymbol{Z_2}$ are incidence matrices for the environment, replication, genotype and genotype x environemnt effects.

Theoretically, all effects should be considered as random effects leading to the following variance and covariance matrices, with the exception of $\boldsymbol{\mu}$:

$$\boldsymbol{V}\begin{bmatrix} \boldsymbol{E_1} \\ \boldsymbol{E_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}\sigma^2_{E1} & \boldsymbol{I}\sigma_{E12} \\ \boldsymbol{I}\sigma_{E12} & \boldsymbol{I}\sigma^2_{E2} \end{bmatrix};\ \boldsymbol{V}\begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}\sigma^2_{r1} & \boldsymbol{I}\sigma_{r12} \\ \boldsymbol{I}\sigma_{r12} & \boldsymbol{I}\sigma^2_{r2} \end{bmatrix};\ \boldsymbol{V}\begin{bmatrix} \boldsymbol{b_1} \\ \boldsymbol{b_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}\sigma^2_{B1} & \boldsymbol{I}\sigma_{B12} \\ \boldsymbol{I}\sigma_{B12} & \boldsymbol{I}\sigma^2_{B2} \end{bmatrix};$$

$$\boldsymbol{V}\begin{bmatrix} \boldsymbol{G_1} \\ \boldsymbol{G_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}\sigma^2_{G1} & \boldsymbol{I}\sigma_{G12} \\ \boldsymbol{I}\sigma_{G12} & \boldsymbol{I}\sigma^2_{G2} \end{bmatrix};\ \boldsymbol{V}\begin{bmatrix} \boldsymbol{ge_1} \\ \boldsymbol{ge_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}\sigma^2_{GE1} & \boldsymbol{I}\sigma_{GE12} \\ \boldsymbol{I}\sigma_{GE12} & \boldsymbol{I}\sigma^2_{GE2} \end{bmatrix}\ and\ \boldsymbol{V}\begin{bmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}\sigma^2_{\varepsilon 1} & \boldsymbol{I}\sigma_{\varepsilon 12} \\ \boldsymbol{I}\sigma_{\varepsilon 12} & \boldsymbol{I}\sigma^2_{\varepsilon 2} \end{bmatrix}$$

However, to reduce computation time, one may take replications, environments and blocks as fixed effects leaving genotype and genotype by environment interaction as random effects. This will yield identical results for balanced data, although with unbalanced data, this analysis will not make use of intertrial information. Since the intertrial information is often

small, sacrifice in efficiency will usually be marginal (Holland 2006; Piepho and Mohring 2005).

Approximate standard error estimates for these coefficients were calculated using the delta method. Below is defined the corresponding variance matrix for the estimation of the standard error of the genetic correlation coefficient:

$$\hat{V}_{(\hat{r}_g)} \approx \begin{bmatrix} \frac{\partial \hat{r}_G}{\partial \hat{\sigma}^2_{Gi}} \\ \frac{\partial \hat{r}_G}{\partial \hat{\sigma}_{Gij}} \\ \frac{\partial \hat{r}_G}{\partial \hat{\sigma}^2_{Gj}} \end{bmatrix}' \begin{bmatrix} V(\hat{\sigma}^2_{Gi}) & C\left(\hat{\sigma}^2_{Gi}, \hat{\sigma}_{Gij}\right) & C(\hat{\sigma}^2_{Gi}, \hat{\sigma}^2_{Gj}) \\ C\left(\hat{\sigma}^2_{Gi}, \hat{\sigma}_{Gij}\right) & V\left(\hat{\sigma}_{Gij}\right) & C\left(\hat{\sigma}_{Gij}, \hat{\sigma}^2_{Gj}\right) \\ C(\hat{\sigma}^2_{Gi}, \hat{\sigma}^2_{Gj}) & C\left(\hat{\sigma}_{Gij}, \hat{\sigma}^2_{Gj}\right) & V(\hat{\sigma}^2_{Gj}) \end{bmatrix} \begin{bmatrix} \frac{\partial \hat{r}_G}{\partial \hat{\sigma}^2_{Gi}} \\ \frac{\partial \hat{r}_G}{\partial \hat{\sigma}_{Gij}} \\ \frac{\partial \hat{r}_G}{\partial \hat{\sigma}^2_{Gj}} \end{bmatrix}$$

$$= \left(\hat{r}_g\right)^2 \begin{bmatrix} \frac{-1}{2\hat{\sigma}^2_{Gi}} \\ \frac{1}{\hat{\sigma}_{Gij}} \\ \frac{-1}{2\hat{\sigma}^2_{Gj}} \end{bmatrix}' \begin{bmatrix} V(\hat{\sigma}^2_{Gi}) & C\left(\hat{\sigma}^2_{Gi}, \hat{\sigma}_{Gij}\right) & C(\hat{\sigma}^2_{Gi}, \hat{\sigma}^2_{Gj}) \\ C\left(\hat{\sigma}^2_{Gi}, \hat{\sigma}_{Gij}\right) & V\left(\hat{\sigma}_{Gij}\right) & C\left(\hat{\sigma}_{Gij}, \hat{\sigma}^2_{Gj}\right) \\ C(\hat{\sigma}^2_{Gi}, \hat{\sigma}^2_{Gj}) & C\left(\hat{\sigma}_{Gij}, \hat{\sigma}^2_{Gj}\right) & V(\hat{\sigma}^2_{Gj}) \end{bmatrix} \begin{bmatrix} \frac{-1}{2\hat{\sigma}^2_{Gi}} \\ \frac{1}{\hat{\sigma}_{Gij}} \\ \frac{-1}{2\hat{\sigma}^2_{Gj}} \end{bmatrix} \quad [13]$$

The corresponding SAS codes used to obtain these estimates were adapted from Holland (2006).

Spearman rank-order correlation

Spearman rank-order correlation is a nonparametric measure of association based on the ranks of the data values. The formula is:

$$\Theta = \frac{\sum_i \left((R_i - \bar{R})(S_i - \bar{S})\right)}{\sqrt{\sum_i (R_i - \bar{R})^2 \sum (S_i - \bar{S})^2}} \quad [14]$$

where R_i is the rank of x_i, S_i is the rank of y_i, $\bar{R}$ is the mean of the R_i values, and $\bar{S}$ is the mean of the S_i values. The CORR procedure computes the Spearman correlation by ranking the data and using the ranks in the Pearson product-moment correlation formula. In case of ties, the averaged ranks are used (SAS 2008a).

The probability values for the Spearman correlation are computed by treating $t = (n-2)^{1/2}\left(\frac{r^2}{1-r^3}\right)^{1/2}$ as coming from a $\boldsymbol{t}$ distribution with *(n-2)* degrees of freedom, where *r* is the sample Spearman correlation. The Fisher's z transformation was used to derive its confidence limits and a p-value under the null hypothesis $H_0: \rho = 0$.

QTL ANALYSIS

Single-Marker analysis

Single Marker analysis is based on comparisons between marker genotypic means through a t-test, an analysis of variance, a likelihood ratio test or a simple regression for trait on coded marker genotype. One marker is analyzed at a time and the QTL is determined to be located near a marker if the phenotypic values for the trait are significantly different among the marker genotypes (Liu 1998). The overall goal of this analysis is simple detection of a QTL linked to a marker, rather than estimation of its position and effects (Lynch and Walsh 1998).

The approach used to test association between the traits and SSR loci *mmc0381, umc1325*, *umc1766*, *bnlg1346*, *phi085*, *bnlg1028* and *umc1477* was a simple regression of the trait values on a dummy variable for the marker genotypes. The regression model is:

$$\boldsymbol{y_i = \beta_0 + \beta_1 x_i + \varepsilon_i} \quad [15]$$

where y_i is the phenotypic trait value of the ith individual of the population; x_i is the dummy variable with value of 1 if the individual is AA or -1 if the individual is aa; β_0 is the intercept for the regression which is the overall mean of the trait, β_1 is the slope for the regression line and ε_i is the error term. The hypothesis test $H_0: \beta_1 = 0$ is equivalent to

testing if the putative QTL and the locus being analyzed are unlinked or that the genetic effects are equal to zero (Liu 1998).

Composite Interval Mapping (CIM)

CIM adds background loci as cofactors in the analysis of a segment between two loci in order to control other QTL effects as well as to reduce the residual variance (Lynch and Walsh, 1998).

The statistical model for CIM analysis on a segment between markers *i* and *i+1* based on a backcross population is described by Zeng (1994) as follows:

$$\boldsymbol{y_j = b_0 + b^* x_j^* + \sum_{k \neq i, i+1} b_k x_{kj} + \varepsilon_j,\ j = 1, 2, \ldots, n} \quad [16]$$

where $\boldsymbol{y_i}$ is the trait value of the jth individual, $\boldsymbol{b_0}$ is the intercept of the model, $\boldsymbol{b^*}$ is the genetic effect of the putative QTL expressed as the difference in effects between the homozygote and heterozygote classes located in the marker interval (*i*, *i+1*), x_j^* is an indicator variable, taking a value 1 or 0 with probability depending on the genotypes of markers *i* and *j* and the position being tested for the putative QTL (double recombination within the marker interval is ignored) , $\boldsymbol{b_k}$ is the partial regression coefficient of the phenotype $\boldsymbol{y}$ on the kth marker, x_{kj} is a known coefficient for the kth marker in the jth individual, taking a value ***1*** or **0** depending on whether the marker type is homozygote or heterozygote, and ε_j is the error term. Under the assumption that ε_j are independent with a N $(0, \sigma^2)$ distribution, the likelihood function is:

$$\boldsymbol{L_1 = \prod_{j=1}^{n} \left[p_j(1) f_j(1) + p_j(0) f_j(0) \right]} \quad [17]$$

where $p_j(1)$ gives a prior probability of x_j^*=1; $p_j(0)$= 1-$p_j(1)$; $f_j(1)$ and $f_j(0)$ are normal density functions for the random variable y_j with $\mu_1 = b_0 + b^* + \sum_{k\neq i,i+1} b_k\, x_{jk}$ and $\mu_2 = b_0 + \sum_{k\neq i,i+1} b_k\, x_{jk}$ for x_j^*=1 and x_j^*=0, respectively; and variance σ^2.

The maximum likelihood (ML) estimates of b^*, $b_{k's}$ and σ^2 are: $\hat{b}^* = \left(Y - X\hat{B}\right)'\hat{P}/\hat{c}$; $\hat{B} = (X'X)^{-1}X'(Y - X\hat{P}b^*)$; $\hat{\sigma}^2 = \left[\left(Y - X\hat{B}\right)'\left(Y - X\hat{B}\right) - \hat{c}\hat{b}^{*2}\right]/n$.

where Y is a $nx1$ vector of y_j's; $\hat{B}$ is a $(t\text{-}1) \times 1$ vector of the estimates of $b_{k's}$; X is an $n\times (t\text{-}1)$ matrix of x_{kj}'s, $\hat{P}$ is a $n\times 1$ vector with the estimates of the posterior probabilities of x_j^*=1 and $\hat{c} = \sum_{j=1}^{n} \hat{P}_{j'}$. Estimates are obtained by the Expectation/Conditional Maximization algorithm (ECM). Iteration of the ECM algorithm consists of one E-step: to compute the probability of $x_j^* = 1$ and three CM steps to estimate the QTL effect, the partial regression coefficients and the variance for the next round of iterations.

The likelihood under the null hypothesis (H_0: $b^* = 0$) is: $L_0 = \prod_{j=1}^{n} f_j(0)$ and the LR statistics is:

$$\boldsymbol{LR = -2\ln(L_0/L_1) = n(ln\,\hat{\hat{\sigma}}^2 - \ln\hat{\hat{\sigma}}^2 + \sum_{j=1}^{m}\left[2\,ln\left[(1-p)\exp\left(\frac{\hat{d}_j}{2}\right) + p\right] - \hat{P}_j\hat{d}_j\right] + \sum_{j=1}^{m}\left[2\,ln\left[p\exp\left(\frac{\hat{d}_j}{2}\right) + (1-p)\right] - \hat{P}_j\hat{d}_j\right]} \quad [18]$$

where $\hat{d}_j = \left[2\hat{b}^*(y_j - \hat{b}_0 - \sum_{k\neq i,i+1} \hat{b}_k\hat{x}_{jk}) - \hat{b}^{*2}\right]/\hat{\sigma}^2$.

The hypotheses are usually tested at every position of an interval and for all intervals of the genome to produce a continuous LRT statistics profile. At every position, the position parameter (p) can be predetermined and only the QTL effect, partial regression coefficient and variance (b^*, B and σ^2, respectively) are involved in the estimation and testing. If the tests are significant in a chromosome region, the position with the largest LRT statistics is

inferred to be the estimate of the QTL position, and the MLEs at this position are the estimates of b^*, B and σ^2 (Kao and Zeng 1997). LR can be expressed in LOD score by a factor of 4.61 ($LR/4.61$).

For the C103/C103AP population, QTL detection by CIM heterozygote genotypes were not considered since it is a population of RILs. Hence, there will be two homozygous classes (i.e. AA and aa) at each marker locus.

Permutation tests

A permutation test is a hypothesis test using the empirical distribution of test statistics obtained by permuting the original sample. The permuted sample is considered to be a sample from the population under the null hypothesis.

In this study, permutation tests were used to calculate a suitable significance level that will give an overall type I error of 0.05 by the procedure described by Churchill and Doerge (1994) as follows: a) individuals in the experiment are indexed from *1* to ***n.*** The data are shuffled by computing a random permutation of the indices *1, . . ., n* and assigning the ith trait value to the individual whose index is given by the ith element of the permutation; b) the shuffled data are then analyzed for QTL effects by CIM. The resulting test statistics at each analysis point are stored and the entire procedure (shuffling and analysis) was repeated 1000 times. At the end of this process there would be results of QTL analyses on 1000 shuffled data sets.

The 95% point of the distribution of the largest likelihood ratio from each reordering gives an overall 0.05 significance level test. The original marker data are used for each permutation; therefore features of the original data set are retained such as the segregation ratios, the number of missing marker genotypes and the pattern of linkage (Hackett 2002).

Multiple Interval Mapping (MIM)

To estimate genetic architecture parameters, including the number, genomic positions, effects and interactions of significant QTL and their contribution to the genetic variance, Multiple Interval mapping analysis (MIM) was conducted. MIM uses multiple marker intervals simultaneously to fit multiple putative QTL directly in the model for mapping QTL. Besides being more powerful and precise in detecting QTL when compared to Interval Mapping and CIM, MIM can readily search for and analyze epistatic interactions, and estimate the individual genotypic values and heritabilities of quantitative traits (Kao et al. 1999).

The MIM model for the analysis of the population of RILs was adapted from Kao et al. (1999) from a backcross population. Cockerham's genetic model is used to define the genetic parameters and to model the relation between the genotypic value and the genetic parameter. Considering m QTL for two genotypic classes (Q_jQ_j and q_jq_j) , each with ½ frequency for a QTL, there are 2^m possible genotypes, and considering only digenic epistatic interactions, the genetic model for the individual i is:

$$\boldsymbol{G_i = \mu + \sum_{j=1}^{m} a_j\, x_{ij} + \sum_{j<k}^{m} w_{jk}\,(x_{ij}, x_{ik})\,,\ \ i = 1, \ldots . 2^m} \quad [19]$$

where x_{ij} is a coded variable with values of ½ or – ½ if the genotype of a given QTL (Q_j) is Q_jQ_j or q_jq_j, respectively; a_j is the additive main effect of Q_j, and w_{ij} is the epistatic effect between Q_j and Q_k.

For the statistical model of MIM, we can assume m QTL (QTL_1, …, QTL_m) located at positions p_1,…, p_m in m different marker intervals, I_1, I_2, …, I_m along the genome that may show epistasis affecting a quantitative trait y on an individual i by the following relationship:

$$\boldsymbol{y_i} = \boldsymbol{\mu} + \sum_{j=1}^{m} \boldsymbol{a_j}\, \boldsymbol{x_{ij}^*} + \sum_{j<k}^{m} \boldsymbol{\delta_{jk}}\, (\boldsymbol{w_{jk}} \boldsymbol{x_{ij}^*} \boldsymbol{x_{ik}^*}_{\boldsymbol{ij}}) + \boldsymbol{\varepsilon_i} \quad [20]$$

where μ is the overall mean, x_{ij}^* is the coded variable for the genotype of Q_j; a_j and w_{ij} are the additive and epistatic effects as defined in [19], δ_{jk} is an indicator variable for epistasis between Q_j and Q_k that will be =1 if both QTL interact, otherwise will be 0; and ε_i is the error term (errors are assumed to be independent with a N (0, σ^2) distribution). Since the putative QTL genotypes (x_{ij}^*) are not observed because the QTL could not be located in the marker intervals and given observed flanking marker genotypes, the conditional distribution of x_{ij}^* for QTL at specific positions, p_j*'s*, can be inferred based on Haldane's mapping function and then the MIM model is a normal mixture model. For each *Qj*, its conditional probabilities are extracted to form a conditional probability matrix used to estimate the QTL position in the intervals.

The distribution of each Q_j in interval I_j is not observed but can be inferred from the flanking markers of I_j based on the recombination frequency between them. Assuming no crossing-over interference, the conditional distributions of the individual putative QTL genotypes, given the flanking marker genotypes are independent, thus, the joint conditional probability of the m QTL is the product of marginal conditional probabilities of individual QTL.

The likelihood function of the MIM model for $\theta = (p_1, p_2, \ldots, p_m, a_1, a_2, \ldots, a_m, \ldots, w_{jk}, \ldots\sigma_2)$ is: $L(\theta|Y,X) = \prod_{i=1}^{n} \left[\sum_{j=1}^{2^m} p_{ij}\, \Phi\left(\frac{y_i-\mu_{ij}}{\sigma}\right)\right]$,where $\Phi(\cdot)$ is a standard normal density function, μ_{ij} 's correspond to the genotypic values of the 2^m different QTL genotypes from [20] and p_{ij}*'s* containing information on QTL positions are the corresponding joint conditional probabilities.

When arbitrarily m putative QTL are considered, the likelihood is a mixture of 2^m normal distributions and at least $2m$ +2 parameters need to be estimated, hence, one of the difficulties with this model is the difficulty in the derivation of the MLE as m increases. MLEs and the asymptotic variance-covariance matrix of the model are obtained based on the formulae derived by Kao and Zeng (1997) based on two matrices, $\boldsymbol{D}$ and $\boldsymbol{Q}$. $\boldsymbol{D}$ is genetic design matrix that characterizes the genetic parameters of the QTL effects with dimensions $2^m \times k$ (k is the column dimension that will contain m main effects and $m\ (m-1)/2$ epistatic interactions). $\boldsymbol{Q}$ is the conditional probability matrix that contains the information on QTL positions with dimensions $4^m \times 2^m$ (at each interval there are 4 possible marker classes: i.e. AABB, AAbb, aaBB and aabb; thus extending to m intervals, there will be 4^m marker classes). At the tested positions, the mixing proportions of the p_{ij}*'s* in the likelihood are fixed and do not need to be estimated. Then, MLES are obtained for the mean, variance, marginal and epistatic effects.

Assessment of pleiotropy and linkage of QTL

This analysis was conducted on those traits with overlapping 1-LOD supporting interval by Multiple trait- MIM using the rational presented by Jiang and Zeng (1995) for multiple trait analysis. Let two QTL, each with an effect on one trait, have positions $\boldsymbol{p_1}$ for the QTL having an effect on trait **1** and $\boldsymbol{p_2}$ for the QTL having an effect on trait **2**. The hypotheses can then be formulated as H_0: $\boldsymbol{p_1} = \boldsymbol{p_2}$ and H_1: $\boldsymbol{p_1} \neq \boldsymbol{p_2}$ **.**

The statistical models for H_1 are the same as [16] for trait 1 and trait 2. The likelihood function is given by

$$\boldsymbol{L_2} = \prod_{j=1}^{n} \sum_{i=0}^{2} \sum_{i=0}^{2} \boldsymbol{p}_{i1i2j} \boldsymbol{f}_{i1i2j}\left(\boldsymbol{y}_j\right) \quad [21]$$

where $f_{i1i2j}(y_j)$ is a bivariate normal density function for y_j with mean $u'_{i1i2j} = \begin{pmatrix} x_j b_1 + i_1 b_1^* \\ x_j b_2 + i_2 b_2^*) \end{pmatrix}$ and variance $V = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$, where $\delta(i_1)$ is an indicator function that will be =1 if i_1=1 and 0 otherwise. The probability p_{i1i2j} can be inferred from the observed genotypic classes. The ECM algorithm is used to calculate the posterior probabilities of individual *j* having genotype i_1 for QTL 1 affecting trait 1 at position p_1 and i_2 for QTL 2 affecting trait 2 at position p_2 (E-step) and calculate the QTL effects, the partial regression coefficients and the variance-covariance matrix for both traits for the next round of iterations (CM-step).

The test statistics is given by the likelihood ratio $LR=-2ln(L_0/L_1)$. It is expected that the likelihood under the alternative hypothesis (L_1) is maximized in the region near the peak indicated by the separate mappings, whereas the maximum likelihood under the null hypothesis (L_0) corresponds to the joint mapping under the same model. As the hypotheses described above are nested hypothesis, the test statistic under H_0 will be asymptotically chi-square distributed with 1 degree of freedom. A further detail on the log likelihood functions for the models is described by Jiang and Zeng (1995).

REFERENCES

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971

Hackett CA (2002) Statistical methods for QTL mapping in cereals. Plant Mol Biol 48:585-599

Holland JB (2006) Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. Crop Sci 46:642-654

Jiang C, Zeng ZB (1995) Multiple Trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111-1127

Kao C-H, Zeng Z-B (1997) General formulas for obtaining the MLEs and the asymptotic variance- covariance matrix in mapping quantitative trait loci when using the EM algorithm. Biometrics 53:653-665

Kao C-H, Zeng Z-B, Teasdale RD (1999) Multiple Interval Mapping for Quantitative Trait Loci. Genetics 152:1203-1216

Littel RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O (2006) SAS for mixed models, second edition. SAS Institute Inc. , Cary, NC

Liu BH (1998) Statistical Genomics: linkage mapping and QTL analysis, First edn. CRC Press LLC, Boca Raton, New York

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

Piepho HP, Mohring J (2005) Best linear unbiased prediction of cultivar effects for subdivided target regions. Crop Sci 45:1151-1159

SAS (2008a) The CORR procedure. SAS/STAT 9.2 User's Guide. SAS Institute Inc., Cary, NC, pp 3-62

SAS (2008b) The MIXED Procedure. SAS/STAT 9.2 User's Guide. SAS Institute Inc., Cary, NC., pp 2559-2851

Zeng ZB (1994) Precision mapping of Quantitative Trait Loci. Genetics 136:1457-1468.