Genetic analysis of hybrid value for silage maize in multiparental designs : QTL detection and genomic selection

Héloïse Giraud

NNT : 2016SACLS013

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE PARIS-SUD

ECOLE DOCTORALE N°581
Agriculture, Alimentation, Biologie, Environnement et Santé

Spécialité de doctorat : Sciences du vivant

Par

Mme Héloïse Giraud

Genetic analysis of hybrid value for silage maize in multiparental designs: QTL detection and genomic selection

Thèse présentée et soutenue à l'UMR Génétique Quantitative Evolution – Le Moulon, le 22 janvier 2016 :

Composition du Jury :
M. Rex BERNARDO, Professeur, University of Minnesota, Rapporteur
Mme Pascale LEROY, Directrice de Recherches, INRA, UMR PEGASE, Rapporteur
M. Philippe BRABANT, Professeur, AgroParisTech, UMR GQE – Le Moulon, Président du jury
M. Pierre CAROLO, Research and Breeding of Maize Manager, Euralis Semences, Examinateur
Mme Zulma VITEZICA, Maître de Conférences, ENSAT, UMR GenPhySE, Examinatrice
M. Alain CHARCOSSET, Directeur de Recherches, INRA, UMR GQE – Le Moulon, Directeur de thèse
Mme Laurence MOREAU, Chargée de Recherches, INRA, UMR GQE – Le Moulon, Co-directeur de thèse

Remerciements

Je souhaite tout d'abord remercier Alain Charcosset et Laurence Moreau pour m'avoir permis de passer ces plus de trois ans et demi au Moulon, d'abord en stage puis en thèse. Merci de m'avoir accordé votre confiance. Je souhaite remercier plus particulièrement Laurence pour son encadrement de tous les jours, le temps passé derrière l'ordinateur et d'avoir toujours été là pour répondre à mes questions. J'ai beaucoup apprécié travailler avec toi et beaucoup appris. Merci de m'avoir aussi laissé liberté de m'organiser et d'avoir su que malgré tous les problèmes de transport et autres affres de déplacement, je réussirais à revenir au Moulon ! Vous m'avez tous les deux fait découvrir un monde stimulant que j'apprécie beaucoup, encore merci.

Je souhaite également remercier les autres membres de l'équipe et particulièrement ceux qui ont œuvré au projet SAM-MCR avant et pendant mon séjour au Moulon ! Merci à Cyril Bauland pour avoir coordonné tout le projet, être allé visiter les essais et m'avoir initiée à la sélection. Merci à Valérie Combes et Delphine Madur pour avoir coordonné les prélèvements chez l'ensemble des partenaires, pour toutes les extractions et le suivi du génotypage réalisé chez Limagrain. Merci à Philippe Jamin et Cécile Monteil pour les comptages de graines, les ajustements de dernière minute et la gestion des envois. Je voulais également remercier Adama Seye pour avoir analysé une partie des données pour laquelle je manquais de temps à consacrer. Bon courage pour la thèse sur SAM-MCR – Le retour. Une mention spéciale à Valérie pour ses cannelés et m'avoir montré comment choisir les meilleurs. Merci Cyril pour les dégustations avec le club œnologique, j'ai beaucoup apprécié même si je ne mettais pas toujours des notes très élevées. Je vais continuer à entraîner mon nez, qui d'après Alain est déjà bon, mais il y a encore du progrès à faire ! Merci également aux autres membres de GQMS pour les échanges que l'on a pu avoir et particulièrement André Gallais et Tristan Mary-Huart (merci pour l'écriture des modèles et toutes les explications de stats !). Merci Stéphane Nicolas pour m'avoir encadré dans mes débuts au Moulon, pour mon stage, et j'espère que bientôt on clôturera cette période. Merci à Matthieu Falque pour tes explications et ton aide sur la cartographie. Un grand merci également aux gens du Moulon en général pour la bonne ambiance qui y règne et les discussions que l'on a pu avoir.

Je souhaite remercier tous les partenaires privés du projet SAM-MCR pour avoir permis sa mise en œuvre et pour les discussions que l'on a pu avoir. Je voudrais également remercier Chris Schön et Christina Lehermeier pour le séjour à la TUM et la collaboration sur les données du NAM.

Je remercie Rex Bernardo, Pascale Leroy, Philippe Brabant, Zulma Vitezica, Pierre Carolo

d’avoir accepté de faire partie de mon jury de thèse. Je souhaite également remercier les membres de mon comité de thèse pour les conseils et ceux de X-Gen pour m’avoir donné un aperçu de la génétique animale. Merci plus spécialement à Andrés Legarra et Zulma Vitezica pour la découverte du monde de génétique quantitative animale hispanique.

Et que n’aurait été ces années au Moulon, sans vous tous, la bande de jeunes. Il y en a qui sont encore là, d’autres qui sont partis, mais en tout cas merci à tous ! Merci à Nico, Fabio, Natalia, Yannick. Merci à Pierrot (le giffois-palmyrien propriétaire de chalet parisien) et Bub (l’incroyable Bub des neiges, pourfendeur des fonds de frigo) qui m’ont adopté, à l’époque où j’étais la jeunette du labo. Qui aurait cru que la longueur de la barbe était si importante ! Merci à mes colocs de bureau, Pierre et Charlotte ; Renaud et Mariangela ; Sandra et Fabien. Merci Renaud pour toutes les discussions que l’on a pu avoir scientifiques ou non, merci Mariangela pour ta bonne humeur et la découverte du risotto aux fraises. Bon courage Fabien pour la suite de ta thèse. Tu fais partie de ceux qui ont un esprit comment dire... mal placé ? Sur ce point, spéciale mention également à l’incredible JT, Chris et Cyril le psy. Merci à Zeineb pour nous avoir donné l’occasion à Margaux-Alison et à moi de découvrir la Tunisie, merci pour l’accueil chaleureux dans ta famille, c’était vraiment superbe. Merci Charlotte entre autres pour notre petite escapade barcelonaise. Merci Steph pour les soirées jeux et Paulina pour m’avoir fait découvrir les spécificités du mexicain. Merci à toute la bande du ski, j’ai vraiment apprécié cette semaine et si, je vous le promets, j’allais vite ! Merci spécialement à Margaux-Alison (vous ne seriez pas sœurs par hasard ? – tu seras au Mexique pour ma soutenance mais je viendrais pour la tienne, pas de rancune) et Chris (mon taxi des derniers mois, fan de durian), qui ont commencé à peu près en même temps que moi, pour le soutien et les vacheries mutuelles ces trois dernières années. Qui aime bien châtie bien comme on dit. Caliméro a vraiment apprécié votre compagnie ces dernières années. Merci à tout le monde pour les moments partagés, les soirées et rires entre le Moulon, le QG d’Orsay, celui des artistes (ou des anglais ?), les logements des uns et des autres... Merci.

Merci aux amis non mouloniens qui m’ont permis de vivre plein de moments forts à Paris, aux quatre coins de France ou même à l’étranger (tous les transports que cela soit train ou avion, passent par Paris, et on ne peut pas dire que je n’en ai pas profité) : Elsa, Laure, Marie, Emeline, Romain, Mathieu, Siva, Maren, Merivan, Louis, Firas, Pieter, Anna, Jesús.

Je remercie tout particulièrement ma famille pour son soutien toutes ces années et ce depuis bien avant la thèse. Merci de m’avoir permis de suivre ma voie, de m’avoir laissé aller un peu partout toutes ces années (désolée pour les peurs quand les évènements imprévus s’en mêlaient) et d’avoir été là pour moi. Merci à mes grands-parents, à ma tante, aux Gexois (Léo, Mélaine

et les filles qui deviennent de plus en plus grandes). Je pense beaucoup à vous tous. Merci plus particulièrement à mes parents à qui je dédie cette thèse.

A mes parents,

Table of contents

Résumé 10
Abstract 12
General Introduction 19
Heterosis 21
Heterosis – history and definition 21
Heterosis – genetics basis 22
Factors affecting heterosis 24
Maize breeding history 25
From open-pollinated maize varieties to maize hybrids 25
Maize hybrid selection schemes – recurrent reciprocal selection 27
Understanding of the genetic basis of the phenotypic traits 29
QTL detection 29
Marker-Assisted Selection 31
Genomic selection 32
Presentation of the phD work 34
Appendix 1: Maize domestication and world diffusion 346
Appendix 2: Expression of the phenotype of a hybrid between two unrelated populations 38
Literature cited 41
Chapter 1 51
Linkage disequilibrium with linkage analysis of multi-line crosses reveals different multi-allelic QTL for hybrid performance in the flint and dent heterotic groups of maize 53
ABSTRACT 55
INTRODUCTION 56
MATERIAL AND METHODS 59
Plant material and phenotypic analysis 59
Genotyping and analysis of genotypic data 59
Clustering analysis of parental inbred lines 60
QTL detection 61
RESULTS 65
Analysis of parental linkage disequilibrium and parental clustering 65
Comparison of the thresholds used in the QTL detection models 66
Comparison of the QTL detected with the different models in the dent and flint designs 66

Allelic effect series and comparison of the different allelic models for the major QTL detected for female flowering time 67
Comparison of the QTL detected in the two heterotic groups analyzed individually and jointly 69
DISCUSSION 71
Linkage disequilibrium and clustering of parental alleles 71
Comparison of the different QTL detection models 73
Comparison between the QTL detected in the two heterotic groups and evolutionary interpretation 74
Overview of detected QTL and comparison with literature studies 75
Acknowledgment 77
Literature cited 78
Chapter 2 97
QTL detection on maize (*Zea mays* L.) hybrids derived from two multiparental designs allowed the simultaneous detection of QTLs for general and specific combining abilities for silage performances 99
ABSTRACT 100
INTRODUCTION 101
MATERIAL AND METHODS 103
Genetic material 103
Genotyping data 104
Field trial design and analysis 106
Variance component analysis 106
Computation of adjusted means and correlations between traits 109
QTL detection 109
RESULTS 112
Genetic variance analysis 112
Adjusted means and correlations between traits 115
QTL detection 116
DISCUSSION 122
Genetic variance components 122
QTL detection 123
Improvement of the QTL detection models 125
Conclusions 126
Acknowledgements 127
Literature cited 128

Chapter 3 135

Can genomic selection deeply modify early stages of hybrid breeding programs? A case study on maize (*Zea mays* L.) silage performance of single-crosses between two multiparental populations 137

INTRODUCTION 138

MATERIAL AND METHODS 140

Plant material 140

Genotyping data 141

Phenotypic evaluation and adjusted means 141

Genomic Best Linear Unbiased Prediction (GBLUP) models and estimation of variance components 142

GBLUP model using kinship matrices based on founder alleles 143

Quality of prediction of the different GBLUP models. Comparison with prediction based on QTL 144

Influence of the calibration set size on the hybrid predictions 145

Influence of the calibration set composition on the hybrid predictions 145

Impact of including half-sib hybrids in calibration set 145

Impact of including hybrids issued from the same founder lines in calibration set 145

RESULTS 147

Estimation of variance components and R^2 of the GBLUP models 147

Predictive abilities of the GBLUP models. Comparison with predictions based on QTL 149

Quality of prediction of GCA 151

Quality of prediction of GCA 151

Effect of the size of the calibration set 151

Calibration set composition 152

Impact of including half-sibs hybrids in calibration set 152

Impact of including hybrids issued from the same founder lines in calibration set 153

DISCUSSION 154

GCA/SCA Variance components 154

Comparison of the GBLUP models 155

Comparison with QTL - Combining the two approaches 155

Optimization of the calibration set for hybrid prediction 156

Reconsidering hybrid breeding selection schemes? 157

Conclusion 158

Acknowledgements 158

Literature cited 160

General discussion 167
QTL detection in multiparental designs 169
Multiallelic QTLs heterotic group dependent 169
Prospects for QTL detection 170
Genomic prediction in multiparental reciprocal designs 171
Applied prospects of the reciprocal multiparental design experiment 173
Prospects for enhancing genomic prediction efficiency 173
Reconsidering early steps of maize hybrid breeding selection programs? 175
Literature cited 177
General appendices 181
Appendix I: Supporting information of Chapter 1 183
Appendix II: Supporting information of Chapter 2 221
Appendix III: Supporting information of Chapter 3 233
Appendix IV: Résumé substantiel en français 237

Résumé

Le maïs (*Zea mays* L.) est une culture majeure présentant un hétérosis important pour les caractères liés à la biomasse. Pour exploiter ce phénomène, les programmes de sélection ont été organisés en groupes hétérotiques et la majorité des surfaces cultivées correspond à des hybrides F1. La valeur hybride peut être décomposée en Aptitudes Générales à la Combinaison (AGC) des lignées parentales et en Aptitude Spécifique à la Combinaison (ASC).

L'objectif de cette thèse était d'apporter de nouveaux éclairages sur la valeur hybride, concernant tant la compréhension de ses déterminismes génétiques sous-jacents que sa prédiction. Deux dispositifs multiparentaux connectés ont été analysés, mettant en jeu deux des principaux groupes hétérotiques utilisés pour la production de maïs fourrage dans le nord de l'Europe : les cornés et les dentés. Le premier dispositif était constitué de deux populations de type Nested Association Mapping (NAM). La population dentée comprenait dix familles biparentales et la cornée onze. Ces populations ont été évaluées pour leur valeur en croisement avec un testeur du groupe complémentaire. Le deuxième dispositif était constitué d'un factoriel entre deux populations multiparentales de lignées : une cornée et une denté, dérivées de deux demi-diallèles entre quatre lignées fondatrices. Les lignées ont été croisées entre elles selon un plan factoriel incomplet afin d'obtenir des hybrides inter-groupes. Pour les deux dispositifs, le phénotypage des hybrides a porté sur le rendement ensilage, le contenu en matière sèche, la date de floraison femelle et la hauteur de plante. Les hybrides des dispositifs NAM ont également été phénotypés pour la date de floraison mâle.

Une détection des locus impliqués dans la variation des caractères quantitatifs (QTL) a été mise en œuvre en utilisant des modèles prenant en compte différents codages alléliques : allèles fondateurs haplotypes ou allèle observé au marqueur. Ces codages ont permis de mettre en œuvre des modèles de types LA (Linkage Analysis) ou LDLA (Linkage Disequilibrium - Linkage Analysis). Une complémentarité des modèles a été observée. Certains modèles ont permis la détection de QTLs multi-alléliques. Des QTLs différents ont été détectés dans les deux groupes hétérotiques, confirmant leur divergence ancienne. Pour le contenu en matière sèche et le rendement, nous n'avons pas détecté de QTL à effet majeur dans aucun des dispositifs à l'exception d'un QTL corné de floraison à effet pléiotropique dans le groupe corné. Bien que l'ASC représentait 20% de la variance génétique intra-

population totale pour la plupart des caractères (sauf la hauteur de plante) le dispositif factoriel n'a permis la détection que d'un faible nombre de QTLs ayant un effet sur la dominance/ l'ASC.

Différents modèles de sélection génomique de type GBLUP ont été mis œuvre dans le dispositif factoriel. Nous avons obtenu de bonnes qualités de prédictions, bien que la majorité des lignées ne fût parentes que d'un seul hybride. Nous avons montré qu'il était important de calibrer les prédictions sur des hybrides issus de lignées apparentées aux hybrides à prédire. Nos modèles ont montré leurs limites quant à la capacité de prédiction de l'ASC mais ont permis d'obtenir de bonnes qualités de prédiction pour les AGC des lignées parentales. Notre étude ouvre de nouvelles perspectives pour reconsidérer les schémas de sélection des hybrides de maïs. Les évaluations des lignées candidates sur testeur pourraient être remplacées par une évaluation d'hybrides obtenus selon un plan de croisement factoriel incomplet entre groupes hétérotiques. Des évaluations complémentaires sont nécessaires pour comparer plus directement notre stratégie avec celle basée sur l'utilisation de testeurs. Cependant des premiers résultats encourageants ont été obtenus.

Mots clés : maïs, valeur hybride, détection de QTL, sélection génomique, ensilage

Abstract

Maize (*Zea mays* L.) is a major crop presenting strong heterosis for traits linked to biomass. To exploit this phenomenon, F1 hybrids are cultivated and breeding programs are organized in heterotic groups. Hybrid value can be decomposed in General Combining Abilities (GCA) of the parental lines and Specific Combining Ability (SCA).

The goal of this thesis was to bring new insights to the understanding and the prediction of hybrid value and its components for silage maize. To do so, two multiparental connected designs, involving hybrids between the dent and flint heterotic groups, main heterotic groups used for silage breeding in Northern Europe, were analyzed. The first design consisted in two Nested Association Mapping (NAM) populations involving test-cross hybrids. The dent population consisted of ten biparental dent families and the flint one of 11 biparental families. The second design consisted of two multiparental populations of inbred lines, one dent and one flint, derived from two half diallels between four founder lines. Inbred lines from both groups were crossed according to an incomplete factorial to produce experimental hybrids. Hybrids were phenotyped for dry matter yield, dry matter content, female flowering date, plant height, and only for the NAM designs male flowering date.

Quantitative Trait Loci (QTL) detection using Linkage Analysis (LA) and Linkage Disequilibrium – Linkage Analysis (LDLA) methods was performed in both designs. QTL detection models were found to be complementary. Some of the QTL detection models allowed us to detect multiallelic series at the QTLs, showing the importance of not directly considering the observed biallelic genotype. Different QTLs were detected in the two heterotic groups, highlighting their long-term divergence. No QTL showed major effect for DMC and DMY except one pleiotropic QTL for flowering time in the flint group. The reciprocal multiparental design allowed the detection of QTLs for dominance/SCA. However, very few QTLs had significant dominance/SCA effects even if SCA represented around 20% of the within-genetic variance, except for PH for which it was lower.

Genomic selection was conducted in the second design using different GBLUP models. We obtained good predictive abilities even though predicted hybrids were obtained through a highly incomplete factorial mating where the majority of the inbred lines were parents of only one hybrid. Variation within populations sharing three founders with populations used for calibration could be predicted efficiently. Our models showed a limited ability for predicting SCA but predicted well the GCA of the parental lines. This opens new prospects for reconsidering maize breeding. Indeed test-cross evaluation, usually used in the breeding

companies, could be replaced by single-cross evaluation, according to an incomplete factorial design, between the two heterotic groups to improve. Further evaluations are needed for comparing the efficiency of this strategy with the one of tester-based designs, but first obtained results are encouraging.

Keywords : maize, hybrid value, QTL detection, genomic selection, silage

General Introduction

Maize is widely used all over the world for food, grain feed, silage and over the last years also for bioethanol. It is grown on 140 million of hectares in the world and is the first cereal in the world in terms of volume of grain. It is the species that prompted the definition of the concept of heterosis (Shull 1914) and the development of F1 hybrids (Shull 1908), leading to the subsequent implementation of advanced breeding methods. In this introduction, we will focus on heterosis, presenting its genetic basis and the factors affecting it. Then, we will present maize breeding history and selection schemes before giving an insight on the contribution of molecular markers to understanding the genetics determinisms of phenotypic traits and improving breeding schemes. Finally, we will briefly present the experiments and approaches that were conducted during this PhD.

Heterosis

Heterosis – history and definition

In 1766, Koelreuter showed that some interspecific crosses in several genus such as *Nicotinia* and *Datura* could lead to hybrids presenting stronger vigor than their parents. Darwin in 1876, observed that for numerous species intraspecific cross-pollination led to better plants than self-pollination with increased performances for quantitative traits. This observation was later theorized by Shull (1908) at Cold Spring Harbor Laboratories in New York and East (1908) at Connecticut State College who defined the concept of heterosis (Shull 1914), after maize (*Zea mays* L. or *Zea mays* ssp *mays*) observations. There is heterosis when a hybrid offspring from genetically diverse individuals or populations shows increased vigor relative to its parents (Fu *et al.* 2014). According to the distance between the parents of the hybrid, three categories of heterosis can be defined depending if it is a cross between: (i) two individuals from two different species, (ii) two individuals from two different subspecies, (iii) two individuals from the same subspecies (Fu *et al.* 2014). Heterosis is widely used in plant and animal breeding. One of its earliest applications was mule breeding, mules being derived from the cross of individuals from two species: a female horse (*Equus caballus*) and a male donkey (*Equus asinus*) (Springer and Stupar 2007). One example in plants of heterosis from wide hybridization is Triticale, a polyploid hybrid of both rye (*Secale* spp.) and wheat (*Triticum* spp.) which is widely sown (Fu *et al.* 2014). The rice hybrid *indica* x *japonica* which presents between 8 to 15% more heterotic potential than intraspecific hybrids for many traits is an example of intersubspecies heterosis. However, at first, these rice hybrids were not widely adopted due to problems of unstable seed setting and poor grain plumpness, problems solved recently (Fu *et al.* 2014). Intraspecies heterosis is certainly the type of heterosis that is the most used in plant breeding and more

specifically in maize breeding (Figure 1). We will focus on intraspecific heterosis in the following and will refer to it as “heterosis”.

Heterosis can be defined in two different ways. Geneticist heterosis (or mid-parent heterosis) is when one genotype is better than the mean of its two parents. Breeder heterosis (or best parent heterosis) is when the hybrid is better than the best of its parents. Note that although not strictly academic farmers may also have a third vision of heterosis which is the difference in performance between the best hybrid available on the market and the best inbred line variety. From a statistical point of view, heterosis is a deviation to additivity. Its conceptual opposite is the inbreeding depression, which is a gradual reduction of vigor after reproduction of related individuals, over several generations.

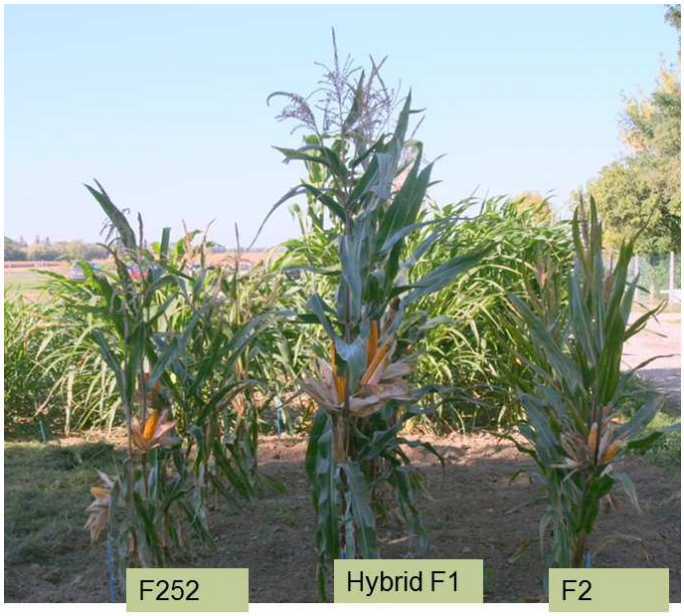

Figure 1: Illustration of phenotypic heterosis in the F2 x F252 maize hybrid, in the center (Photo: Julie Fiévet).

Heterosis – genetics basis

Even if heterosis is widely used in plant and animal breeding, our knowledge of the mechanisms underlying it is partial. Three main non-exclusive hypotheses exist regarding the phenomena underlying heterosis for a given trait in a given environment: dominance, overdominance and epistasis (Figures 2 and 3).

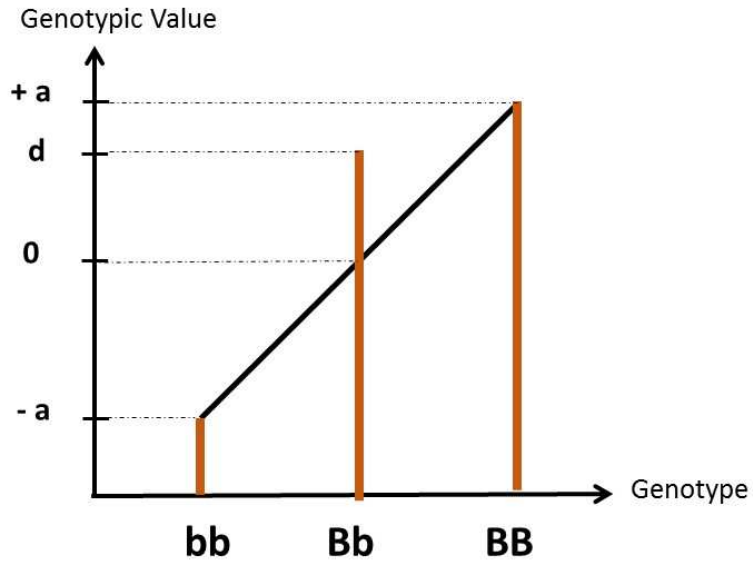

Figure 2: Definition of genotypic value in a biallelic situation for one locus. Genotypic values of the genotypes bb, Bb and BB, are defined by the parameters a and d. a is the biological additive effect of the locus and d is the biological dominance effect of the locus. The genotypic value of the genotype bb is $-a$, genotypic value of the genotype BB is a. When there is strict additivity, genotypic value of genotype Bb is 0, whereas when there is dominance of allele B over allele b (situation presented here) it is equal to d. If $^{d}/_{a} = 1$, there is complete dominance, if $0 < {}^{d}/_{a} < 1$ there is partial dominance and if $^{d}/_{a} > 1$there is overdominance.

In the dominance hypothesis, heterosis is explained by the fact that each parental line carries recessive unfavorable alleles at different loci so that their effects are masked by their dominant counterpart when assembled in the hybrid (Figure 3). The hybrid benefits from the complementation of these deleterious alleles (Davenport 1908; Jones 1917). Overdominance is an intralocus complementation for which the heterozygote state at one locus is phenotypically superior to both homozygote states (Hull 1946) (Figure 3). Few examples of overdominance exist and at our knowledge there is only one overdominance action which was fully proved in maize. Hollick and Chandler (1998) provided evidence for overdominance at the *Pl* locus (purple plant locus) which controls the leaves anthocyanin synthesis. Strong correlation between two locus (or Linkage Disequilibrium LD) can be a problem for distinguishing dominance from overdominance gene action on phenotypes (Springer and Stupar 2007) since apparent overdominance can be the result of tight linkage between two loci with favorable dominant alleles in repulsion. This situation is called pseudo-overdominance (Jones 1917) (Figure 3). Note that in situations where the favorable allele at a locus depends on the environment and is dominant, one can observe an overdominance effect for the average performance, referred to as marginal overdominance (Wallace 1968). Epistasis which

corresponds to the interaction between alleles at two or more loci can also have an impact on heterosis (Richey 1942; Powers 1944; Jinks and Jones 1958; Williams 1959).

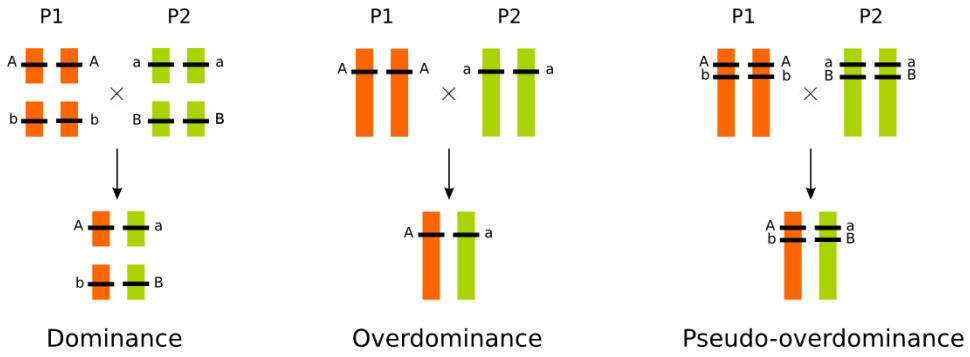

Figure 3: Genetic mechanisms proposed to explain heterosis. P1 and P2 correspond to the two parental lines. Alleles a and b are recessive and unfavorable, alleles A and B are dominant and favorable.

Various molecular mechanisms may explain these genetic mechanisms. In addition to non-synonymous single base mutations that can cause loss of function, maize presents a lot of structural variation as difference in genome organization and presence-absence variations. These presence-absence variations can be at the gene level, gene regulating regions level or components of gene families level (Springer and Stupar 2007). Combination of inbred lines presenting such variations can have an impact on the level of gene expression. One example is the expression in the B73 x BSSS53 hybrid of the *zein1C* gene family controlling the development of endosperm tissue. B73 expresses six genes of the family, BSSS53 seven, but only three are shared between the two genotypes. The hybrid expresses all 10 *zein1C* genes (Springer and Stupar 2007).

Factors affecting heterosis

Plant heterosis is highly variable depending on species. In general, heterosis is stronger in allogamous species (such as maize, onion and alfalfa) than in autogamous species (such as wheat, rice, *Arabidopsis*, tomato). This is because allogamy allows the development of the genetic load as recessive slightly deleterious alleles at low frequency are masked at the heterozygote state. In autogamous species, recessive unfavorable alleles cannot be masked and are thus counter selected, which explains why genetic burden is much less important (Gallais 1989; Gallais 2009). For allogamous species, when comparing inbred lines to the F_1 generation, heterosis for grain yield or biomass is of 100 to 400% whereas for autogamous species it is between 0 and 100%. However, this result needs to be nuanced as for allogamous species it would be more logical to compare hybrid value with that of populations. Inbred lines are indeed

created by forcing allogamous plants to become autogamous. The advantage of hybrids compared to populations from which parental lines are extracted is lower, between 10 and 15% (Morrow and Gardner 1893). As heterosis is due in part to dominance, it is expected to be stronger when "distant" individuals are crossed. However, the relationship between genetic distance of the parents and importance of heterosis in the F_1 is not straightforward. It has been observed that when genetic distance between the parents is too important, heterosis can be lower than for crosses between closer individuals (Moll *et al.* 1965 for an example on maize).

Heterosis is dependent on the considered trait. Indeed, in maize, heterosis is more important for grain yield than for plant height, which itself present more heterosis than leave width. Number of seeds presents more heterosis than thousand grain weight. Heterosis has a minor impact on number of leaves and number of ranks per kernel, which are mainly additive traits (Gallais 2009). Thus, heterosis tends to be more important for traits linked to fitness than for other traits. It is more important for complex traits as grain yield or silage yield than for traits with simple genetic determinism (Gallais 2009). One explanation can be that many complex traits are multiplicative and can be decomposed in elementary components. If elementary traits present heterosis, the complex trait corresponding to their product will present more heterosis. In addition, multiplication of elementary components having complementarity characteristics in both parents (as one parent with long and narrow leaves and one parent with short and wide leaves), and not necessarily presenting heterosis, can conduct at heterosis at the hybrid level for the complex trait (Gallais 2009). Heterosis does not only affect the young or adult plants but is manifest already in the early stages, at the beginning of embryogenesis. For instance, size of the hybrid embryo six days after fecundation, speed of root development and some embryo enzymatic activities present mid-parent heterosis (Gallais 2009).

Maize breeding history

From open-pollinated maize varieties to maize hybrids

The theorization by East and Shull of the heterosis phenomenon, has accompanied the transition of cultivated maize from landraces and open-pollinated varieties to hybrids in the USA. Open-pollinated varieties correspond to a population of individuals all different one from each other, resulting from random crosses of plants from the previous generation. These populations have been created by mass selection and adaptation to environmental conditions from ancestral populations originated from Mexico (see Appendix 1 for a brief overview of the history of maize). F_1 hybrids were seen as a way to produce at a large scale the best combination of gametes that could be derived from such populations (Shull 1908). As noted by Shull this idea

is an extension of the isolation concept that was applied to autogamous species in order to make it applicable in the presence of inbreeding depression. Maize morphology allows an easy production of hybrids seeds as male flowers are on the top of the plant which can thus be easily castrated which allows a large-scale production of hybrid seeds (Morrow and Gardner 1893). However, at the beginning, one main issue was the low rate of seeds produced by the first inbred lines derived from populations, due to the inbreeding depression. To circumvent this limitation and allow a production in quantity of identical seeds, Jones (1918) proposed the use of triple-cross hybrids and double-cross hybrids. From 1922-1924, best hybrids, as for example Cooper Cross, presented a repeatable advantage of around 10% compared to the best open-pollinated varieties (Charcosset 2002). Such practical results led in the 1930s to the domination of triple-cross and double-cross hybrids over open-pollinated varieties (Troyer 1999). For instance, in Iowa in 1935 less of 10% of the surface was sown with hybrids whereas it was over 90% in 1939 (Reif *et al.* 2005). Then, in the 1960s, with the improvement of inbred lines *per se* value, simple hybrids could be developed (Troyer 1999) and are now widely used (Figure 4). This transition from open-pollinated varieties to double-cross and then single-cross hybrids was concomitant to an increase in maize yield (Figure 4). In the 1950s (North Central Regional Corn Improvement Conference 1949; Lamkey and Lorenz 2014), heterotic groups, structuring the genetic diversity, were created in order to increase heterosis, by avoiding relatedness between parental lines used as parents of hybrids. Heterotic groups can be defined "as a group of related or unrelated genotypes from the same or different populations, which display similar combining ability and heterotic response when crossed with genotypes from other genetically distinct germplasm groups" (Melchinger and Gumber 1998).

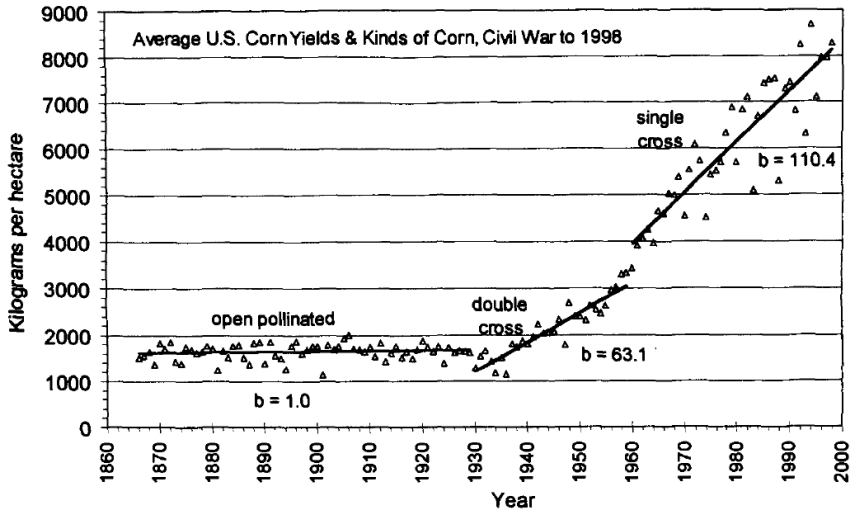

Figure 4: Evolution of U.S. maize yields and kinds of corn from 1860 to 1998; periods dominated by open-pollinated, by double cross and single cross hybrids are shown. (From Troyer 1999).

First hybrids cultivated in Europe, before the late 1950s, were imported from USA and had only a limited success. After World War II, programs were started to develop hybrids adapted to Northern European conditions. In France, new lines were developed from the flint French Southwestern populations and crossed to dent inbred lines from the Corn Belt dent region in USA, as Wisconsin inbreds (Troyer and Hendrickson 2007). The most successful flint inbred lines, F2 and F7, were derived from the Lacaune population (Tenaillon and Charcosset 2011). In regions at the North of Loire valley, the resulting hybrids over yielded and were better adapted than the early flowering hybrids imported from Northern USA (Troyer and Hendrickson 2007; Tenaillon and Charcosset 2011). Currently, these dent and flint heterotic groups are still the ones mainly used for maize breeding in Northern Europe, for silage as well as for grain maize, although there is a growing tendency to introgress the flint pool with dent material. In breeding companies, these main heterotic groups are structured in smaller heterotic groups derived from the most popular inbred lines of their breeding programs.

Maize hybrid selection schemes – recurrent reciprocal selection

One main concern of hybrid breeding is to select the parental lines which in combination will give the best hybrid. One first option is to select the parental lines based on their *per se* value (Jenkins 1929). However, correlation is low between *per se* value of an inbred line and the values of hybrids that can be derived from it for traits showing heterosis (Richey 1924; Richey and Mayer 1925). *Per se* value selection generally allows the elimination of only the worst parents (Gallais 2009). In 1932, Jenkins and Bruson showed that the average value of the hybrids derived from an inbred line is better correlated to the value of the progeny of this inbred line with another inbred line, an hybrid or a population than to its *per se* value. Sprague (1939) and Jenkins (1940) suggested that combining ability of the tested inbred lines should be taken into account at early stages in the inbreeding process. Sprague and Tatum (1942) clearly defined the notions of combining abilities and introduced the partitioning of hybrid values in terms of General and Specific Combining Abilities (GCA and SCA respectively). The GCA of a line corresponds to its average performance in hybrid combinations. The SCA of a pair of parental lines is the difference between the hybrid value and its value predicted based on GCAs. It corresponds to the cases “in which certain combination do relatively better or worse than would be expected on the basis of the average performance of the lines involved” (Sprague and Tatum 1942).These notions can be extended to the case of hybrids between two heterotic groups (see Appendix 2).

Hayes and Garber (1919) are credited as the first users of recurrent selection for maize breeding

(Hallauer *et al.* 2010, Chapter 6), that is to say methods of selection conducted recurrently, *i.e.* when similar procedures are repeated in successive cycles of selection. Recurrent selection methods are designed to increase the frequency of favorable alleles for quantitative traits while maintaining genetic variability allowing continuous genetic improvement (Hallauer *et al.* 2010, Chapter 6). Jenkins (1940) suggested a method, modified by Hull in 1945, for improving SCA of a heterozygous population thanks to the use of a tester with a narrow genetic base. The used tester may be a pure line or a hybrid and is one of the parents of the future commercial hybrid. Hull's method, which did not allow for much improvement on the tester side, would be efficient in presence of overdominance but inefficient for dominance when the tester carries favorable alleles (Hull 1945; Comstock *et al.* 1949). Another possibility is to perform recurrent selection for GCAs. In this case, plants from one population are evaluated in cross with a ester presenting broad genetic base, identical for all populations. The tester used consisted in at least two single crosses between pure lines or in a variety. Then, selected plants of each populations are crossed and a new cycle can be initiated (Comstock *et al.* 1949; Lonnquist 1957). These method is more efficient than the previous one in presence of dominance but not in presence of overdominance (Comstock *et al.* 1949). Comstock *et al.* proposed in 1949 a reciprocal recurrent method which is efficient regardless of the genetic mechanisms underlying heterosis and which maximize use of GCAs and SCA (Figure 5). Their procedure consisted in the simultaneous improvement of the two parental populations of the hybrid, testing plants from one population by crossing them with plants of the other population. Each plant of one population is crossed with four or five plants of the other population and the resulting hybrids are evaluated. In each population, plants are selected based on the value of their bulked test-cross progeny. Reciprocal recurrent selection proved its efficiency in several selection programs initiated in the 1960s and 1970s (see Hallauer *et al.* 2010, Chapter 7). Variations of the reciprocal recurrent selection procedure proposed by Comstock *et al.* (1949) were proposed as for instance reciprocal recurrent selection based on testcrosses of half-sib families (Paterniani 1967) which reduces the effort for making testcrosses, reciprocal recurrent selection based on half-sib progenies of prolific plants (Hallauer *et al.* 2010, Chapter 12) and reciprocal full-sib selection on prolific plants (Hallauer and Eberhart 1970) or on one-ear plants (Hallauer *et al.* 2010, Chapter 12). In reciprocal full-sib selection, contrary to the method proposed by Comstock *et al.* (1949), full-sib progenies are evaluated rather than half-sib progenies (see Hallauer *et al.* 2010, Chapter 12 for more details on the procedure). These selection schemes showed their interest for maize hybrid breeding. Coors (1999) synthetized a large number of studies for several breeding method. He showed that reciprocal half-sib and full-sib recurrent selection methods were the ones allowing the

highest direct responses for interpopulation improvement on a cycle basis and led to the highest gains in grain yield on an annual basis.

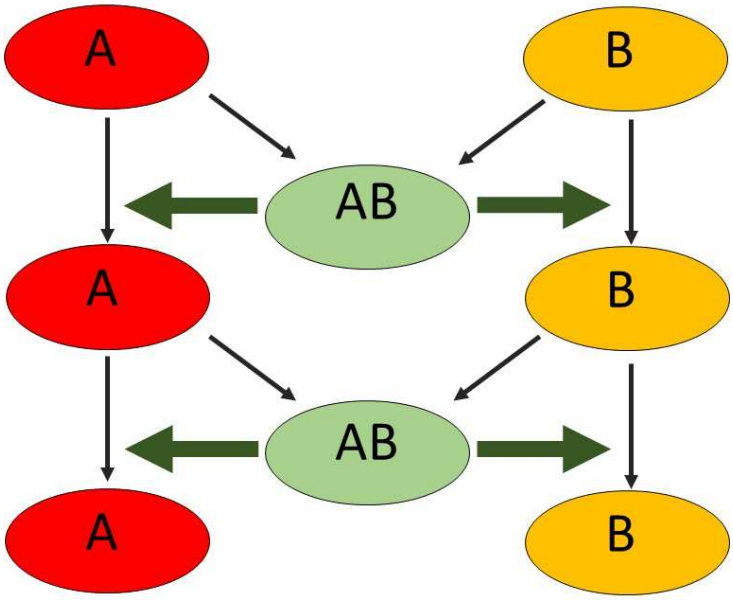

Figure 5: Reciprocal recurrent selection for two complementary populations A and B (adapted from Varona and Toro, 2012).

In the breeding companies, generally a modified recurrent reciprocal selection procedure is used. Within each heterotic group; breeders cross pairs of inbred lines to generate biparental segregating populations that are evaluated for their test-cross values. Generally, a few testers (two or three) of the opposite heterotic group are used and correspond to good parental lines, which will be one of the future parents of the released hybrids. Selected plants are then selfed during several generations to derive new inbred lines. In a second stage, inbred lines selected in the different groups are crossed in order to identify the best hybrid combinations.

Understanding of the genetic basis of the phenotypic traits

QTL detection

Conventional breeding methods can be considered as “blind” approaches. When they were developed little was known on the genes or loci (or Quantitative Trait Loci, QTL) involved in the variation of quantitative traits. Detecting these QTLs is important for understanding the genetic basis of the traits. Molecular markers, when correlated to the trait, can bring information on location and polymorphism of these loci. The first molecular markers used were protein variants (isozymes) identified by electrophoresis and developed in the 1960s. They presented codominant variations but were not very polymorphic and did not cover the entire genome.

Then, new approaches appeared, enabling to detect polymorphism at the DNA level: first Restriction Fragment Length Polymorphisms (RFLP) in the 1980s, then Amplified Fragment Length Polymorphisms (AFLP) and Simple Sequence Repeats (SSR) in the 1990s and finally Single Nucleotide Polymorphisms (SNP) and Copy-Number Variations (CNV). Two main methods exist for detecting QTLs: QTL detection or Linkage Analysis (LA) mapping and association mapping or Linkage Disequilibrium (LD) mapping. These methods use the fact that polymorphisms at two loci of the genome are in LD (not randomly associated) depending of a certain number of factors as genetic distance, selection, drift etc.. Consequently, allelic information on one locus (*e.g.* a marker) can give information on the allele at another locus (*e.g.* the gene) with which it is in LD.

In plants, first QTL detections were carried out in segregating populations derived from crosses between two inbred lines (Paterson *et al.* 1988 on RFLP).This approach relies on the fact that in these populations, correlations between markers and QTL is a simple function of the recombination rate. It has a limited resolution (Darvasi and Soller 1997) due to the low number of recombination events that occurred in such population. Moreover each biparental population represents only a small part of the generic diversity available. To circumvent these limitations, with the increase of marker density and the development of genotyping chips, it became possible to detect QTLs in a more diverse material, with no family relationships. It is the association mapping approach. This approach takes advantage of the LD present in panel of individuals that has been broken along generations by historical recombinations (Pritchard and Przeworski 2001) which allows a better resolution than QTL detection in biparental populations. However, in association mapping populations, contrary to linkage mapping populations, LD is not only due to genetic linkage but also to the history of the population. It can be caused by structure, relatedness, drift and selection. To reduce false-positive associations caused by the history of the population, additional random or fixed terms (corresponding to structure or relatedness) can be added in the association mapping model (Yu *et al.* 2006).

Approaches were developed for improving QTL detection based on LA-mapping. For instance, QTL detection can be carried out in highly recombining inbred lines or intermated lines (Darvasi and Soller 1995; Huang *et al.* 2010). The higher number of recombinations in comparison to a traditional LA-mapping population allows a better accuracy of estimated QTLs positions. Synthesis of the information of several biparental QTL detection studies is also possible by performing meta-analysis (Goffinet and Gerber 2000; see Truntzler *et al.* 2010 for an example on silage maize) which can allow to explore more diversity and improve the resolution of the QTL mapping. Multiparental designs can also be a solution for improving the

precision of QTL location while exploring more diversity. Among these designs, multiple connected biparental families designs can be developed by assembling biparental populations that have one parents in common (Rebaï *et al.* 1997). Nested Association Mapping (NAM) designs (Yu *et al.* 2008) are a specific case of such designs in which all biparental populations shred the same parent. Another possibility is the development of multiparental advanced-generation intercross (MAGIC) populations (Huang *et al.* 2015) where the initial crosses of the various founder lines are followed by several generations of inter-mating. Two types of analyses can be performed on multiparental designs: traditional linkage based analyses when looking at the parental alleles or LDLA analyses which synthesize both LD and LA approaches by looking at ancestral alleles or at the observed marker alleles.

Marker-Assisted Selection

Knowing position of QTL and of markers physically liked to the QTL is important from a fundamental point of view and offers interesting perspectives for Marker-Assisted Selection (MAS). Interest of markers in breeding was first discussed by Neimann-Sorensen and Robertson in 1961, in an animal breeding context. It is only in the 1980s that MAS became a more tangible reality in animal and plant breeding. One application of MAS is for traits determined by a major gene and difficult or expensive to phenotype. When the favorable allele at the major gene is associated to a specific allele at a marker, it can be cheaper and easier to genotype and screen for the marker than to phenotype. For more complex traits, influenced by many genes, Lande and Thompson (1990) proposed to estimate the genetic value of the individuals based on the sum of the effects of markers significantly associated to QTLs. Phenotypic information can also be integrated to the selection scheme and integration of phenotype and genotype information lead to different MAS schemes. Hospital *et al.* (1997) shown that one of the main interest of MAS was an increase of genetic grain per unit of time when cycles with only marker information and with marker and phenotypic information were alternated. MAS is expected to be more interesting than conventional selection based on phenotype for traits with low heritability provided that QTLs can be detected (Hospital *et al.* 1997; Moreau *et al.* 1998). One limit of these approaches is that selection on markers only lead to the quick fixation of favorable alleles at the biggest QTLs (Hospital *et al.* 1997; Moreau *et al.* 2004) whereas unfavorable alleles can be fixed at the smallest QTLs. Another limit is that effects of the detected QTLs are often overestimated due to the Beavis effect (Beavis 1998) which lead to wrong weightings of the QTL effects in predictions based on markers and a reduced efficiency of MAS (Moreau *et al.* 1998; Melchinger *et al.* 1998). When using detected

QTLs, only a limited proportion of genetic variance is taken into account and "missing heritability" (Maher 2008) is important. Even if some studies reported a limited efficiency of MAS in biparental populations (Moreau *et al.* 2004) others, notably some issued from the private sector, reported its interest (Eathington *et al.* 2007). MAS appears particularly interesting when considering connected multiparental populations to assemble favorable alleles issued from several founder lines. Interest of this approach was tested by simulations (Blanc *et al.* 2008) and validated experimentally (Moreau and Charcosset 2011).

Genomic selection

To circumvent some of QTL detection limitations, when the studied trait is controlled by many QTLs, one possibility is to use genomic selection. The basic principle of this method is to estimate all marker effects simultaneously, and use these effects for prediction, without conducting first a QTL detection step. The idea behind this approach is that if enough markers are available, the LD between markers and QTLs will enable markers to capture QTLs effects. However, as the number of markers is generally higher than the number of performances, conventional fixed-effect models used for QTL detection cannot be used. In 2000, Whittaker *et al.* proposed to use the ridge regression to estimate marker effects and showed its interest compared to predictions based on QTL detection. Meuwissen *et al.* (2001) defined the concept of genomic selection and proposed to use additional approaches, based on Bayesian statistics, for estimating maker effects. He also proposed one approach of genomic selection (called GBLUP) which consists in using markers to estimate kinship relationships between individuals and use this matrix in a BLUP model to predict values of non-phenotyped individuals using performances of phenotyped ones. This GBLUP model was proved to be statistically equivalent to a random ridge regression (RR-BLUP) (Habier *et al.* 2007). It has to be noted that a very similar approach to the GBLUP model was proposed already in 1994 by Bernardo who used marker-based distances between parental lines of single-crosses for predicting performances of non-phenotyped single-crosses using performances of a related set of single-crosses. Development of GS was favored by advances in high-throughput genotyping methods that are now available at a reasonable cost for most species of agronomical interest.

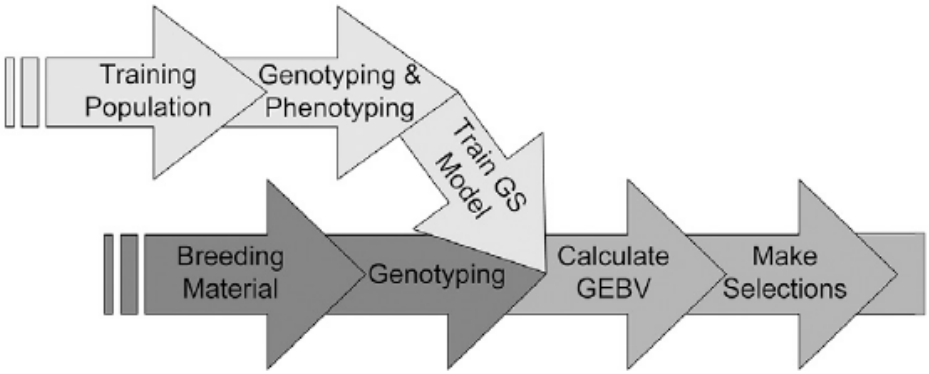

Figure 6: Diagram of genomic selection process with a single occurrence of model training, starting from the training population and selection candidates continuing through to genomic estimated breeding value (GEBV)–based selection (From Heffner *et al.* 2009).

In a first step, a statistical model is calibrated on a calibration set consisting in a certain number of individuals which are phenotyped and genotyped. In a second step, the calibrated model is used to predict the genetic value of individuals which are only genotyped and who belonged to what is called the prediction set (Figure 6). To obtain the prediction equation, a lot of methods exist (see Heslot *et al.* 2012 for a review). They can be parametric as GBLUP (equivalent to RR-BLUP) or Bayesian methods, or non-parametric as Reproducing Kernel Hilbert Spaces methods or neural networks.

Since 2006, GS showed its practical interest in dairy cows where it is now largely implemented. It was facilitated by the pooling between different countries of phenotyping, genotyping and pedigree information in order to create a huge reference population used for calibration (through for instance the EuroGenomics consortium which groups European private and public actors of Holstein breeding). In plant breeding, simulations and fields experiments gave encouraging results for implementation of GS in populations with variable levels of diversity. For instance, Bernardo and Yu (2007) showed using stimulations that GS provided 18 to 43% more genetic gain per cycle than traditional marker-assisted recurrent selection based on QTLs in biparental populations.

For maximizing the quality of genomic predictions, some critical parameters must be taken into account when designing GS procedures. Choices of the statistical model and of the marker density are important. Size of the calibration set (Technow *et al.* 2013; Lehermeier *et al.* 2014) and choice of the individuals in it (Rincent *et al.* 2012; Akdemir *et al.* 2015), and for instance the genetic distance between the calibration set and the prediction set (Riedelsheimer *et al.* 2013; Lehermeier *et al.* 2014) need to be considered.

Some limits of the current GS approaches are under investigation such as the incorporation in

the model of GxE interactions (Burgueño *et al.* 2012; Heslot *et al.* 2014). Several recent studies have been published on the prediction of hybrids between two populations and the inclusion of dominance and SCA in the GS models. This question is of high interest for animal genetics (Toro and Varona, 2010; Amuzu-Aweh *et al.* 2013; Ertl *el al.* 2014; Sun *et al.* 2014) and for plant genetics (see Zhao *et al.* 2013 for an example in wheat; Reif *et al.* 2013 for an example in sunflower; Xu *et al.* 2014 for an example in rice). In maize, where heterosis is strong, first exploratory studies based on GBLUP were carried out on small factorial designs (Bernardo 1994) or on hybrids from advanced selection staged of breeding programs (Bernardo 1996a; Bernardo 1996b). More recently, more important datasets were studied using GBLUP and alternative models, based on simulations (Technow *et al.* 2012) or data from the last steps of maize breeding programs (Maenhout *et al.* 2007; Maenhout *et al.* 2010; Massman *et al.* 2013; Technow *et al.* 2014). More studies are still needed on hybrid prediction especially for hybrids between inbred lines developed directly from segregating populations available at early selection stages.

Presentation of the phD work

In France, in 2013 silage maize was cultivated over 1.49 million of ha representing around 44% of the total superficies of maize. French silage maize yield was multiplied by two over the last 50 years passing from 150 000 Hg/Ha to 312 175 Hg/Ha (FAO stats). During decades, maize was bred for silage or grain indistinctively and breeding criteria were based on grain yield performances and not on silage performance (Surault *et al.* 2005). Since 1986, for the registration in the French Official Catalogue of Species and Varieties, maize varieties are classified into two groups: grain maize and silage maize. Since 1998, feed quality is taken into account for registration (Surault *et al.* 2005) through the Milk Forage Unit (MFU), which quantified the calorific energy brought to dairy cows by one kilo of forage. This late preoccupation for feed quality and digestibility led to a decrease of the MFU values (Surault *et al.* 2005) as well as Neutral Detergent Fiber (NDF) digestibility values (Barrière *et al.* 2004) for the hybrids registered in the French catalogue of varieties or in the European catalogue between 1958 and 2002. Understanding of the genetic determinisms below the traits subjected to silage breeding is of main importance.

The main objective of this phD work was to analyze the genetics of the hybrid value of silage maize in multiparental designs, using two different approaches: (i) QTL detection and (ii) genomic selection and two different strategies for hybrid production: (j) single tester approach and (jj) a “no tester” approach with a highly incomplete factorial mating design between two

populations of candidate lines. To do so, two multiparental silage maize designs were studied. The dent and flint heterotic groups, corresponding to the major heterotic groups used for silage breeding in Northern Europe are involved in both designs. The first design is a NAM design, actually composed of two NAM designs, corresponding to the dent and the flint heterotic groups. The dent design consisted of ten biparental dent families and the flint one of 11 biparental families. Inbred lines were evaluated as test-cross value, using for each heterotic group the central line of the other group as tester. This allowed us to analyze these designs separately and also jointly. Five phenotypic traits were studied: dry matter yield (DMY), dry matter content (DMC), female flowering (DtSILK), male flowering (DtTAS) and plant height (PH). For each heterotic group, we detected QTLs using a LA model taking into account connections between populations and three different LDLA models. The second studied design consisted of two multiparental designs, one dent and one flint, derived from two half diallel between four founder lines and crossed according to an incomplete factorial. In this case, phenotypic evaluation was carried out on hybrids between the dent and flint parental lines and not on test-cross hybrids. DMC, DMY, DtSILK and PH were phenotyped. To our knowledge, few studies exist concerning QTL detection or genomic selection directly on hybrids and none were carried out on material directly issued from the cross of segregating families available at early selection stages. Our design allowed us to perform QTL detection using LA and LDLA models and genomic selection on such a material.

The first chapter of this manuscript is dedicated to the QTL detection in the two NAM designs. It highlights the complementarities of the different QTL detection models which were performed and puts into evidence different multiallelic QTLs in the two heterotic groups. These findings complement the analysis of the same dataset in a genomic selection context (Lehermeier *et al.* 2014). The second and third chapters are devoted to the analyses carried out in the second design. First, results of QTL detection in this design are presented. We developed models for performing LA and LDLA QTL detection directly on hybrids between unselected lines of two heterotic groups. We found that some of the QTLs for GCA were different in both groups and that a few QTLs had an effect on SCA at an individual risk level of 5%. The third chapter is devoted to the implementation of GS prediction in this design. It also includes some elements of discussion on the interest of such design in selection compared to conventional tester designs. The three main chapters are organized as scientific articles. The first one was published in Genetics in 2014, the second one will be soon submitted to Genetics and the third one is organized in view of submission. Finally, limits and perspective of the present work will be discussed.

Appendix 1: Maize domestication and world diffusion

Archeology (Piperno *et al.* 2009) and genetics (Matsuoka *et al.* 2002; van Heerwaarden *et al.* 2011) shown that maize was domesticated in a valley in Southwestern Mexico 9000 years ago from the wild species teosinte *Zea mays* ssp *parviglumis*. During its domestication, it went through strong phenotypic transformations due to strong selection of a number of genes (Hufford *et al.* 2012). Following two different southwards roads, maize reached the Andeans mountains around 6000 years ago and South Brazil around 6700 years ago (Tenaillon and Charcosset 2011) (Figure A1). It expanded northwards, arriving in the South-West of USA around 4100 years ago and then in Northern USA and Canada 800 years later. Recent hybridization between subtropical Southern dent material (from South of USA) and Northern flint material (from North-West of USA), adapted to short cycles, took place in 1800 AC, creating Corn Belt dent material (Tenaillon and Charcosset 2011). Maize landraces can now be found from 40°S in South of Chile (close to Chiloe Island) to 50°N in Canada (close to Gaspé peninsula) and between + 0 meters above the seal level (Caribbean islands) to +3400meters in the Andean mountains.

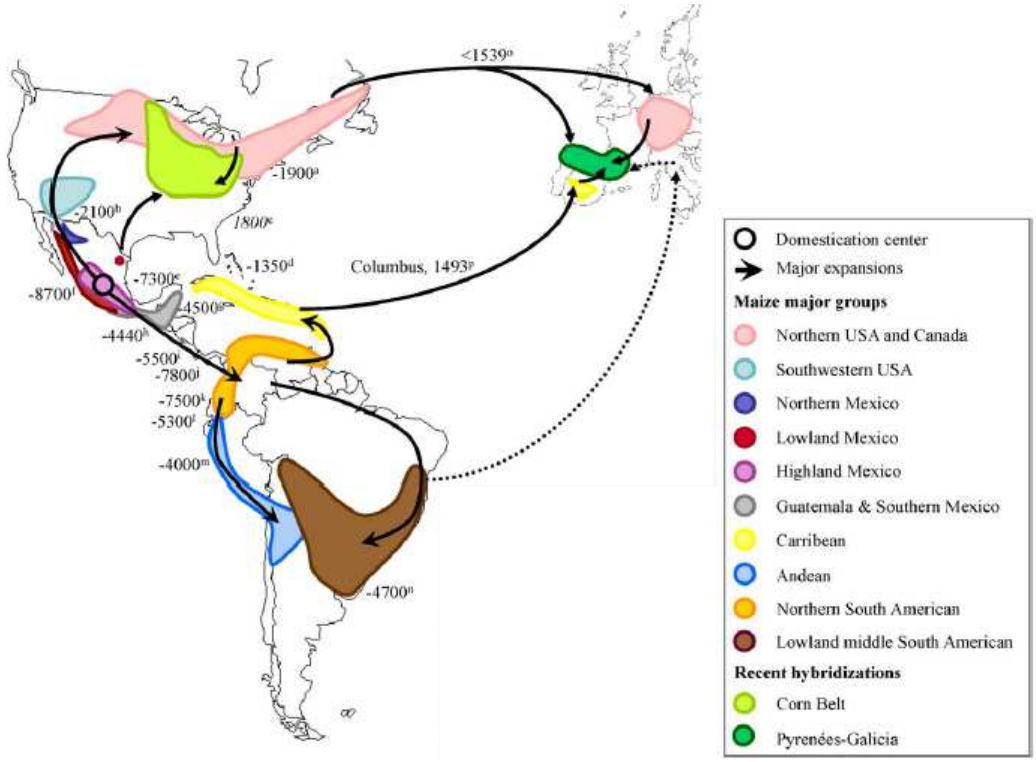

Figure A1: Domestication center and hypothetical diffusion of maize through the Americas and Europe (From Tenaillon and Charcosset 2011).

From 1493, maize spread all over the world, reaching Europe in 1493 and Africa and Asia since 1496 (Mir *et al.* 2013). Maize landraces were present in Europe long time before the broad cultivation of maize hybrids. They present a large range of morphological variation but all of them have flint kernel and white cob color (Rebourg *et al.* 2001). Maize was introduced in Europe through two main roads (Tenaillon and Charcosset 2011) (Figure A1). The first introduction is due to Christopher Columbus who brought in Spain, in 1493, Caribbean maize from Hispaniola Island (presently the Dominican Republic and Haiti). Due to its climatic needs, its cultivation probably remained confined to Southern Spain (Rebourg *et al.* 2003). A second introduction took place before 1539 from official or unofficial expeditions from Europe to the Northern American coast. It is at the origin of the presence of Northern flint maize in Germany and Northern Europe (Dubreuil *et al.* 2006; Tenaillon and Charcosset 2011). A third introduction in Italy probably took place in the XVIth century (Tenaillon and Charcosset 2011). Consequently to these introductions, maize was widely cultivated in Europe in the late XVIth, early XVIIth century, with less variability in Northern and Eastern Europe probably because of selection for tolerance to lower temperatures (Rebourg *et al.* 2001). After these main introductions, a new maize genetic group was created from hybridization between maize from Southern Spain and maize from France and Germany in the Pyreneans mountains (Tenaillon and Charcosset 2011). Nowadays, European maize diversity represents around 75% of the American diversity (Tenaillon and Charcosset 2011).

Appendix 2: Expression of the phenotype of a hybrid between two unrelated populations

The following developments are based on Gallais (1989). Other expressions for GCA and SCA, based on the notation of Hayman (1954), can be found in Charcosset and Essioux (1994).

Let us consider a biallelic locus presenting the alleles B and b, for a hybrid individual issued from the cross between two gametes, one issued from population 1 carried the allele i and one issued from population 2 that carried allele j. The allele B is present in the population 1 at frequence p_1 and in the population 2 at frequence p_2. The allele b is present in the population 1 at frequence q_1 and in the population 2 at frequence q_2, with $p_1 + q_1 = 1$ and $p_2 + q_2 = 1$. The genotypic value associated to the hybrid can be written as: $y_{i_1 j_2} = \mu + \alpha_{i_1} + \alpha_{j_2} + \beta_{i_1 j_2}$ with α_{i_1} the additive effect of the allele i from the population 1, α_{j_2} the additive effect of the allele j from the population 2, $\beta_{i_1 j_2}$ the dominance deviation between the allele i from the population 1 and the allele j from the population 2, with $i = B \text{ or } b$; $j = B \text{ or } b$. $\alpha_{i_1} + \alpha_{j_2}$ corresponds to the additive genetic value of the considered hybrid.

The additive effect of one allele from one population corresponds to the average of the hybrid genotypes carrying the considered allele coming from the considered population. Additive effects can be expressed in function of a, d, and the frequencies of the alleles in the populations. Thus, we have for the hybrid population:

$\alpha_{B_1} = p_2 a + q_2 d - \mu$: additive effect of the allele B from population 1

$\alpha_{b_1} = p_2 d - q_2 a - \mu$: additive effect of the allele b from population 1

$\alpha_{B_2} = p_1 a + q_1 d - \mu$: additive effect of the allele B from population 2

$\alpha_{b_2} = p_1 d - q_1 a - \mu$: additive effect of the allele b from population 2

Substitution effects in each population can be defined as $\alpha_1 = \alpha_{B_1} - \alpha_{b_1}$ and $\alpha_2 = \alpha_{B_2} - \alpha_{b_2}$

Thus we have: $\alpha_{B_1} = q_1 \alpha_1$; $\alpha_{b_1} = -p_1 \alpha_1$; $\alpha_{B_2} = q_2 \alpha_2$; $\alpha_{b_2} = -p_2 \alpha_2$

The dominance deviations for the hybrid individual can be written as following, depending of its genotype: $\beta_{B_1 B_2} = -2 q_1 q_2 d$; $\beta_{b_1 b_2} = -2 p_1 p_2 d$; $\beta_{b_1 B_2} = 2 p_1 q_2 d$; $\beta_{B_1 b_2} = 2 q_1 p_2 d$.

We define the GCA of an individual from a population as the average of the value of its descendants when crossed to the other population. Thus, considering one locus, we have for the hybrid from the cross of a plant with genotype x from population 1 and a plant with genotype y from population 2: $Y_{(12)} = \mu + GCA_{x1} + GCA_{y2} + SCA_{xy(12)}$, with μ the average of the

hybrid population, GCA_{x1} the GCA of the genotype x from population 1 respectively to the population 2, GCA_{y2} the GCA of the genotype y from population 2 respectively to the population 1, $SCA_{xy(12)}$ the SCA between the two parental genotypes.

The GCA at one locus of an individual with genotype $x = i_1 j_1$ from population 1 relatively to population 2 is equal to $GCA_{x1} = \frac{1}{2}(\alpha_{i_1} + \alpha_{j_1})$. For an individual with genotype $y = i_2 j_2$ from population 2 relatively to population 1, it will be: $GCAy_2 = \frac{1}{2}(\alpha_{i_2} + \alpha_{j_2})$.

Thus, we can write the GCA at the locus level of the individual from population 1 depending of its genotype x: $GCA(BB)_1 = \alpha_{B_1}$; $GCA(Bb)_1 = \frac{1}{2}(\alpha_{B_1} + \alpha_{b_1})$; $GCA(bb)_1 = \alpha_{b_1}$. The GCA of an individual from the population 2 can be expressed in a similar way.

The general expression for SCA of the hybrid between the populations 1 and 2 at the gene level is: $SCAxy = \frac{1}{4}(\beta_{i_1 i_2} + \beta_{i_1 j_2} + \beta_{j_1 i_2} + \beta_{j_1 j_2})$

Then, SCA can be expressed depending of the genotype of the parental lines of the hybrid:

		Individual from population 1		
		BB	Bb	bb
Individual from population 2	BB	$\beta_{B_1B_2}$	$\frac{1}{2}(\beta_{B_1B_2} + \beta_{b_1B_2})$	$\beta_{b_1B_2}$
	Bb	$\frac{1}{2}(\beta_{B_1B_2} + \beta_{B_1b_2})$	$\frac{1}{4}(\beta_{B_1B_2} + \beta_{b_1B_2} + \beta_{B_1b_2} + \beta_{b_1b_2})$	$\frac{1}{2}(\beta_{b_1b_2} + \beta_{b_1B_2})$
	bb	$\beta_{B_1b_2}$	$\frac{1}{2}(\beta_{b_1b_2} + \beta_{B_1b_2})$	$\beta_{b_1b_2}$

These expressions at the locus level can be extended at the multilocus level. Thus, GCAs involve biological additive effects (a), biological dominant effects (d) and additive x additive epistatic effects (interaction between two non-homologous genes) (not shown here for the epistatic effects). SCA does not involve biological additive effect but involves biological dominant effect (d) and epistatic effects (not shown here for the epistatic effects). SCA involves the three types of epistasis: additive x additive (between two non-homologous genes), additive x dominance (between three genes: two homologous genes at one locus and one gene at the other locus) and dominance x dominance (between four genes: two homologous genes at one

locus and two homologous genes at the other locus) (not shown here for the epistatic effects). Allelic frequencies in the two populations have an impact on the GCAs and SCA values.

Literature cited

Akdemir, D., Sanchez, J. I.,and Jannink, J. L., 2015 Optimization of genomic selection training populations with a genetic algorithm. Genet. Sel. Evol. 47: 38.

Amuzu-Aweh, E. N., P. Bijma, B. P. Kinghorn, A. Vereijken, J. Visscher *et al.*, 2013 Prediction of heterosis using genome-wide SNP-marker data: application to egg production in white Leghorn crosses. Heredity 111: 530-538.

Barrière, Y., J. C. Emile, R. Traineau, F. Surault, M. Briand *et al.*, 2004 Genetic variation for organic matter and cell wall digestibility in silage maize. Lessons from a 34-year long experiment with sheep in digestibility crates. Maydica 49: 115-126.

Beavis, W. D., 1998 QTL analyses: power, precision and accuracy, pp. 145-162 in *Molecular dissection of complex traits*, edited by A. H. Paterson. CRC Press, New York.

Bernardo, R., 1994 Prediction of maize single-cross performance using RFLPs and information from related hybrids. Crop Sci. 34: 20–25.

Bernardo, R., 1996a Best linear unbiased prediction of the performance of crosses between untested maize inbreds. Crop Sci. 36: 50–56.

Bernardo, R., 1996b Best linear unbiased prediction of maize single-cross performance. Crop Sci. 36: 872–876.

Bernardo, R., and J. Yu, 2007 Prospects for genomewide selection for quantitative traits in maize. Crop Sci. 47: 1082-1090.

Blanc, G., A. Charcosset, J.-B. Veyrieras, A. Gallais, and L. Moreau 2008 Marker-assisted selection efficiency in multiple connected populations: a simulation study based on the results of a QTL detection experiment in maize. Euphytica 161: 71-84.

Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa, 2012 Genomic prediction of breeding values when modeling genotype x environment interaction using pedigree and dense molecular markers. Crop Sci. 52: 707-719.

Charcosset, A., and L. Essioux, 1994 The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. Theor. Appl. Genet. 89: 336-343.

Charcosset, A., 2002 Le fait hybride, conditions de l'innovation et choix stratégiques. "Colloque l'amélioration des plantes, continuités et ruptures", Montpellier, France, October 2002.

Comstock, R.E., H.F. Robinson, and P.H. Harvey, 1949 A breeding procedure designed to make maximum use of both general and specific combining ability. J. Am. Soc. Agron. 41:360–367.

Coors, J. G., 1999 Selection methodology and heterosis. In *The genetics and exploitation of heterosis in crops. Proceedings of an international symposium, CIMMYT, Mexico City, Mexico, 17—22 August 1997.*

Darvasi, A., and M. Soller, 1995 Advanced intercross lines, an experimental population for fine genetic mapping. Genetics 141: 1199-1207.

Darvasi, A., and M. Soller, 1997 A simple method to calculate resolving power and confidence interval of QTL map location. Behavior Genetics 27: 125-132.

Darwin, C., 1876 The effects of cross and self fertilization in the vegetable kingdom. John Murray, London.

Davenport, C. B., 1908 Degeneration, albinism and inbreeding. Science 28: 454-455.

Dubreuil, P., M. Warburton, M. Chastanet, D. Hoisington, and A. Charcosset, 2006 More on the introduction of temperate maize into Europe: large-scale bulk SSR genotyping and new historical elements. Maydica 51: 281-291.

East, E. M., 1908 Inbreeding in corn, pp.419-428 in *Reports of the Connecticut Agricultural Experimental Station for years 1907-1908*. Connecticut Agricultural Experiment Station, New Haven.

Eathington, S. R., T. M. Crosbie, M. D. Edwards, R. D. Reiter RD, and J. K. Bull, 2007 Molecular markers in a commercial breeding program. Crop Sci. 47: 154–163.

Ertl, J., A. Legarra, Z. G. Vitezica, L. Varona, C. Edel *et al.*, 2014 Genomic analysis of dominance effects on milk production and conformation traits in Fleckvieh cattle. Genet. Select. Evol. 46: 40.

Fu, D., M. Xiao, A. Hayward, Y. Fu, G. Liu *et al.*, 2014 Utilization of crop heterosis: a review. Euphytica 197: 161-173.

Gallais, A., 1989 *Théorie de la sélection en amélioration des plantes*. Ed. Masson, Paris

Gallais, A., 2009 *Hétérosis et variétés hybrides en amélioration des plantes*. Ed. Quae.

Goffinet B., and S. Gerber, 2000 Quantitative trait loci: a meta-analysis. Genetics 155: 463-473.

Habier, D., R. L. Fernando, and J. C. M. Dekkers, 2007 The impact of genetic relationship information on genome-assisted breeding values. Genetics 177: 2389–2397.

Hallauer, A. R., and S. A. Eberhart, 1970 Reciprocal full-sib selection. Crop Sci. 10: 315–316.

Hallauer, A. R., M. J. Carena, and J. B. Miranda Filho, 2010 *Quantitative genetics in maize breeding.* Handbook of plant breeding, 6. Springer.

Hayes, H. K., and R. J. Garber, 1919 Synthetic production of high protein corn in relation to breeding. J. Am. Soc. Agron. 11: 308–18.

Hayman, B. I., 1954 The theory and analysis of diallel crosses. Genetics 39: 789-809.

Heffner, E. L., M. E. Sorrells, and J. L. Jannink, 2009 Genomic selection for crop improvement. Crop Sci. 49: 1-12.

Heslot, N., H. P. Yang, M. E. Sorrells, and J. L. Jannink, 2012 Genomic selection in plant breeding: A comparison of models. Crop Sci. 52: 146-160.

Heslot, N., Akdemir, D., Sorrells, M. E., and Jannink, J.L., 2014 Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Genet. 127: 463–480.

Hollick, J. B., and V. L. Chandler, 1998 Epigenetic allelic states of a maize transcriptional regulatory locus exhibit overdominant gene action. Genetics 150: 891-897.

Hospital, F., L. Moreau, F. Lacoudre, A. Charcosset, and A. Gallais, 1997 More on the efficiency of marker-assisted selection. Theor. Appl. Genet. 95: 1181-1189.

Huang, B. E., K. L. Verbyla, A. P. Verbyla, C. Raghavan, V. K. Singh *et al.*, 2015 MAGIC populations in crops: current status and future prospects. Theor. Appl. Genet. 128: 999-1017.

Huang, Y. F., D. Madur, V. Combes, C. L. Ky, D. Coubriche *et al.*, 2010 The genetic architecture of grain yield and related traits in *Zea maize* L. revealed by comparing intermated and conventional populations. Genetics 186: 395-404.

Hufford, M. B., X. Xu, J. van Heerwaarden, T. Pyhäjärvi, J-M. Chia, *et al.*, 2012 Comparative population genomics of maize domestication and improvement. Nat. Genet. 44: 808-811.

Hull, F. G., 1945 Recurrent selection for specific combining ability in corn. J. Am. Soc. Agron. 37: 134–145.

Hull, F. H., 1946 Overdominance and corn breeding where hybrid seed is not feasible. J. Am. Soc. Agron. 38: 1100-1103.

Jenkins, M. T., 1929 Correlation studies with inbred and crossbred strains in maize. J. Agric. Res. 39: 677-722.

Jenkins, M. T., and A. M. Brunson, 1932 Methods of testing inbred lines of maize in crossbred combinations. J. Am. Soc. Agron. 24: 523-530.

Jenkins, M. T., 1940 The segregation of genes affecting yield of grain in maize. J. Am. Soc. Agron. 32:55-63.

Jinks, J. L., and R. M. Jones, 1958 Estimation of the components of heterosis. Genetics 43: 223-234.

Jones, D. F., 1917 Dominance of linked factors as a means of accounting for heterosis. Genetics 2: 466-479.

Jones, D. F., 1918 The effect of inbreeding and crossbreeding upon development. Proc. Natl. Acad. Sci. U.S.A. 4: 246-250.

Koelreuter, J.G., 1766 *Vorlaüfigen Nachrcht von einigen das Geschlecht der Pflanzen betreffenden Versuchen und Beobachtungen* (cited *in* Gowen, J. W., 1952 *Heterosis*. Iowa State College Press.)

Lamkey, C. M., and A. J. Lorenz, 2014 Relative effect of drift and selection diverging populations within a reciprocal recurrent selection program. Crop Sci. 54: 576-585.

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

Lehermeier, C., N. Krämer, E. Bauer, C. Bauland, C. Camisan *et al.*, 2014 Usefulness of multiparental populations of maize (*Zea mays* L.) for genome-based prediction. Genetics 198: 3-16.

Lonnquist, J. H., 1957 Project report on recurrent and reciprocal selection in corn. National poultry breeders roundtable proceedings, Part II, section I.

Maenhout, S., B. De Baets, G. Haesaert, and E. Van Bockstaele, 2007. Support vector machine regression for the prediction of maize hybrid performance. Theor. Appl. Genet. 115: 1003-1013.

Maenhout, S., B. De Baets, and G. Haesaert, 2010 Prediction of maize single-cross hybrid performance: support vector machine regression versus best linear prediction. Theor.

Appl. Genet. 120: 415-427.

Maher, B., 2008 The case of the missing heritability. Nature 456: 18-21.

Massman, J. M., A. Gordillo, R. E. Lorenzana, and R. Bernardo, 2013 Genomewide predictions from maize single-cross data. Theor. Appl. Genet. 126: 13-22.

Matsuoka, Y., Y. Vigouroux, M. M. Goodman, G. J. Sanchez, E. Buckler *et al.*, 2002 A single domestication for maize shown by multilocus microsatellite genotyping. Proc. Natl. Acad. Sci. U.S.A. 99: 6080–6084.

Melchinger, A. E., and R. K. Gumber, 1998 Overview of heterosis and heterotic groups in agronomic crops. pp. 29-44. *In* Lamkey, K.R., and J.E. Staub (Eds.), Concepts and Breeding of Heterosis in Crop Plants. CSSA, Madison, WI.

Melchinger, A. E., H. F. Utz, and C. C. Schön, 1998 QTL mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149: 383–403.

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

Mir, C., T. Zerjal, V. Combes, F. Dumas, D. Madur *et al.*, 2013 Out of America: tracing the genetic footprints of the global diffusion of maize. Theor. Appl. Genet. 126: 2671-2682.

Moll, R. H., J. H. Lonnquist, J. V. Fortuna, and C. E. Johnson, 1965 The relationship of heterosis and genetic divergence in maize. Genetics 52: 139-144.

Moreau, L., A. Charcosset, F. Hospital, and A. Gallais, 1998 Marker-assisted selection efficiency in populations of finite size. Genetics 148: 1353-1365.

Moreau, L., A. Charcosset, and A. Gallais, 2004 Experimental evaluation of several cycles of marker-assisted selection in maize. Euphytica 137: 111-118.

Moreau, L., and A. Charcosset, 2011 Marker-assisted selection in maize. In *Advance in Maize* Ed. J. L. Prioul, Society for Experimental Biology, UK.

Morrow, G. E., and F. D. Gardner, 1893. Field experiments with corn 1892. Illinois Agricultural Experimental Station Bulletin 25: 173-203.

Neimann-Sorensen, A., and A. Robertson, 1961 The associations between blood groups and several production characteristics in three Danish cattle breeds. Acta. Agric. Scand. 11: 163-196.

Paterniani, E., 1967 Interpopulation improvement: Reciprocal recurrent selection variations. Maize 8, CIMMYT.

Paterson, A. H., E. S. Lander, J. D. Hewitt, S. Peterson, S. E. Lincoln *et al.*, 1988 Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335: 721-726.

Piperno, D. R., A. J. Ranere, I. Holst, J. Iriarte, and R. Dickau, 2009 Starch grain and phytolith evidence for early ninth millennium B. P. maize from the Central Balsas River Valley, Mexico. Proc. Natl. Acad. Sci. U.S.A. 106: 5019-5024.

Powers, L., 1944 An expansion of Jones's theory for the explanation of heterosis. Am. Nat. 78: 275-280.

Pritchard, J. K., and M. Przeworski, 2001 Linkage disequilibrium in humans: models and data. Am. J. Hum. Genet. 69: 1-14.

Rebaï, A., P. Blanchard, D. Perret, and P. Vincourt, 1997 Mapping quantitative trait loci controlling silking date in a diallel cross among four lines of maize. Theor. Appl. Genet. 95: 451-459.

Rebourg, C., B. Gouesnard, and A. Charcosset, 2001 Large scale molecular analysis of traditional European maize populations. Relationships with morphological variation. Heredity 86: 574-587.

Rebourg, C., M. Chastanet, B. Gouesnard, C. Welcker, P. Dubreuil *et al.*, 2003 Maize introduction into Europe: the history reviewed in the light of molecular data. Theor. Appl. Genet. 106: 895–903.

Reif, J., A. Hallauer and A. Melchinger, 2005 Heterosis and heterotic patterns in maize. Maydica 50: 215-223.

Reif, J. C., Y. Zhao, T. Würschum, M. Gowda, and V. Hahn, 2013 Genomic prediction of sunflower hybrid performance. Plant Breeding 132: 107-114.

Richey, F. D., 1924 Effects of selection on the yield of a cross between varieties of corn. U.S.D.A. Bul. 1209.

Richey, F. D., and L. S. Mayer, 1925. The productiveness of successive generations of self-fertilized lines of corn and of crosses between them. USDA Bull. 1354.

Richey, F. D., 1942 Mock-dominance and hybrid vigor. Science 96: 280-281.

Riedelsheimer, C., J. B. Endelman, M. Stange, M. E. Sorrells, J.-L. Jannink *et al.*, 2013

Genomic predictability of interconnected bi-parental maize populations. Genetics 194: 493–503.

Rincent, R., D. Laloë, S. Nicolas, T. Altmann, D. Brunel *et al.*, 2012 Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). Genetics 192: 715-28.

Shull, G. H., 1908 The composition of a field of maize. Am. Breed. Assoc. Rep. 5: 51-59.

Shull, G. H., 1914 Duplicate genes for capsule-form in *Bursa bursa pastoris*. Zeitschrift ind. Abst. U. Verebsgl. 12: 97-149.

Sprague, G. F., 1939 An estimation of the number of top-crossed plants required for adequate representation of a corn variety. J. Am. Soc. Agron. 31: 11-16.

Sprague, G. F., and L. A. Tatum, 1942 General *vs.* specific combining ability in single crosses of corn. J. Am. Soc. Agron. 34: 923-932.

Springer, N. M., and R. M. Stupar, 2007 Allelic variation and heterosis in maize: How do two halves make more than a whole? Genome Res. 17: 264-275.

Sun, C., P. M. VanRaden, J. B. Cole, and J. R. O'Connell 2014 Improvement of prediction ability for genomic selection of dairy cattle by including dominance effects. PLoS One 9: e103934.

Surault, F., J. C. Emile, M. Briand, Y. Barrière, and R. Traineau, 2005 Variabilité génétique de la digestibilité *in vivo* d'hybrides de maïs. Bilan de 34 années de mesures. Fourrages 183: 459-474.

Technow, F., C. Riedelsheimer, T. A. Schrag, and A. E. Melchinger, 2012 Genomic prediction of hybrid performance in maize with models incorporating dominance and population specific marker effects. Theor. Appl. Genet. 125: 1181-1194.

Technow, F., A. Bürger, and A. E. Melchinger, 2013 Genomic prediction of Northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. G3 (Bethesda) 3: 197–203.

Technow, F., T. A. Schrag, W. Schipprack, E. Bauer, H. Simianer *et al.*, 2014 Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. Genetics 197: 1343-1355.

Tenaillon, M. I., and A. Charcosset, 2011 A European perspective on maize history. C. R. Biol. 334: 221–228.

Toro, M. A., and L. Varona, 2010 A note on mate allocation for dominance handling in genomic selection. Genet. Select. Evol. 42: 33.

Troyer, A. F., 1999 Background of U.S. hybrid corn. Crop Sci. 39: 601-626.

Troyer, A. F., and L. G. Hendrickson, 2007 Background and importance of "Minnesota 13" corn. Crop. Sci. 47: 905-914.

Truntzler, M., Y. Barrière, M. C. Sawkins, D. Lespinasse, J. Bertran *et al.*, 2010 Meta-analysis of QTL involved in silage quality of maize and comparison with the position of candidate genes. Theor. Appl. Genet. 121: 1465-1482.

van Heerwaarden, J., J. Doebley, W. H. Briggs, J. C. Glaubitz JC, M. M. Goodman *et al.*, 2011 Genetic signals of origin, spread, and introgression in a large sample of maize landraces. Proc. Natl. Acad. Sci. U.S.A. 108: 1088–1092.

Varona, L., and M. A. Toro, 2012 Non-additive genetic effects in genomic evaluation. Wroclaw, Poland.

Wallace, B. E., 1968 Selection in favor of heterozygotes. pp. 203-216 in *Topics in Population Genetics*. Norton, New York.

Whittaker, J. C., R. Thompson, and M. C. Denham, 2000 Marker-assisted selection using ridge-regression. Genet. Res. Camb. 75: 249-252.

Williams, W., 1959 Heterosis and the genetics of complex characters. Nature 184: 527-530.

Xu, S., D. Zhu, and Q. Zhang, 2014 Predicting hybrid performance in rice using genomic best linear unbiased prediction. Proc. Natl. Acad. Sci. U. S. A. 111: 12456-12461.

Yu, J., G. Pressoir, W. H. Briggs, I. Vroh Bi, M. Yamasaki *et al.*, 2006 A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38: 203-208.

Yu, J., J. B. Holland, M. D. McMullen, and E. S. Buckler, 2008 Genetic design and statistical power of nested Association mapping in maize. Genetics 178: 539-551.

Zhao, Y., J. Zeng, R. Fernando, and J. C. Reif, 2013 Genomic prediction of hybrid wheat performance. Crop Sci. 53: 802-810.

Chapter 1

Linkage disequilibrium with linkage analysis of multi-line crosses reveals different multi-allelic QTL for hybrid performance in the flint and dent heterotic groups of maize

Héloïse Giraud[*], Christina Lehermeier[†], Eva Bauer[†], Matthieu Falque[§], Vincent Segura[‡], Cyril Bauland[§], Christian Camisan**, Laura Campo[§§], Nina Meyer[††], Nicolas Ranc[‡‡], Wolfgang Schipprack***, Pascal Flament**, Albrecht E. Melchinger***, Monica Menz[‡‡], Jesús Moreno-González[§§], Milena Ouzunova[††], Alain Charcosset[§], Chris-Carolin Schön[†], Laurence Moreau[§1]

*Univ Paris-Sud, UMR 0320 / UMR 8120 Génétique Végétale, F-91190 Gif-sur-Yvette, France

[†]Plant Breeding, Technische Universität München, D-85354 Freising, Germany

[§]INRA, UMR 0320 / UMR 8120 Génétique Végétale, F-91190 Gif-sur-Yvette, France

[‡]INRA, UR0588, F-45075 Orléans, France

**Limagrain Europe, F-63720 Chappes, France

[§§]Centro Investigacións Agrarias Mabegondo (CIAM), 15080 La Coruña, Spain

[††]KWS SAAT AG, D-37555 Einbeck, Germany

[‡‡]Syngenta Seeds S.A.S., F-31790 Saint-Sauveur, France

***Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, D-70593 Stuttgart, Germany

[1]Corresponding author

Running title: QTL detection in multi-parental families

Key words: QTL detection; LDLA analysis; allelic series; multi-parental families; maize biomass production

Corresponding author:
Dr. Laurence Moreau
INRA, UMR 0320 / UMR 8120 Génétique Végétale
F-91190
Gif-sur-Yvette
France
Phone: +33 1 69 33 23 37
Fax: +33 1 69 33 23 40
Email: moreau@moulon.inra.fr

ABSTRACT

Multi-parental designs combined with dense genotyping of parents have been proposed as a way to increase the diversity and resolution of quantitative trait loci (QTL) mapping studies, using methods combining linkage disequilibrium information with linkage analysis (LDLA). Two new Nested Association Mapping designs adapted to European conditions were derived from the complementary dent and flint heterotic groups of maize (*Zea mays* L.). Ten biparental dent families (N=841) and 11 biparental flint families (N=811) were genotyped with 56,110 single nucleotide polymorphism markers and evaluated as testcrosses with the central line of the reciprocal design for biomass yield, plant height and precocity. Alleles at candidate QTL were defined as (i) parental alleles, (ii) haplotypic identity by descent and (iii) single marker groupings. Between five and 16 QTL were detected depending on the model, trait and genetic group considered. In the flint design, a major QTL (R^2=27%) with pleiotropic effects was detected on chromosome 10, whereas other QTL displayed milder effects (R^2<10%). On average, the LDLA models detected more QTL but generally explained lower percentages of variance, consistent with the fact that most QTL display complex allelic series. Only 15% of the QTL were common to the two designs. A joint analysis of the two designs detected between 15 and 21 QTL for the five traits. Out of these, between 27 for silking date to 41% for tasseling date were significant in both groups. Favorable allelic effects detected in both groups open perspectives for improving biomass production.

INTRODUCTION

Most traits of agronomic interest present a continuous variation resulting from the sum of the effects of various quantitative trait loci (QTL). Mapping these QTL is a first step towards elucidating their molecular nature and offers important application perspectives for marker-assisted breeding. QTL mapping started in plants with segregating families derived from the cross of two inbred lines (Lander and Botstein 1989). However, such biparental designs address only a small portion of the diversity available (a maximum of two alleles can segregate at a given QTL) and the accuracy of QTL positions is usually poor. To overcome these limitations, Rebai and Goffinet (1993) and Charcosset *et al.* (1994) proposed models for joint QTL detection in several biparental families connected to each other by the use of common parental lines. When the number of parents is less than the number of families, connections can be taken into account in the detection model to reduce the number of allelic effects to estimate. This increases power and accuracy of detection when QTL behave additively (see Blanc *et al.* 2006). However, such a model makes the assumption that each parental line carries a different allele, which limits its benefit when the number of parental lines is high relative to the number of families, a situation commonly encountered in breeding programs.

Recent advances in sequencing and genotyping technologies make it possible to genotype individuals for a large number of markers at reduced costs, so that one can expect to have markers closely linked to any QTL. This has paved the way towards association mapping, in which marker-trait associations are directly detected in populations composed of diverse inbred lines without the need to develop experimental segregating families. Association mapping, also often referred to as Linkage Disequilibrium (LD) mapping, has been widely used with success in the plant community (see for instance Bouchet *et al.* 2013 and Romay *et al.* 2013 for recent results of association mapping in maize). In this approach, it is important to use models accounting for potential underlying population structure and relatedness between individuals to prevent spurious QTL detection due to associations between loci that are not linked physically (Yu *et al.* 2006). As a consequence, the power to detect associations is low for causal polymorphisms correlated with the underlying population structure or when they are present in the population at a low frequency (Rincent *et al.* 2014). In addition, associations are generally tested at SNP (Single Nucleotide Polymorphism) markers which leads to the implicit assumption that the QTL are biallelic. These limitations can be alleviated by combining information coming from LD at the level of the parents and linkage within families, as first proposed for animal populations by Meuwissen and Goddard (2001). In this approach, referred

to as Linkage Disequilibrium and Linkage Analysis (LDLA), dense genotyping of parents is used to detect identity by descent (IBD) at putative QTL, *i.e.* the fact that two individuals carry the same allele transmitted by a common ancestor. Different types of LDLA analyses have been proposed to account for the LD component. The simplest is to consider that parents carrying the same allele at a given marker are IBD (Yu *et al.* 2008; Liu *et al.* 2012) as done in association mapping. Haplotype based approaches also have been proposed to group parental alleles and tested by simulations (for instance Jansen *et al.* 2003; Bink *et al.* 2012; Leroux *et al.* 2014). Advantages of LDLA have been shown experimentally in maize notably by using the nested association mapping (NAM) design developed in the USA (Yu *et al.* 2008; McMullen *et al.* 2009). This design consists of 25 biparental recombinant inbred line (RIL) populations derived from the cross of the inbred B73 with 25 diverse lines representing the diversity of maize (tropical, temperate, sweet corn, and popcorn lines). This design was studied with a linkage analysis model (Buckler *et al.* 2009; Kump *et al.* 2011; Tian *et al.* 2011) where QTL effects were nested within each family and each parental line was assumed to carry a different allele, and with LDLA through a genome-wide association mapping model (Kump *et al.* 2011; Tian *et al.* 2011) including allelic effects observed at individual SNP of the parents to identify IBD alleles. This design successfully led to the detection of numerous QTL and use of LDLA permitted in some cases to resolve QTL detection up to the gene level (Kump *et al.* 2011; Poland *et al.* 2011; Tian *et al.* 2011; Cook *et al.* 2012). Recently, Bardol *et al.* (2013) applied the haplotype-based approach of Leroux *et al.* (2014) to detect QTL in two datasets coming from an applied maize (*Zea mays L.*) breeding program and compared it to models considering each parental allele as different (linkage model) or considering that parents carrying the same allele at a given marker are IBD. Results showed that when parental lines are all issued from the same breeding program and related by pedigree, LDLA models were more powerful than linkage approaches. Bardol *et al.* (2013) also showed that the different ways of modeling allelic variation (either using haplotypes or single marker information) had variable efficiencies depending on the QTL and trait considered and were therefore complementary. It is thus important to further evaluate the ability of diverse LDLA models to detect QTL in multi-parental populations with different diversity levels.

The central line of the US NAM (B73) is too late flowering for evaluation in Northern Europe and founder lines cover a very broad range of geographical origins, including even later tropical materials. This prevents the evaluation of the whole design for productivity traits in Northern European conditions and due to diversity of the lines it is difficult to use a single tester to investigate hybrid values. To overcome these limitations and expand the genetic pool

investigated in maize QTL mapping studies, two parallel complementary NAM designs were developed within the European project CornFed. Each was derived from inbred lines representing the main diversity available for breeding in each of the two major heterotic groups (dent and flint) used in Northern Europe. Both designs were genotyped with a 50k SNP array (Ganal *et al.* 2011) and genotyping information was used to build individual population maps (Bauer *et al.* 2013). The two NAM designs were crossed with the central line of the opposite group to produce hybrids, that were analyzed for traits related to biomass production as described in Lehermeier *et al.* (in press). Increasing biomass production is of key interest in Northern Europe where maize has been extensively used for decades for silage and more recently for bioenergy production. To our knowledge no QTL mapping experiment has been carried out so far for traits related to biomass production in multi-parental design assembling such large diversity. Note that both hybrid designs address variation compared to the same hypothetical reference hybrid (the one produced by crossing the two central lines), with each experimental hybrid of each group sharing on average 75% of its genome with the reference hybrid. In this context, effects of all segregating genotypes at a QTL (11 on the dent side and 12 on the flint side) are compared to a same genotype (having received alleles from two central lines). This makes this design particularly adapted for deciphering loci involved in genetic variation on the dent and flint sides for productivity traits.

The present study aimed at comparing different methods of QTL detection in these two European NAM designs for five traits of agronomical interest for biomass production in maize: whole plant dry matter yield, whole plant dry matter content at harvest, female flowering, male flowering and plant height. We compared a linkage approach with two LDLA approaches either considering haplotypic IBD or single marker groupings. This allowed us to investigate the performance of the different LDLA approaches in two complementary heterotic groups in a more diverse context than a simple breeding program. A second important objective of this work was to compare the results of QTL detection conducted separately in the two heterotic groups or jointly for the whole design, in order to better understand the contribution of each group to trait variation..

MATERIAL AND METHODS

Plant material and phenotypic analysis

Two maize NAM designs composed of half-sib families from the two major heterotic groups (dent and flint) used for breeding in Europe were analyzed. The two designs are described in Bauer *et al.* (2013). In short, the dent and flint designs were respectively composed of 10 and 11 doubled haploid (DH) families, derived from the cross of respectively 10 and 11 diverse founder lines with a common central line: F353 for the dent and UH007 for the flint. F353 and UH007 represent very promising European lines created by public institutes in their respective heterotic groups. The parental lines were chosen to cover the diversity available within the two groups with a combination of ancestral and more recent material. From each cross, DH lines were generated resulting in 919 lines for the dent and 1009 for the flint (Bauer *et al.* 2013) (Table S1). For phenotypic evaluation (see below), the segregating DH lines of a given group were crossed with the central line of the other group. 841 hybrids were produced for the dent group and 811 for the flint group (Lehermeier *et al*. in press) (Table S1). The number of dent lines for which test-crossed progenies were phenotyped per family was 84 on average and varied between 53 and 104, depending on the family. For the flint group, the number of DH lines per family that were phenotyped for test-cross values ranged from 17 to 133 with an average of 73. As the hybrids of each group were obtained by crossing DH lines with the central line of the other group, all the hybrids shared a large proportion of their genome and were expected to be heterozygotes F353 / UH007 for 50% of their genome. Hybrids were evaluated in 2011 in four (dent) and six (flint) European locations. Five traits were considered: biomass dry matter yield (DMY, decitons per hectare, $dt.ha^{-1}$) at the whole plant level, whole plant dry matter content (DMC, %) at harvest, days to tasseling (DtTAS, in days, d), and days to silking (DtSILK, in days, d) measured as the number of days from sowing until tasseling and silking, respectively. Field trial design is described in Lehermeier *et al.* (in press. Individual field plot measures were analyzed (Lehermeier *et al.* in press) to compute for each hybrid the adjusted means over the different trials that were used in this study.

Genotyping and analysis of genotypic data

The 1,928 DH lines and the 23 parental lines were genotyped with the Illumina® MaizeSNP50 BeadChip containing 56,110 SNPs (Ganal *et al.* 2011a). Markers with a call frequency < 0.9, a GenTrainScore < 0.7, a minor allele frequency (MAF) < 0.01, or more than 10% missing values were discarded as in Lehermeier *et al.* (in press).

Consensus maps for the flint and the dent multi-populations were obtained following the same procedure. We considered for each consensus map the list of markers present in at least one of the 10 dent individual maps (respectively 11 flint individual maps) from Bauer *et al.* (2013). The flint DH family resulting from the cross of EP44 and UH007 was not used due to small population size. For each marker of this list and for each individual genetic map, we computed the relative genetic position of this marker in this map by starting from its physical coordinate on the B73 genome assembly and converting it into a genetic coordinate with the spline smoothing interpolating procedure described in Bauer *et al.* (2013). These genetic coordinates were then normalized between zero and one to obtain relative genetic positions. For the present study, each consensus map was built by computing the consensus relative genetic position of each marker as the average of its relative genetic positions in all individual maps involved, weighted by the numbers of individuals in the corresponding populations. Finally, the consensus genetic coordinate of each marker was obtained by multiplying its consensus relative genetic position by the genetic length of the consensus map, taken as the average of the genetic lengths of all maps, weighted by the numbers of individuals in the corresponding populations. The two consensus maps obtained are available at Maize GDB (MaizeGDB data record). A consensus map for the dent and flint multi-populations was built with the same procedure.

For the QTL detection we only considered in the analysis the PANZEA markers which were mapped on the consensus maps. PANZEA markers result from the alignment of sequences coming from resequencing data of the 27 lines used as parents of the US NAM design *(*McMullen *et al.* 2009) and mapped against the B73 genome v2 (Gore *et al.* 2009). We discarded the other markers, mainly defined by comparing the sequences of the inbred lines B73 and Mo17, as they are known to create an ascertainment bias in diversity analyses (Ganal *et al.* 2011; Frascaroli *et al.* 2013).The dent and flint consensus genetics maps obtained were composed of respectively 21,878 and 20,406 PANZEA markers, corresponding respectively to 6,808 and 7,272 genetic positions on the consensus maps. The dent-flint consensus map was composed of 25,472 PANZEA markers, corresponding to 8,124 genetic positions (Table 1).

Clustering analysis of parental inbred lines

Clustering of the parental inbred lines was carried out with the R package “clusthaplo” (Leroux *et al.* 2014), separately on the dent and flint parents. This clustering was based on genomic similarities computed between each pair of individuals in a sliding window along the genome. In order to get insight into the length of the sliding window to use, we evaluated how fast LD between pairs of markers decays with the genetic distance. LD between pairs of markers was

estimated for the 11 dent founder lines and for the 12 flint founder lines, according to Hill and Robertson (1968) as $r^2 = D_{AB}^2 / (p_A(1-p_A)p_B(1-p_B))$, with $D_{AB} = p_{AB} - p_A p_B$ where p_{AB} denotes the haplotype frequency of AB, p_A the frequency of allele A at one marker locus, and p_B the frequency of allele B at the other locus. The LD decay was estimated using the Hill and Weir model (1988). The choice of the sliding window size was based on the LD decay observed in the dent and flint material considering the length in genetic distance needed to reach an r^2 below 0.2. Two values were chosen, 2 cM and 5 cM, each based on the LD decay observed for the flint and dent group, respectively. For facilitating comparisons between results obtained in the two groups, the clustering was carried out in each group using the two window sizes.

For each window size at each genotyped position, the similarity score between two parental lines i and j at a position t (center of the window) was calculated according to the formula described in Leroux *et al.* (2014) and used in Bardol *et al.* (2013). This formula is adapted from Li and Jiang (2005) and combined the number of alleles alike-in-state between the two lines inside the sliding window and the length of their longest common segment centered on t. Based on the similarity score curves obtained along each chromosome, a hidden Markov model (HMM) was used to determine at each position *t* if the two lines were similar and thus carried the same ancestral allele or not. After the clustering process, the number of ancestral alleles per position was plotted along chromosomes. We also computed similarities between inbred lines as the percentage of ancestral alleles shared over the genome and compared them with the similarities obtained from the SNP markers. A graphical representation of these similarities and a classification of the parental lines were carried out using the "heatmap" function in R (R Core Team 2013).

QTL detection

Analyses were first performed separately for each trait on the dent and flint multi-family designs, using their respective consensus map. Four statistical models were tested: one based on linkage analysis and three others combining linkage and LD information. All the models were multi-locus models in which the significance of each QTL was tested conditional on the inclusion of other QTL positions used as cofactors.

The first model corresponded to a conventional multi-family connected model. This model considered the connections between families through the sharing of the central inbred line and relied on the assumptions that each parental inbred line carried a different QTL allele and that each allelic effect was independent of the family:

$$y = J\,.\mu + X_q\,.\,a_q + \sum_{c \neq q} X_c\,.a_c + e$$

where y was the vector $(N \times 1)$ of the adjusted phenotypic means of the N individuals of the dataset, J was a $(N \times P)$ matrix of 0 and 1 that linked each individual to the family it belonged to with P being the total number of families, μ was the column vector $(P \times 1)$ of family means, X_q and X_c were $(N \times K)$ matrices with K being the number of parents. Each element (ranging from 0 to 2) of these matrices corresponded to the expected number of alleles of the parent k at QTL q and cofactor c for each individual, according to the genotyping information at the position of q and c when this information was available (*i.e.* when these positions correspond to markers polymorphic in the population the individual belong to) or at flanking markers otherwise. a_q and a_c were the column vectors $(K \times 1)$ of the additive intra-family effects associated with QTL q and cofactor c, respectively. e was a column vector $(N \times 1)$ of the residuals of the model. This model will be further referred to as "connected". Note that this model is close to the "joint inclusive composite interval mapping" (JCIM) model proposed by Buckler *et al.* (2009) and used on the US NAM design.

The second and third models were LDLA multi-family connected models which used the results of the clustering of parental alleles carried out with "clusthaplo":

$$y = J\,.\mu + \ X_q\,.\,Q_q\,.\,h_q + \sum_{c \neq q} X_c\,.Q_c\,.\,h_c + e$$

where y, J, μ, X_q, X_c and e were the same as described as in the previous model. Q_q and Q_c were $(K \times A_q)$ and $(K \times A_c)$ matrices with A_q and A_c being the number of ancestral alleles at QTL q and cofactor c. Each element (0 or 1) of these matrices linked the parental alleles at QTL q and cofactor c to the ancestral alleles identified by the clustering approach. h_q and h_c were column vectors $(A_q \times 1)$ and $(A_c \times 1)$ of the additive effects of the ancestral alleles associated with QTL q and cofactor c. Two models were considered, one based on the clustering approach using a window size of 2cM and further referred to as "LDLA – 2 cM", and one based on the clustering approach using a window size of 5 cM and further referred to as "LDLA – 5 cM".

QTL detection using the three models described above were performed using the MCQTL_LD software (Jourjon *et al.* 2005) using an iterative composite interval QTL mapping method (iQTLm) (Charcosset *et al.* 2000). For these models, genotypic information of markers located

at the same position of the consensus genetic map, was concatenated to indicate which parental allele was transmitted. For missing data, MCQTL_LD software estimated the probability of transmission of each parental allele based on the information of flanking markers. At each tested position, the presence of a QTL was assessed based on the - log10 of the Fisher test p-value (-log10(p-value)). Thresholds for considering a QTL as significant were computed for each trait and each dataset using 5,000 intra-family permutations of the phenotypes for a type I risk of 10% across all families and total genome. In the iQTLm approach, the initial set of cofactors was chosen using a multiple regression with a forward selection of marker positions with a threshold equal to 80% of the QTL significance threshold value. At the end of the detection process, for the conventional connected model, confidence intervals at 95% were estimated on the basis of a 1 LOD unit fall. The confidence intervals were not estimated for the LDLA models as there is no established method proposed for these models.

The fourth model, referred to as single marker LDLA model ("LDLA – 1-marker"), considered that two parental lines carrying the same allele at a marker were IBD for this marker:

$$y = J.\mu + M_q . g_q + \sum_{c \neq q} M_c . g_c + e$$

y, J, μ and e were as described in the previous model. M_q and M_c were $(N \times 2)$ matrices whose elements (0 or 1) corresponded to the genotyping information at QTL q and cofactor c for each individual. g_q and g_c were column vectors (2×1) of the additive effects of marker alleles associated with QTL q and cofactor c. This model can be viewed as a multi-locus genome-wide association study with population structure controlled by family membership. It is equivalent to the association mapping model used to analyze the US NAM design (Yu *et al.* 2008; Tian *et al.* 2011; Kump *et al.* 2011) except that in our model dense marker genotyping information is directly available for the progenies and does not need to be inferred from the parental genotypes.

The analysis with the fourth model was performed in R (R Core Team 2013) using an R-script derived from the one used for the multi-locus mixed model approach presented in Segura *et al.* (2012). We used a multi-locus forward-backward stepwise linear regression model and selected the most appropriate model using the extended Bayesian information criterion (Segura *et al.* 2012). Loci of the selected model which had p-values below the Bonferroni threshold for a genome-wide risk of 10% were considered as QTL. For this model, imputation of the genotyping data for marker with missing data was done using the software BEAGLE (Browning

and Browning 2009) family by family. Even if we considered the same type I error risk at the genome level than for other models, the threshold used for the LDLA – 1-marker model was not obtained by permutations and is possibly more conservative compared to other models.

Analyses were then performed jointly for each trait on the two designs using the dent-flint consensus map. The model used corresponded to a conventional multi-family connected model except that all the dent and flint families were considered jointly. As the central line of the dent is used as tester in the flint design and reciprocally, the F353-UH007 genotype segregates against an alternative genotype in each population. This enabled us to connect allelic effects estimated in the two designs. QTL detection was performed using the MCQTL_LD software (Jourjon *et al.* 2005) following the same procedure as group specific QTL detection. Thresholds for considering a QTL as significant were computed for the joint dataset for each trait using 5,000 intra-family permutations of the phenotypes for a type I risk of 10% across all families and total genome. To test whether effects were significant in a single group or in both groups, the effects of the QTL detected in the joint analysis were tested in each of the separate datasets. They were considered as significant if the - log10 of the Fisher test p-value was above the thresholds of the studied trait in the separate dataset (estimated with the dent or flint consensus maps, respectively).

For each analysis, variances explained by each QTL (partial R^2_{QTL}) were defined as the ratio between the sum of squares associated with the QTL effect in the model including the other detected QTL, and the residual sums of squares of a linear model considering only the family effects. Total percentage of variance explained by the detected QTL (R^2_{total}) was defined as the ratio between the sum of squares of all the detected QTL, and the residual sums of squares of a linear model considering only the effects of the families. All the R^2 were adjusted by the number of degrees of freedom of the considered models (Charcosset and Gallais 1996). Differences in effects among pairs of alleles at a given QTL was tested a posteriori using a t-test (α=5%). For facilitating comparisons between models and the interpretation of the QTL results, the allelic effect of the central lines were set to zero and the other allelic effects were estimated accordingly.

Comparison of the positions of the QTL detected separately in the two groups and in the joint analysis was based on the results of the connected model. QTL detected in each separate group

and on the joint dataset were projected on the dent-flint consensus map using BioMercator V4.2 (Sosnowski *et al.* 2012). A QTL was considered common for a trait when the confidence intervals of the QTL after projection were overlapping.

RESULTS

Analysis of parental linkage disequilibrium and parental clustering

The average genetic distance to reach a LD below r^2 =0.2 was 1.2 and 0.65 cM for the dent and flint groups, respectively (Table 1). This distance varied according to the chromosome between 0.45 cM (chromosome 5) and 2.51 cM (chromosome 2) for the dent group, and 0.35 cM (chromosome 5) and 0.76 cM (chromosomes 1, 7, 9, 10) for the flint group. The two different sliding window sizes that we considered for computing the similarity score with "clusthaplo" approximately correspond to two times the distance beyond which LD becomes negligible for all the chromosomes. Note that 2 cM was the minimum window size that we could consider since the HMM based clustering approach did not converge for smaller window sizes.

The 5 cM sliding window size led to a higher number of ancestral alleles than the 2 cM one for the two designs. For dent, the average number of ancestral alleles along the genome was 5.6 per genetic position for the 2 cM sliding window size and 6.5 for the 5 cM window. For flint, the average number of ancestral alleles was 5.9 per genetic position for the 2 cM sliding window size and 7.2 for the 5 cM window. It has to be noted that the number of ancestral alleles varied along the genome. For both window sizes, clustering was more important in telomeric than in centromeric regions, where quite often the number of ancestral alleles equaled the number of parental lines (Figure 1).

For both sliding window sizes, similarities between the parental inbred lines estimated based on ancestral alleles sharing showed a structured pattern (Figure 2). Within the dent group, pairs of lines involving (i) UH250, D09 and D06 and (ii) F353 and UH304 shared the same ancestral alleles for more than 47% of the genetic positions for both sizes of sliding window. In the flint group, with the 5cM window, closest pairs of lines involved UH006, UH007 and UH009. With the 2 cM window size, this expanded to F03802, D152 and F2. The classifications of parental lines based on single markers were globally consistent with those based on ancestral alleles, at least for grouping the most similar lines. Only positions of inbred lines which showed low levels of similarities with the other lines slightly changed in the dendrogram depending on the allele definition considered. In the dent group, three related lines UH250, D09, and D06 are clearly

separated from a non structured group among which only F353 (the central line of the dent design) and UH304 were related. In the flint group, similarities separated a sub-group composed of F64, EC49A, EZ5, and EP44 from the other lines that appeared to be more closely related to each other. In this sub-group, UH009 and UH006 are both related to UH007, the central line of the flint design.

Comparison of the thresholds used in the QTL detection models

For the separate datasets analyses, threshold values (- log10(p-value)) were higher for the LDLA models than for the linkage model (Table S2). For LDLA models, the threshold increased as the size of the considered window decreased. This suggests that reducing the size of the window decreases the dependence between tests. For every model, threshold values were lower for DMC and higher for DtSILK and DtTAS (except for the conventional connected model for the flint group). This might be due to heterogeneity of within family variances for some traits. For instance, for DtSILK, for the dent dataset, genetic variances varied from 0.95 to 4.93 (see Lehermeier *et al.* in press for an estimation of these variances). Like for the separate datasets thresholds, for the joint dataset, threshold values for the connected model were lower for DMC and higher for DtSILK and DtTAS.

Comparison of the QTL detected with the different models in the dent and flint designs

For a given trait and group, the number of detected QTL varied according to the model (Table 2, Table S3, Table S4, Table S5, Table S6, Table S7, Table S8, Table S9 and Table S10). Between 5 (for DMY with LDLA – 5 cM and LDLA – 1-marker models) and 16 (for DMC with LDLA – 2cM model) QTL were detected in the dent design and between 7 (for DMC with LDLA – 1-marker model) and 16 QTL (for DtSILK and DtTAS with LDLA – 1-marker model) in the flint design.

For the dent group, the LDLA – 1-marker model detected fewer QTL over all traits (45 QTL in total) and explained the smaller percentage of variance (33.8% on average). In this group, the LDLA models using “clusthaplo” information detected more QTL (56 in total for the LDLA – 5 cM, 55 for the LDLA – 2 cM) than the conventional connected model (52 QTL in total). This advantage of the LDLA models in terms of number of QTL detected was found for DMC, DtSILK, and DtTAS. On the contrary, for DMY and PH the connected model detected more QTL. Even if more QTL were detected on average with the LDLA models, the connected model explained a higher percentage of variance (46.9%) than the other models.

For the flint group, the LDLA – 1-marker model detected more QTL (59 QTL in total) but

explained a smaller percentage of variance (47.3% on average) than the other models. In this group, the conventional connected model detected the smallest number of QTL (55 in total). The LDLA models using "clusthaplo" information detected an intermediate number of QTL (58 and 56 for the LDLA – 5 cM and LDLA – 2 cM models, respectively). The ranking of the models in terms of number of detected QTL varied depending on the trait. For instance, the two LDLA models using "clusthaplo" information detected more QTL than the conventional connected model for DtTAS, PH, DMC (with the LDLA – 5 cM model only), and for DMY (with the LDLA – 2 cM model only). For the flowering traits, the LDLA – 1-marker model detected more QTL than the other models. As for the dent group, the connected model explained a higher percentage of variance (56.3%) compared to the other models even if it did not detect a higher number of QTL.

One can note that the – log10(p-values) curves showed relatively noisy patterns along the genome, especially for the LDLA models (Figure 3, Figure S1, Figure S2, Figure S3, Figure S4). However, curves displaying evolution of –log10(p-values) along the genome were globally highly consistent across models and all models detected the same major QTL (Figure 3, Figure S1, Figure S2, Figure S3, Figure S4). This was true even in cases when they detected a different number of QTL on the same chromosome. For instance, in the flint design, for DMC, all models detected a major QTL at 45 – 46 cM on chromosome 10 but two models detected other QTL in the region without challenging the position of the major QTL: the LDLA – 2 cM model at 69.9 cM and the LDLA – 1-marker model at 68.9 cM (Figure S1, Table S3, Table S4, Table S5, Table S6).

Considering the QTL which were detected by different models, the ranking of the models according to their – log10(p-value) varied with the QTL. For instance, for the QTL detected with all models for DtSILK in the dent group at 70 – 74 cM on chromosome 6, the highest – log10(p-value) was found with the LDLA – 2 cM model (17.5) and the lowest with the connected model (13) (Figure 3). On the contrary, for the QTL detected with all models for DMY in the dent group on chromosome 6 at 14 – 17 cM, the highest – log10(p-value) was found with the connected model (14.9) (Figure S2, Table S7) and the lowest with the LDLA – 2 cM model (13.3) (Table S9).

Allelic effect series and comparison of the different allelic models for the major QTL detected for female flowering time

Visualization of allelic effects of the connected model through heat maps (Figure S5, Figure S6, Figure S7, Figure S8, Figure S9, Figure S10, Figure S11, Figure S12, Figure S13, and Figure

S14) illustrated a continuous range of effects for all QTL. The central line had an intermediate value for most of the loci in both designs. Each parental line carried alleles with either positive or negative effects compared to the central line. LDLA models are expected to outperform the connected model if the clustering process correctly identifies underlying allelic series at QTL. To get further insight into this point, we compared allelic effects estimated by the different models for the two major DtSILK QTL found in this study.

The allelic effects of the DtSILK major QTL detected in the flint group on chromosome 10 at 38 – 50 cM clearly showed an allelic series (Figure 4). The four models detected QTL in this region but at slightly different positions. For the QTL detected with the connected model, at least three classes of effects were identified based on t-tests. F283 and DK105 carried a late allele (3.7 d and 3.5 d compared to UH007), UH006 an intermediate allele (2.07 d), and D152, UH009, F2, UH007, and F03802 an early allele (between -0.29 and 0.4 d), the three other parental lines showing effects between the early and the intermediate classes. For the QTL detected with the LDLA - 5 cM and LDLA – 2 cM models, allelic effects were globally consistent with those found for the QTL detected with the connected model except for EZ5 which had the earliest allele with the LDLA – 5cM model. Note that the family derived from this parent was one of the smallest of the design. The LDLA – 1-marker model detected two QTL in this region: one at position 45.9 cM (close to the position of the QTL found with the other models) and one, of smaller effect, 7 cM apart at the position 38.6 cM. For the marker detected at position 45.9 cM, the late allele (2.44d) was shared by F283, DK105 and UH006, which also carried the latest alleles according to the other models. All the other lines shared the same early allele (0d). For the marker detected at position 38.6 cM, the late allele (1.1 d) was shared by DK105, F283 (the lines carrying the latest alleles in the other models), EC49A, and F64 (which carried alleles classified as intermediate). All the other lines shared the early allele (0 d). So, when considered jointly, these two markers account for the allelic series observed for the QTL detected with the other models: DK105 and F283 carrying the late alleles at the two markers; UH006 carrying the late allele for the marker with the strongest effect and the early allele for the other marker; EC49A and F64 carrying the late allele at the marker with the smallest effect and the early allele for the other one, and D152, UH009, F2, UH007, and F03802 carrying at both markers the early alleles. The two QTL detected with the LDLA – 1-marker model individually explained 2.2% and 11.1% of the variance for the marker at positions 38.6 and 45.9 cM, respectively, but they jointly explained 26.8% of the variance, only slightly less than the variance explained by the QTL detected with the other models (between 27.5 and 28.2%).

The allelic effects of the DtSILK QTL detected in the dent group, on chromosome 8 at 45 – 58 cM also clearly showed an allelic series and the same type of pattern (Figure 5). With the connected model, allelic effects showed a continuous variation and at least two classes of alleles could be identified. Four inbred lines (D06, D09, UH250, and F618) carried early alleles compared to the group consisting of F353 (central line), EC169, and Mo17. The other parental alleles were not clearly classified but had intermediate effects. In this chromosome region, the two LDLA models based on ancestral allele clustering both identified a QTL. With both window sizes D06, D09, and UH250 which carried the earliest alleles in the connected model, were attributed to the same ancestral allele with an early effect (-1.77 with LDLA – 5 cM and -1.76 with LDLA – 2 cM compared to F353). Mo17, EC169 (the two lines with latest allelic effects in the connected model), UH304, and F353 were attributed to the same or to different ancestral alleles depending on the window size but in both cases their allelic effects were equal or close to zero. With these models, B73 was attributed the latest effect (0.4 or 0.49) but this effect was not significantly different from zero. The other lines had allelic effects consistent with the effects estimated with the connected model. Two QTL were detected in this region with the LDLA – 1-marker model: one at 45.5 cM and the other at 57.3 cM, on either side of the QTL detected with the other models. D06, D09 and UH250 which carried the earliest allele of the connected model and were attributed to the same early ancestral allele with LDLA-2cM and LDLA-5cM models, carried the early allele at both QTL. Mo17, EC169, B73 and F353, the lines with the latest allelic effects with the other models, carried the late allele at both QTL. The other lines, which had intermediate allelic effects with the other models, carried the late allele at one QTL and the early allele at the other QTL. Thus, marker effects at these two QTL jointly mimic the allelic series identified by the other models. The two QTL detected with the LDLA – 1-marker individually explained 1.5 and 2.9% of the variance but they jointly explained 7.9% of the variance, which is only slightly less than the other models (8.9% for the LDLA – 5 cM and LDLA – 2 cM models, and 9.6% for the connected model).

Comparison of the QTL detected in the two heterotic groups analyzed individually and jointly

In total, for the connected model, 52 QTL were detected in the dent design for all traits and 55 in the flint design (Table 2). More QTL were found in the dent than in the flint design for DMC and PH, whereas the reverse was observed for DtSILK, DtTAS, and DMY.

Based on overlap of their confidence intervals, when comparing results obtained in the two separate datasets only seven QTL were common between the two groups. Two of these QTL

were for DMC (chromosomes 8 and 10), three for DtSILK (chromosomes 1, 2 and 3), one for DtTAS (chromosome 3), and one for PH (chromosome 1). No common QTL were found for DMY. In addition, some chromosome regions carried QTL detected in the two groups but not for the same trait (Figure 6).

The distribution of QTL effects (in terms of R^2) differed in the two groups (Figure 7). In the dent group, all the QTL had low to medium effect (R^2< 10%). The QTL with the biggest effect was detected on chromosome 3 at 63 cM for DtTAS and explained 10.4% of the variance (Table S7). A QTL was also detected at this position for DMC but with a smaller effect. The second biggest QTL was detected on chromosome 8 at position 50 cM for DtSILK and explained 9.4% of the genetic variance. This region was also detected for the other traits but with smaller effects. On the contrary, in the flint group, one region located on chromosome 10 around position 44 – 50 cM showed a major effect on all the traits (Table S3). Depending on the trait considered this region explained between 14% of the variation for DMY and 27.5% for DtSILK. All the other QTL detected in this group showed milder effects with R^2 <10%. It is interesting to note that the QTL which exhibited a strong effect in one group (the QTL detected on chromosome 10 in the flint group and the QTL detected on chromosome 3 and 8 in the dent group) did not have such a strong effect in the other group for the same traits.

87 QTL were detected in total with the joint analysis, which is less than the sum of the QTL found in the two separate datasets (107) (Table 2, Table S11). For each trait, the number of QTL detected with the joint analysis was equal or superior to that detected in each single dataset analysis. For DMC and PH, QTL detected with the joint analysis explained a larger fraction of variance than the one explained in the separate datasets analysis. On the contrary, for DMY, DTSILK and DtTAS, more variance was explained in the flint dataset analysis than in the joint analysis.

QTL found in the joint analysis were generally found at the same position or close to QTL detected in one or both separate analyses (Figure 6). In some cases, they were detected between two QTL detected in a single dataset analysis (for instance QTL on chromosome 5 for DtSILK), or between one QTL detected in the dent dataset and one detected in the flint dataset (QTL at 130 cM on chromosome 2 for DMC). In some cases, no QTL was detected with the joint analysis although QTL were detected in the separate datasets (for instance flint QTL at 9 cM on chromosome 1 or dent QTL on chromosome 2 for DtTAS). Other QTL were detected only with the joint analysis (and not close to or between two QTL detected with the separate analysis), as the one detected for DMC on chromosome 7.

When testing the effects of these 87 QTL in the separate datasets, 30 were significant in both

datasets, 52 in a single dataset only and five in none of the datasets (Table 3). So the number of QTL with effect in both dataset varied between 27% for DtSILK to 41% for DtTAS.
Concerning the seven QTL found common when comparing the dent and flint separate analyses, the joint analysis always found a QTL in the region nearby (not necessarily with overlapping of the confidence regions but really close). Except for the QTL found on chromosome 2 for DtSILK, these QTL were significant in both groups.

DISCUSSION

Our study aimed at comparing genetic determinism of biomass related traits in two complementary flint and dent genetic pools that are often used to produce commercial hybrids in Northern Europe. To do so, a new NAM DH population was developed for each group. Both NAM populations display intermediate levels of diversity compared to the US NAM design and classical elite breeding programs. Data from each design were analyzed with four models: a connected model where parents are assumed to carry different alleles, an LDLA model based on single marker information close to the one successfully used for the US NAM design, and two LDLA models based on ancestral allele modeling previously used with success by Leroux *et al.* (2014) and Bardol *et al.* (2013). In addition data of the two designs were analyzed jointly with the connected model, considering that the central line of one design was used as tester in the other design and reciprocally.

Linkage disequilibrium and clustering of parental alleles

The haplotype clustering approach of Leroux *et al.* (2014) requires the definition of a window size according to genetic map units (cM). We defined it based on the estimation of the LD extent at the level of the parental lines. This showed that LD decreased below r^2=0.2 after approximately 1 cM and 2 cM in the flint and dent parental lines, respectively. Although estimated with only 11 and 12 inbred lines, for the dent and flint group respectively, these values were consistent with the LD extent observed for these groups by van Inghelandt *et al.* (2011). Based on this result, we considered two window sizes for the parental clustering, one of 2 cM, more adapted to the flint group and one of 5 cM, more adapted to the dent group. Note that a 1 cM window was also considered but the HMM approach did not converge with the R version we used for this study. These values are smaller than the 10 cM window size used in Bardol *et al.* (2013) to analyze a multi-parental design derived from highly related founders.

In both flint and dent groups, the clustering process identified on average six and seven ancestral alleles per position for the 2 cM and 5 cM window sizes, respectively. The percentage of genome detected as IBD was in agreement with the marker-based similarities between inbred line pairs and pedigree information. These results showed that among dent lines, there were two groups of related lines: (i) D09, D06, and UH250, which came from the breeding program of the University of Hohenheim, and (ii) UH304 and F353, which share a common Iodent background (Bauer *et al.*, 2013). For the flint, there was a separation between EC49A, EZ5, EP44 (the three lines with Spanish origin), and F64 (Argentinean origin) and all the other lines. The number of ancestral alleles detected after clustering with "clusthaplo" varied along the genome, first at the local level from one position to the next. This results in a variation in model dimension along the genome that certainly explains the erratic pattern of the –log(p-values) curves of the LDLA models (see below). Beyond this local variation we observed that on average more ancestral alleles were detected in the centromeric than in the telomeric regions. This result is probably related to the higher number of marker loci per cM in centromeric regions than in telomeric ones. It may be also related to a higher divergence between lines in centromeric regions. The similarity score used in "clusthaplo" is expected to be robust against the difference of marker density inside the sliding windows (Leroux *et al.* 2014). Our results suggest however that we reached here limits in this robustness. As most of the lines were not closely related, the size of IBD segments was expected to be limited, which made them difficult to detect. Visual inspection of the graphs of IBD segments (results not shown) indeed revealed that the segments were in general shorter than in Bardol *et al.* (2013) except for related lines such as D06 and D09. The method implemented in the "clusthaplo" software should therefore be adapted to cope with more diverse sets of lines than the one considered in Leroux *et al.* (2014), possibly by reducing window sizes in regions of the genome where marker density is high and local LD is low relative to the genetic map.

Adapting the method to cope with populations with limited LD also raises issues regarding the genetic map to be considered for the clustering process. Bauer *et al* (2013) showed that even if the individual maps of the families of a given group had globally consistent order, putative inversions were found in some areas. This is in agreement with recent studies which showed copy number variations (Springer *et al.* 2009; Swanson-Wagner *et al.* 2010), chromosomal inversions or translocations between the different maize lines. Ganal *et al.* (2011) also suggested that some regions of the physical map of B73 v2 are not correctly assembled. This may have affected our consensus maps since information from the physical map was used for positioning the markers and this may have affected the clustering process. It appears thus

important to further evaluate the properties of the clustering approach when using denser genotyping data and also evaluate its potential interest in the context of the rapid emergence of sequencing data that may enable a more direct identification of conserved haplotypes between inbred lines.

Comparison of the different QTL detection models

The highest total number of QTL was detected by one of the three LDLA models in both designs. We noted however different trends for the two designs. For the dent, LDLA – 2 cM and LDLA – 5 cM detected very similar numbers of QTL (55 and 56, respectively), more than for both the connected and LDLA – 1-marker models (52 and 45, respectively). Note that Bardol *et al.* (2013) also found that in an elite dent breeding pool, the LDLA method based on ancestral alleles detected on average more QTL than the LDLA – 1-marker model. Our results suggest that the genotyping data and window sizes used for "clusthaplo" were well suited for LDLA models for the dent design. For the flint design, the connected model detected fewer QTL (55) than the LDLA – 5 cM, the LDLA – 2 cM and the 1-marker model (58, 56 and 59 respectively), but differences between models were small on average This suggests that the the available density of genotyping data and/or window size we could use with the HMM approach were not necessarily optimal for this design. Interestingly, although the connected model was globally outperformed by LDLA models in terms of number of QTL detected, it explained a higher percentage of variance than the other models for nearly all the traits. Conversely, the LDLA – 1-marker model explained a smaller percentage of variance even when detecting more QTL. As the estimations of the percentages of variation explained were adjusted for the number of parameters, this cannot be due to model over fitting. One can thus hypothesize that a large part of the QTL showed allelic series that are not completely accounted for by local similarities or single marker-information. This is consistent with Würschum *et al.* (2012) who compared by simulation different models for joint linkage association mapping. They concluded that, even if the single SNP model was more powerful in terms of detection, the model considering one allele per parent was better adapted to estimate QTL effects in case of multi-allelic series, corroborating experimental results of Liu *et al.* (2011).

Globally, LDLA models and linkage analysis detected QTL in the same chromosome regions although fine comparison of QTL positions was complicated by the relatively noisy pattern of the LDLA –log10(p-value) curves. We noted that the number of QTL in a given genomic region could either be the same or vary across models. In cases when a single QTL position is detected by all models, one can assume that variation is most likely due to a single QTL with two alleles

well reflected by a single biallelic marker. On the opposite, a variable number of QTL across models suggests a more complex situation with linkage between several QTL or allelic series at a single QTL. This can be exemplified by the DtSILK QTL detected on chromosome 8 in the dent design. In this region, the LDLA – 1-marker model detected two QTL 12 cM apart and located on both sides of the single QTL detected with the connected model. This suggests that either the two marker loci were needed to account for the allelic series at a single QTL or conversely that the connected model failed at distinguishing the two underlying QTL due to limited recombination in DH families.

The different models thus showed variable efficiency depending on the trait and region considered, which highlights complementarities of different allele coding methods in deciphering allelic series in genetic studies.

Comparison between the QTL detected in the two heterotic groups and evolutionary interpretation

Similar numbers of QTL were detected in the two groups with the separate dataset analyses, showing that both can contribute genetic variation useful for breeding in Northern Europe. Less than 15% of the QTL were common between the dent and flint design when comparing the positions of the QTL detected in the separate dataset analyses. This is consistent with the long time divergence between the dent and flint heterotic groups: more than 500 years (Tenaillon and Charcosset 2011). Part of this low value can be due to power issues. Indeed the joint analysis enabled us to detect additional QTL compared to single group analysis and among the detected QTL with the joint analysis, 34% on average were significant in both groups. However, some QTL detected in individual designs disappeared in the joint analysis which suggests that they were really specific of one group and that variation within the other group diminished power at these QTL in the joint analysis. Some of the QTL detected in the joint analysis were found at an intermediate position between the positions of design specific QTL. This may correspond to a gain in precision but one cannot exclude that these QTL might also correspond to an artifact “ghost” QTL between actual QTL.

Note that in addition to the common QTL, some chromosome regions had an effect in both designs but for different traits. These QTL could be pleiotropic QTL for which effects on some traits were not detected in one of the designs, due to a lack of power, diversity, etc.

When comparing the single dataset analyses, QTL common to flint and dent designs were observed for DMC, DTSILK, DtTAS and PH. It is interesting to note that no common QTL was observed for DMY. With the joint analysis, a smaller percentage of QTL significant in both

datasets was for DtSILK and DMY (27% and 31%, respectively), than for the other traits (33% for DMC to 41% for DtTAS). For traits subjected to directional selection such as DMY, several alleles must have been fixed over time but there is no reason that the same alleles were fixed in both groups, especially considering that selection for hybrid value certainly favored fixation of complementary alleles in each group (Larièpe *et al.* 2012; Schön *et al.* 2010). This may explain why only few common QTL or QTL significant in both groups were detected for DMY. On the contrary, for traits for which a stabilizing selection is performed, the same polymorphisms are more likely to be maintained in both groups. This is the case for PH, DtTAS and also indirectly for DMC since DMC at harvest of a genotype depends on its precocity and its drying speed. Interestingly, common DMC QTL between groups and most of the DMC QTL detected with the joint analysis and significant in both datasets were detected in regions also carrying QTL for flowering time (DtSILK or DtTAS).

The few common QTL between dent and flint groups that we detected could explain the low predictive abilities of the prediction between dent and flint in genomic selection (Meuwissen *et al.* 2001; Jannink *et al.*2010) when dent are in the estimation set and flint in the test set and *vice versa* (Lehermeier *et al.* in press). The presence of a major effect QTL in the flint group might also partly explain this result.

Overview of detected QTL and comparison with literature studies

For the single dataset analyses, between 20 QTL for DMY and 28 QTL for DtSILK were detected in total over the two groups when considering the model which detected the highest number of QTL. For the joint analysis, between 15 QTL for DtSILK and 21 QTL for PH were detected.

For DtSILK, although high, the number of detected QTL is less than the one reported for the US NAM design (39 QTL detected with the multiple family joint stepwise model, 52 with JCIM) (Buckler *et al* 2009; Li *et al.* 2011). This is also less than the total number of QTL estimated through meta-analysis for flowering time (62 and 59 in Chardon *et al.* 2004 and Salvi *et al.* 2009, respectively). QTL detected in our study explained a smaller proportion of the variance (for the connected model the detected QTL explained 52.3%, for the dent design, 59.7 % for the joint analysis, and 69.3%, for the flint design of the within family variability) than the one detected on the US NAM design (89%) (Buckler *et al* 2009; Li *et al.* 2011). Similar trends were observed for male flowering (DtTAS). In our study, all QTL explained 10% or less of variation, with the exception of the main QTL found in the flint design on chromosome 10 (45-50 cM with the connected model). In the joint analysis, this QTL was significant for female

flowering when tested in both datasets whereas for male flowering it was significant only in the flint dataset. This QTL was also found by Blanc *et al.* (2006) and is close to the *ZmCCT* gene which was fine mapped as a major flowering time QTL by Ducrocq *et al.* (2009) and validated by Coles *et al.* (2011). In the flint design, for the connected model, this QTL explained 18.7% and 27.5% of male and female flowering time, respectively. In the joint analysis, it explained 12% and 15.2% of male and female flowering time, respectively. This value is higher than reported for the same region in the US NAM (1.1% for male flowering and 1.3% for female flowering with Joint Linkage Stepwise Model in Buckler *et al.* 2009) and in Blanc *et al.* (2006) (18% for female flowering). These differences can be explained by the fact that several lines in our flint design share a late allele and possibly suggest that the expression of the effect of this QTL is amplified in early flowering backgrounds compared to the later US NAM background. In the dent design analyzed separately, the most significant DtSILK QTL was found on chromosome 8. This QTL does not seem to be located in the region where two major flowering time QTL, *vgt1* and *vgt2* (*ZCN8*), have been fine-mapped (Salvi *et al.* 2007; Bouchet *et al.* 2013). It seems to be close to an area where other studies also found QTL for flowering time (Ducrocq *et al.* 2008; Salvi *et al.* 2009; Bouchet *et al.* 2013).

For plant height (PH), we detected in total 25 QTL which explained 55.0% and 57.1% of the variation for the flint and dent designs, respectively. With the joint analysis, we detected 21 QTL which explained 61.2% of the variation. A recent study (Peiffer *et al.* 2014) based on the US NAM and IBM family (Lee *et al.* 2002) reported 89 family-nested markers detected with an adaptation of JCIM and 277 associations through a joint-linkage-assisted genome wide association study (Tian *et al.* 2011). Except the QTL found on chromosome 10 in the flint design and that likely corresponds to a pleiotropic effect of a major flowering time QTL, no QTL explained more than 10% of the variation, in the separate or joint datasets. As in Peiffer *et al.* (2014), none of the QTL detected in this study seem to be located in the vicinity of known candidate genes for plant height.

For DMY, with the separate analyses, we detected in total 20 QTL which is lower than the number of QTL detected for the other traits. With the joint analysis, we detected 16 QTL which is one of the lowest number of QTL detected. This may be explained by the lower heritability of this trait and the fact that variation for this trait may involve numerous QTL of small effects that are difficult to detect. For DMC, we detected in total 27 QTL with the separate analyses and 18 with the joint analysis. Only few studies address QTL detection for biomass yield and dry matter content, mainly in biparental populations (e.g. Lübberstedt *et al.* 1998; Méchin *et al.* 2001; Barriere *et al.* 2010; Barriere *et al.* 2012). They reported only limited number of QTL

and are not easily comparable with our results. Our study, which led to the detection of many QTL in a multi-parental context, therefore represents a large advance towards understanding the genetics of biomass yield.

Thus globally, although high compared to the number of QTL indentified in biparental populations, the number of QTL detected in this study appears lower than those detected in most comprehensive designs and meta-analysis. Several explanations can be given for this result. First, compared with the US NAM design, our experimental designs explore less diversity and included fewer individuals (841 and 811 DH lines for the dent and flint designs, respectively compared to 5,000 RILs for the US NAM design). Moreover, as DH lines were used instead of RILs, the number of recombination events in our designs is expected to be two times lower per family. This certainly impacted the power and resolution of our designs for deciphering trait variation even with LDLA models. One cannot exclude that QTL detected in our study may indeed correspond to clusters of linked QTL that could have been individually detected using a higher number of individuals, higher number of markers and progenies exhibiting more crossovers (Huang *et al.* 2010). The main specificity of our study compared to the US NAM design, was that the different families were evaluated through their testcross progeny in order to evaluate traits related to biomass production at usual productivity levels. Under the hypothesis of additivity, the genetic variance is expected to be four times lower for testcross value than for *per se* value. In addition, the two central lines of each group that were used as testers for the other group belong to two complementary heterotic pools, so one expects to observe some dominance effects between the flint and the dent alleles at QTL. Such dominance effects may have masked part of the variability in each group. Despite these limitations, as progenies were evaluated based on testcross performance, the QTL detected in this study directly reflect the genetic variation present in each of the two main heterotic groups that is useful for breeding in European conditions.

Acknowledgment

Results have been achieved within the framework of the Transnational Cooperation within the PLANT-KBBE Initiative CirnFed, with funding from the Federal Ministry of Education and Research (BMBF, Germany), Agence Nationale de la Recherche (ANR, France), and Ministry of Science and Innovation (MICINN, Spain). Part of this research was funded by the Federal Ministry of Education and Research (BMBF, Germany) within the Agro-ClustEr Synbreed-Synergistic plant and animal breeding (FKZ 0315528A).

Literature cited

Bardol, N., M. Ventelon, B. Mangin, S. Jasson, V. Loywick, et al., 2013 Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. Theor. Appl. Genet. 126: 2717–2736.

Barriere, Y., V. Mechin, D. Denoue, C. Bauland, J. Laborde, 2010 QTL for yield, earliness, and cell wall quality traits in topcross experiments of the F838 x F286 early maize RIL progeny. Crop Sci. 50: 1761-1772.

Barriere, Y., V. Méchin, B. Lefevre, S. Maltese, 2012 QTLs for agronomic and cell wall traits in a maize RIL progeny derived from a cross between an old Minnesota13 line and a modern Iodent line. Theor. Appl. Genet. 125: 531-549.

Bauer, E., M. Falque, H. Walter, C. Bauland, C. Camisan, et al., 2013 Intraspecific variation of recombination rate in maize. Genome Biol. 14: R103.

Bink, M. C. A. M., L. R. Totir, C. J. F. ter Braak, C. R. Winkler, M. P. Boer, et al., 2012 QTL linkage analysis of connected populations using ancestral marker and pedigree information. Theor. Appl. Genet. 124: 1097–1113.

Blanc, G., A. Charcosset, B. Mangin, A. Gallais, and L. Moreau, 2006 Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. Theor. Appl. Genet. 113: 206–224.

Bouchet, S., B. Servin, P. Bertin, D. Madur, V. Combes, et al., 2013 Adaptation of maize to temperate climates: mid-density genome-wide association genetics and diversity patterns reveal key genomic regions, with a major contribution of the Vgt2 (ZCN8) locus. PloS One 8: e71377.

Browning, B. L., and S. R. Browning, 2009 A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84: 210–223.

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, et al., 2009 The genetic architecture of maize flowering time. Science 325: 714–718.

Charcosset, A., M. Causse, L. Moreau, A. Gallais, 1994 Investigation into effect of genetic background on QTL expression using three connected maize recombinant inbred lines (RIL) populations, pp. 75-84 in: Biometrics in plant breeding: applications of molecular markers, edited by J. W. van Oijen, J. Jansen. Proceedings of the ninth Meeting of the

Eucarpia Section Biometrics in Plant Breeding, 6-8 July 1994, Wageningen, the Netherlands.

Charcosset, A. and A. Gallais, 1996 Estimation of the contribution of quantitative trait loci (QTL to the variance of a quantitative trait by means of genetic markers. Theor. Appl. Genet. 93: 327-333.

Charcosset, A., B. Mangin, L. Moreau, L. Combes, M-F. Jourjon, et al., 2000 Heterosis in maize investigated using connected RIL populations, pp. 89-98 in Quantitative genetics and breeding methods: the way ahead, edited by INRA, Paris, France.

Chardon, F., B. Virlon, L. Moreau, M. Falque, J. Joets, et al., 2004 Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and synteny conservation with the rice genome. Genetics 168:2169-2185.

Coles, N. D., C. T. Zila, and J. B. Holland, 2011 Allelic effect variation at key photoperiod response quantitative trait loci in maize. Crop Sci. 51: 1036-1049.

Cook, J. P., M. D. McMullen, J. B. Holland, F. Tian, P. Bradbury, et al., 2012 Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. Plant Physiol. 158: 824–834.

Ducrocq, S., D. Madur, J.-B. Veyrieras, L. Camus-Kulandaivelu, M. Kloiber-Maitz, et al., 2008 Key impact of Vgt1 on flowering time adaptation in maize: evidence from association mapping and ecogeographical information. Genetics 178: 2433–2437.

Ducrocq, S., C. Giauffret, D. Madur, V. Combes, F. Dumas, et al., 2009 Fine mapping and haplotype structure analysis of a major flowering time quantitative trait locus on maize chromosome 10. Genetics 183: 1555–1563.

Frascaroli, E., T. A. Schrag, A. E. Melchinger, 2013 Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. Theor. Appl. Genet.126: 133-141.

Ganal, M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler, et al., 2011 A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS ONE 6: e28334.

Gore, M. A., J.-M. Chia, R. J. Elshire, Q. Sun,E. S. Ersoz, et al., 2009 A first-generation haplotype map of maize. Science 326: 1115-1117.

Hill, W. G., and A. Robertson, 1968 Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38: 226–231.

Hill, W. G., B. S. Weir, 1988 Variances and covariances of squared linkage disequilibria in finite populations. Theor. Pop. Biol. 33: 54-78.

Huang, X., X. Wei, T. Sang, Q. Zhao, Q. Feng, et al., 2010 Genome-wide association studies of 14 agronomic traits in rice landraces. Nat. Genet. 42:961-967.

van Inghelandt, D., J. C. Reif, B. S. Dhillon, P. Flament, and A. E. Melchinger, 2011 Extent and genome-wide distribution of linkage disequilibrium in commercial maize germplasm. Theor. Appl. Genet. 123: 11–20.

Jannink, J. L., A. J. Lorenz and H. Iwata, 2010 Genomic selection in plant breeding: from theory to practice. Brief. Funct. Genomics 9:166-177.Jansen, R. C., J.-L. Jannink, and W. D. Beavis, 2003 Mapping quantitative trait loci in plant breeding populations: use of parental haplotype sharing. Crop. Sci. 43: 829-834.

Jourjon, M.-F., S. Jasson, J. Marcel, B. Ngom, and B. Mangin, 2005 MCQTL: multi-allelic QTL mapping in multi-cross design. Bioinformatics 21: 128–130.

Kump, K. L., P. J. Bradbury, R. J. Wisser, E. S. Buckler, A. R. Belcher, et al., 2011 Genome-wide association study of quantitative resistance to Southern leaf blight in the maize nested association mapping population. Nat. Genet. 43: 163–168.

Larièpe, A., B. Mangin, S. Jasson, V. Combes, F. Dumas, et al., 2012 The genetic basis of heterosis: multiparental quantitative trait loci mapping reveals contrasted levels of apparent overdominance among traits of agronomical interest in maize (Zea mays L.). Genetics 190: 795-811.

Lee, M., N. Sharopova, W. D. Beavis, D. Grant, M. Katt et al., 2002 Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. Plant Mol. Biol. 48: 453-461.

Lehermeier, C., N. Krämer, E. Bauer, C. Bauland, C. Camisan, et al., 2014 Usefulness of multi-parental populations of maize (*Zea mays* L.) for genome-based prediction of testcross performance. Genetics (in press).

Leroux, D., A. Rahmani, S. Jasson, M. Ventelon, F. Louis, et al., 2014 Clusthaplo: a plug-in for MCQTL to enhance QTL detection using ancestral alleles in multi-cross design. Theor. Appl. Genet. 1–13 (DOI 10.1007/s00122-014-2267-1).

Li, H., P. Bradbury, E. Ersoz, E. S. Buckler, and J. Wang, 2011 Joint QTL linkage mapping for multiple-cross mating design sharing one common parent. Plos One 6: e17573.

Li, J., and T. Jiang, 2005 Haplotype-based linkage disequilibrium mapping via direct data mining. Bioinformatics 21: 4383–4393.

Liu, W., M. Gowda, J. Steinhoff, H. P. Maurer, T. Würschum, et al., 2011 Association mapping in an elite maize breeding population. Theor. Appl. Genet. 123: 847–858.

Liu, W., J. C. Reif, N. Ranc, G. D. Porta, and T. Würschum, 2012 Comparison of biometrical approaches for QTL detection in multiple segregating families. Theor. Appl. Genet. 125: 987–998.

Lübberstedt, T., A.E. Melchinger, S. Fahr, D. Klein, A. Dally, et al., 1998 QTL mapping in testcrosses of flint lines of maize: III. Comparison across populations for forage traits. Crop Sci. 38: 1278-1289.

MaizeGDB data record. http://cur.maizegdb.org/cgi-bin/displayrefrecord.cgi?id=9024747 (temporary)

McMullen, M. D., S. Kresovich, H. S. Villeda, P. Bradbury, H. Li, et al., 2009 Genetic properties of the maize nested association mapping population. Science 325: 737–740.

Méchin, V., O. Argillier, Y. Hebert, E. Guingo, L. Moreau, et al., 2001 Genetic analysis and QTL mapping of cell wall digestibility and lignification in silage maize. Crop Sci. 41: 690-697.

Meuwissen, T. H., and M. E. Goddard, 2001 Prediction of identity by descent probabilities from marker-haplotypes. Genet. Sel. Evol. 33: 605.

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

Peiffer, J. A., M. C. Romay, M. A. Gore, S. A. Flint-Garcia, Z. Zhang, et al., 2014 The genetic architecture of maize height. Genetics (DOI 10.1534/genetics.113.159152).

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

R Core Team, 2013 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

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

Rincent, R., L. Moreau, H. Monod, E. Kuhn, A. E. Melchinger, et al., 2014 Recovering power in association mapping panels with variable levels of linkage disequilibrium. Genetics (DOI 10.1534/genetics.113.159731).

Romay, M. C., M. J. Millard, J. C. Glaubitz, J. A. Peiffer, K. L. Swarts, et al., 2013 Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 14: R55.

Salvi, S., G. Sponza, M. Morgante, D. Tomes, X. Niu, et al., 2007 Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize.

Proc. Natl. Acad. Sci. 104: 11376–11381.

Salvi, S., S. Castelletti, R. Tuberosa, 2009 An updated consensus map for flowering time QTLs in maize. Maydica 54: 501-512.

Schön, C., B. Dhillon, H. Utz, and A. Melchinger, 2010 High congruency of QTL positions for heterosis of grain yield in three crosses of maize. Theor. Appl. Genet. 120: 321–332.

Segura, V., B. J. Vilhjálmsson, A. Platt, A. Korte, Ü. Seren, et al., 2012 An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat. Genet. 44: 825–830.

Sosnowski, O., A. Charcosset, and J. Joets, 2012 BioMercator V3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. Bioinformatics 28: 2082–2083.

Springer, N. M., K. Ying, Y. Fu, C. T. Yeh, Y. Jia, et al., 2009 Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. PloS Genet. 11: e1000734

Swanson-Wagner, R. A., S. R. Eichten, S. Kumari, P. Tiffin, J. C. Stein, et al., 2010 Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. Genome Res. 20: 1689-1699.

Tenaillon, M. I., and A. Charcosset, 2011 A European perspective on maize history. C. R. Biol. 334: 221–228.

Tian, F., P. J. Bradbury, P. J. Brown, H. Hung, Q. Sun, et al., 2011 Genome-wide association study of leaf architecture in the maize nested association mapping population. Nat. Genet. 43: 159-162.

Valente, F., F. Gauthier, N. Bardol, G. Blanc, J. Joets, et al., 2013 OptiMAS: a decision support tool for marker-assisted assembly of diverse alleles. J. Heredity 104: 586-590.

Würschum, T., W. Liu, M. Gowda, H. P. Maurer, S. Fischer, et al., 2012 Comparison of biometrical models for joint linkage association mapping. Heredity 108: 332-340.

Yu, J., G. Pressoir, W. H. Briggs, I. Vroh Bi, M. Yamasaki, et al., 2006 A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38: 203–208.

Yu, J., J. B. Holland, M. D. McMullen, and E. S. Buckler, 2008 Genetic design and statistical power of nested association mapping in maize. Genetics 178: 539–551.

List of Figures

Figure 1 Number of ancestral alleles along the genome after clustering with "clusthaplo" using a 2 cM sliding window size and number of markers in the 2 cM sliding window along the genome for the dent design (6,808 unique positions on the genome – 1,343.3 cM in total) and the flint design design (7,272 unique positions on the genome – 1,586.3 cM in total. The black points correspond to the number of ancestral alleles. The green line corresponds to the number of markers in the 2 cM sliding window alog the genome. Horizontal red lines correspond to the average number of ancestral alleles along the whole genome. The vertical black dotted lines correspond to the limits of each chromosome. 85

Figure 2 Similarities between the dent (left) and the flint parental lines (right), computed based on direct marker genotyping (top) and on ancestral allele sharing (using "clusthaplo" and a 2cM window size) (bottom). Yellow colors correspond to a low similarity, red colors correspond to a high similarity (color scale on the top-right corner). Lines were ordered according to their position in the dendogram (on the top and on the left of each graph) obtained by a hierarchical clustering based on similarities.
A Similarities between the dent parental lines computed based on direct marker genotyping
B Similarities between the flint parental lines computed based on direct marker genotyping
C Similarities between the dent parental lines computed based on ancestral allele sharing (using "clusthaplo" and a 2cM window size)
D Similarities between the flint parental lines computed based on ancestral allele sharing (using "clusthaplo" and a 2cM window size) 86

Figure 3 Results of the QTL detection for with each model for DtSILK for A the dent design and B the flint design. The - log10(p-values) of the connected model are represented by black lines, the QTL positions of the connected models by black dots. The - log10(p-values) of the LDLA – 5 cM model are represented by tblue lines and the QTL positions by blue diamonds. The - log10(p-values) of the LDLA – 2 cM model are represented by red lines and the QTL positions by red crosses. The - log10(p-values) of the QTL detected by the LDLA – 1-marker model are represented by green stars. Horizontal lines correspond to the threshold values of the different models. 87

Figure 4 Allelic effects for the different flint lines for the QTL detected on chromosome 10 at 38 – 50 cM for DtSILK with all the QTL detection models. Allelic effects are estimated in contrast to the central line allelic effect (UH007), which was set to zero. The same letter was given to allelic effects not significantly different at a 5% risk level. Alleles with intermediate effects may be attributed to more than one letter. The last column corresponds to the joint effect of the two QTL detected in the region with LDLA – 1-marker model. Allelic effects estimated for EP44 were not shown because the population where it segregates was too small (17 individuals) to obtain a reliable estimation. Inbred lines are ranked according to their allelic effects obtained with the connected model. 88

Figure 5 Allelic effects for the different dent lines for the QTL detected on chromosome 8 at 45 – 58 cM for DtSILK with all the models. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero. The same letter was given to allelic effects not significantly different at a 5% risk level. Alleles with intermediate effects may be attributed to more than one letter. The last column corresponds to the joint effect of the two QTL detected in the region with LDLA – 1-marker model. Inbred lines are ranked according to their allelic effects obtained with the connected model. 89

Figure 6 QTL projection on the flint-dent consensus map of the QTL detected in the dent dataset, the flint dataset and in the joint dataset for DMC, DMY, DtSILK, DtTAS and PH. Each QTL is displayed by one horizontal line bound by two vertical lines representing the confidence region and a vertical line proportional to the QTL adjusted R^2 symbolizing the QTL position. QTL common to dent and flint according to the overlap of their confidence region on the dent-flint consensus map are represented in red. For the QTL detected in the joint analysis, the letters d and f written below the QTL indicate that the QTL was significant when tested in the dent or flint dataset respectively. 90

Figure 7 Distribution of the percentage of variance ($RQTL2$) explained by the QTL detected in A the dent design and B the flint design, with the connected model and for the five traits. 91

FIGURES

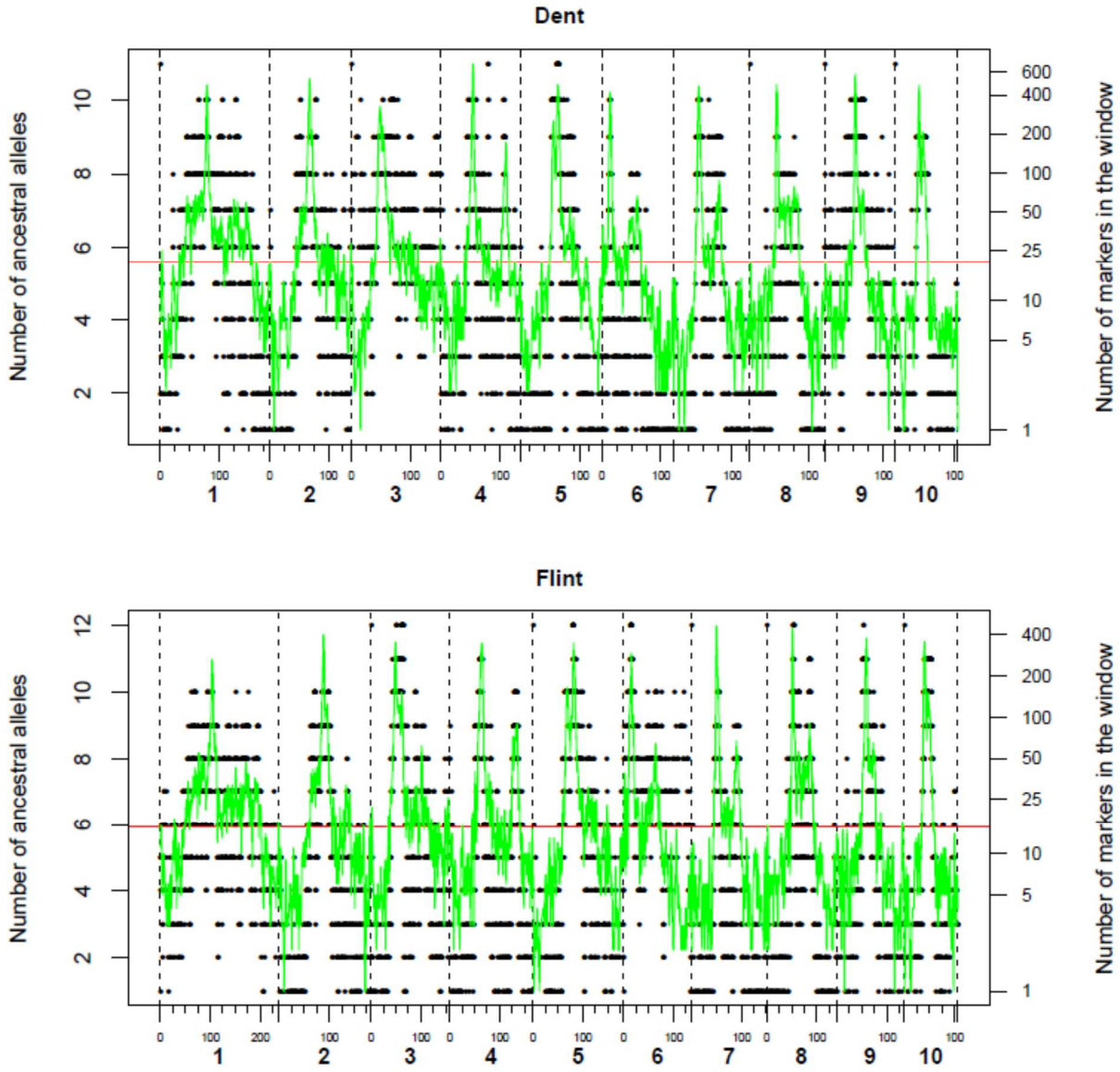

Figure 1 Number of ancestral alleles along the genome after clustering with "clusthaplo" using a 2 cM sliding window size and number of markers in the 2 cM sliding window along the genome for the dent design (6,808 unique positions on the genome – 1,343.3 cM in total) and the flint design design (7,272 unique positions on the genome – 1,586.3 cM in total. The black points correspond to the number of ancestral alleles. The green line corresponds to the number of markers in the 2 cM sliding window alog the genome. Horizontal red lines correspond to the average number of ancestral alleles along the whole genome. The vertical black dotted lines correspond to the limits of each chromosome.

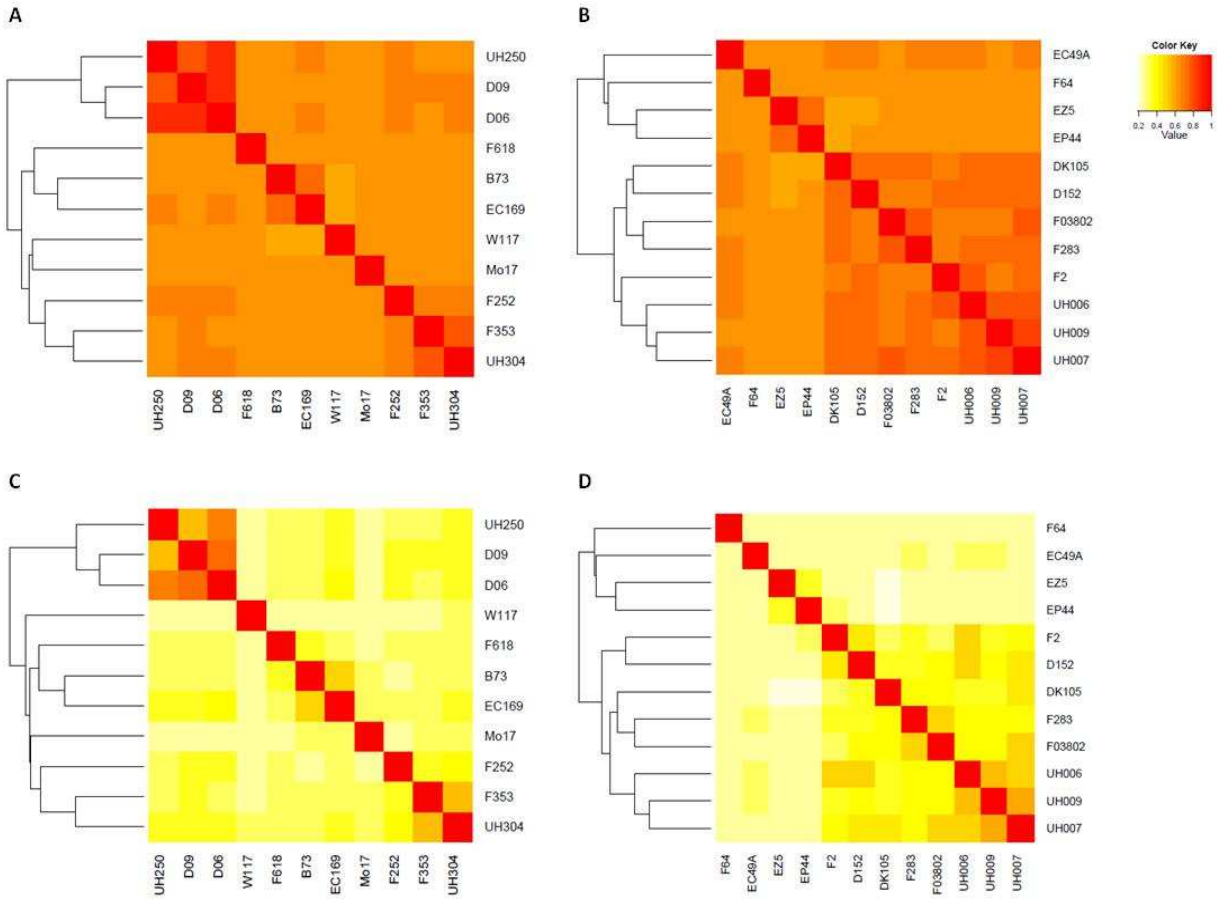

Figure 2 Similarities between the dent (left) and the flint parental lines (right), computed based on direct marker genotyping (top) and on ancestral allele sharing (using "clusthaplo" and a 2cM window size) (bottom). Yellow colors correspond to a low similarity, red colors correspond to a high similarity (color scale on the top-right corner). Lines were ordered according to their position in the dendogram (on the top and on the left of each graph) obtained by a hierarchical clustering based on similarities.

A Similarities between the dent parental lines computed based on direct marker genotyping.

B Similarities between the flint parental lines computed based on direct marker genotyping.

C Similarities between the dent parental lines computed based on ancestral allele sharing (using "clusthaplo" and a 2cM window size).

D Similarities between the flint parental lines computed based on ancestral allele sharing (using "clusthaplo" and a 2cM window size).

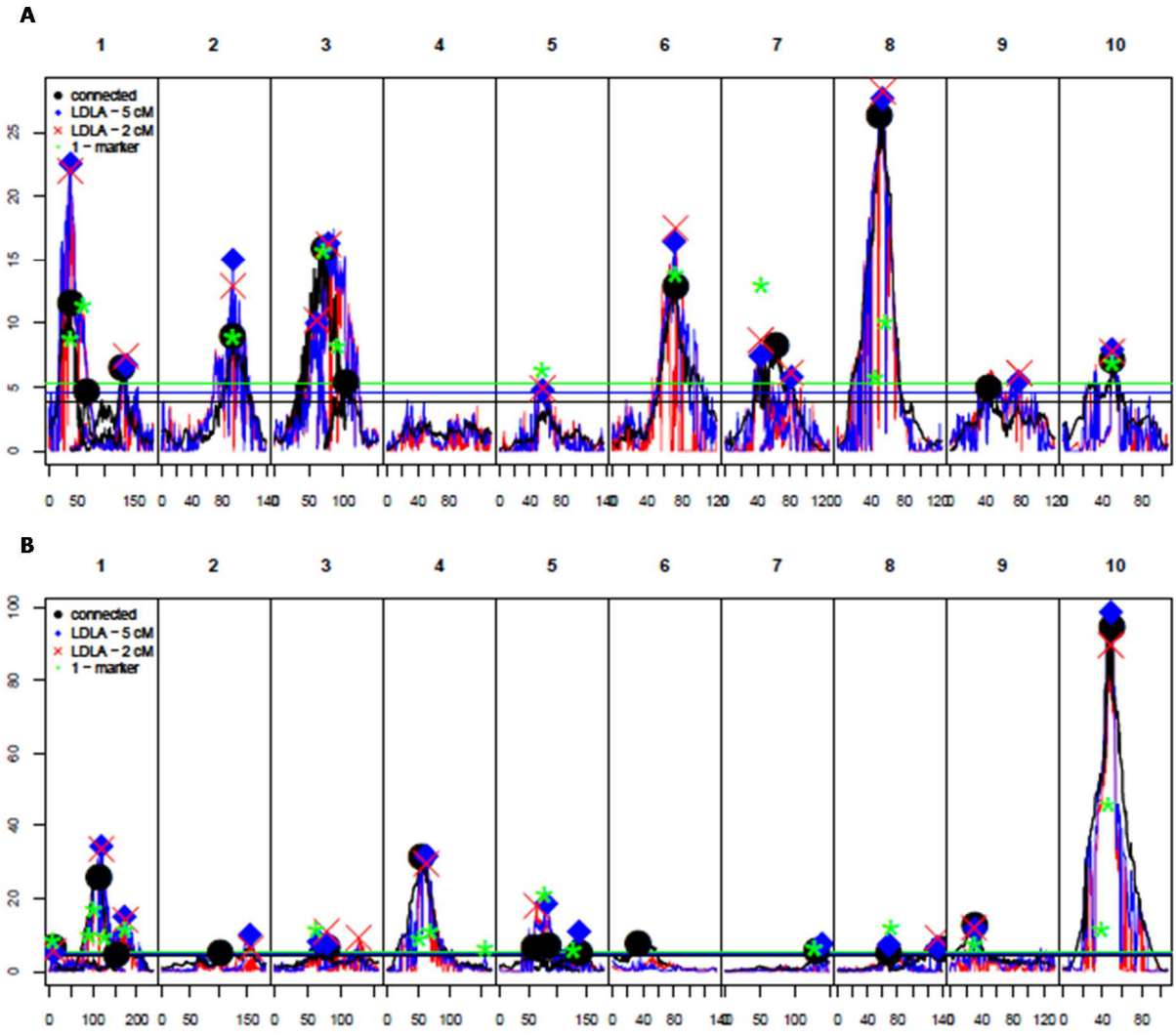

Figure 3 Results of the QTL detection for with each model for DtSILK for A the dent design and B the flint design. The - log10(p-values) of the connected model are represented by black lines, the QTL positions of the connected models by black dots. The - log10(p-values) of the LDLA – 5 cM model are represented by tblue lines and the QTL positions by blue diamonds. The - log10(p-values) of the LDLA – 2 cM model are represented by red lines and the QTL positions by red crosses. The - log10(p-values) of the QTL detected by the LDLA – 1-marker model are represented by green stars. Horizontal lines correspond to the threshold values of the different models.

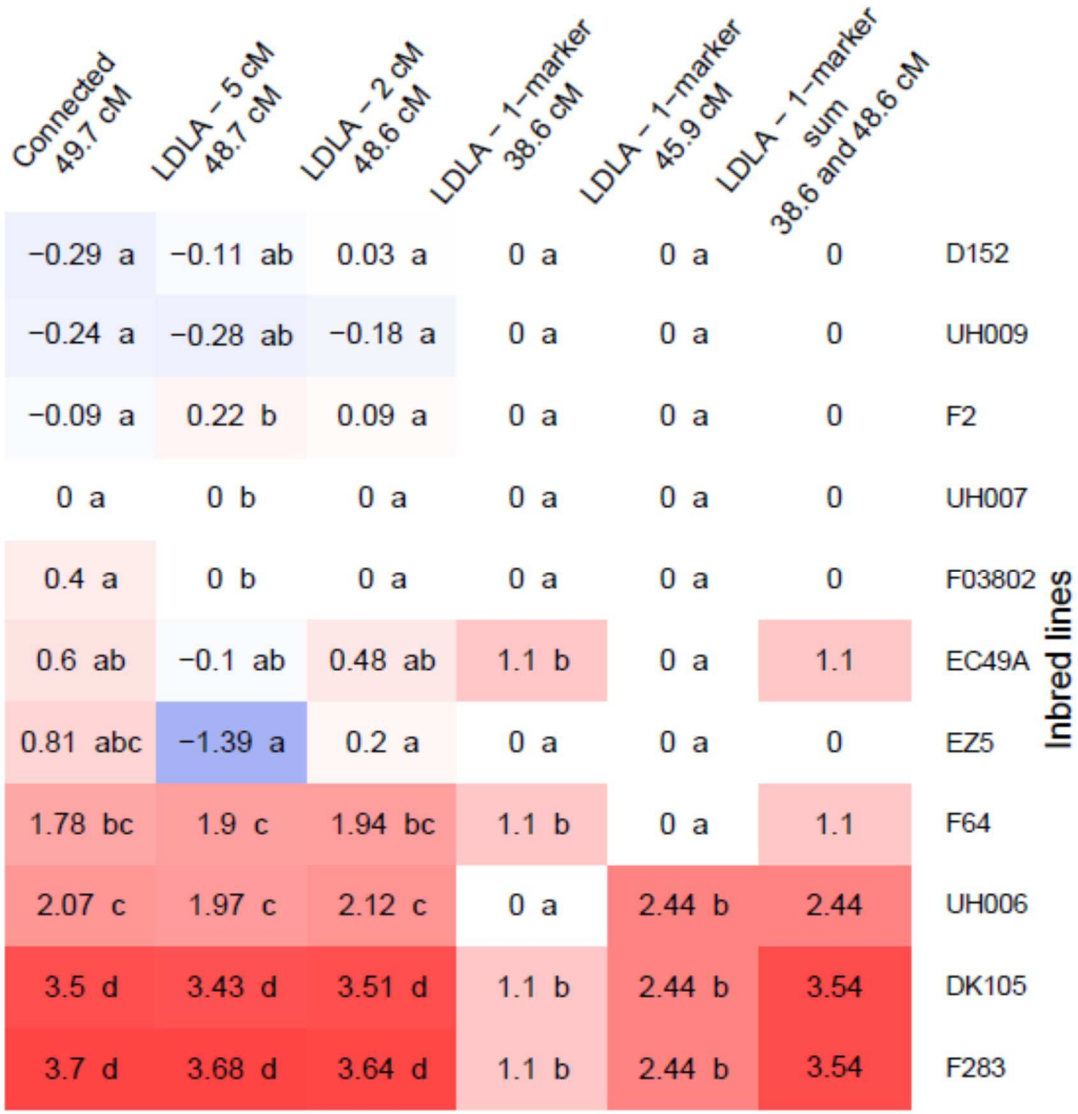

Figure 4 Allelic effects for the different flint lines for the QTL detected on chromosome 10 at 38 – 50 cM for DtSILK with all the QTL detection models. Allelic effects are estimated in contrast to the central line allelic effect (UH007), which was set to zero. The same letter was given to allelic effects not significantly different at a 5% risk level. Alleles with intermediate effects may be attributed to more than one letter. The last column corresponds to the joint effect of the two QTL detected in the region with LDLA – 1-marker model. Allelic effects estimated for EP44 were not shown because the population where it segregates was too small (17 individuals) to obtain a reliable estimation. Inbred lines are ranked according to their allelic effects obtained with the connected model.

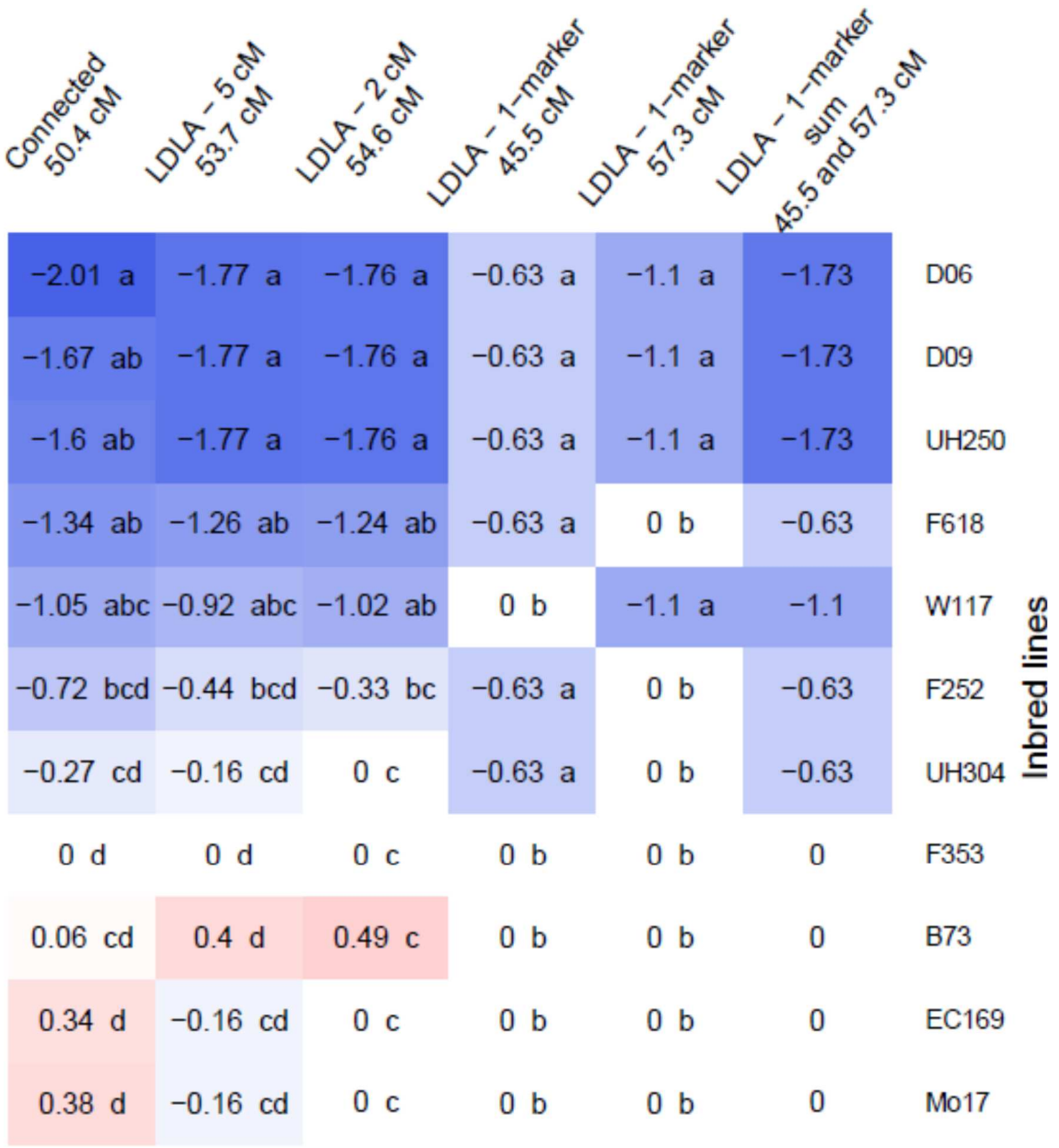

Figure 5 Allelic effects for the different dent lines for the QTL detected on chromosome 8 at 45 – 58 cM for DtSILK with all the models. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero. The same letter was given to allelic effects not significantly different at a 5% risk level. Alleles with intermediate effects may be attributed to more than one letter. The last column corresponds to the joint effect of the two QTL detected in the region with LDLA – 1-marker model. Inbred lines are ranked according to their allelic effects obtained with the connected model.

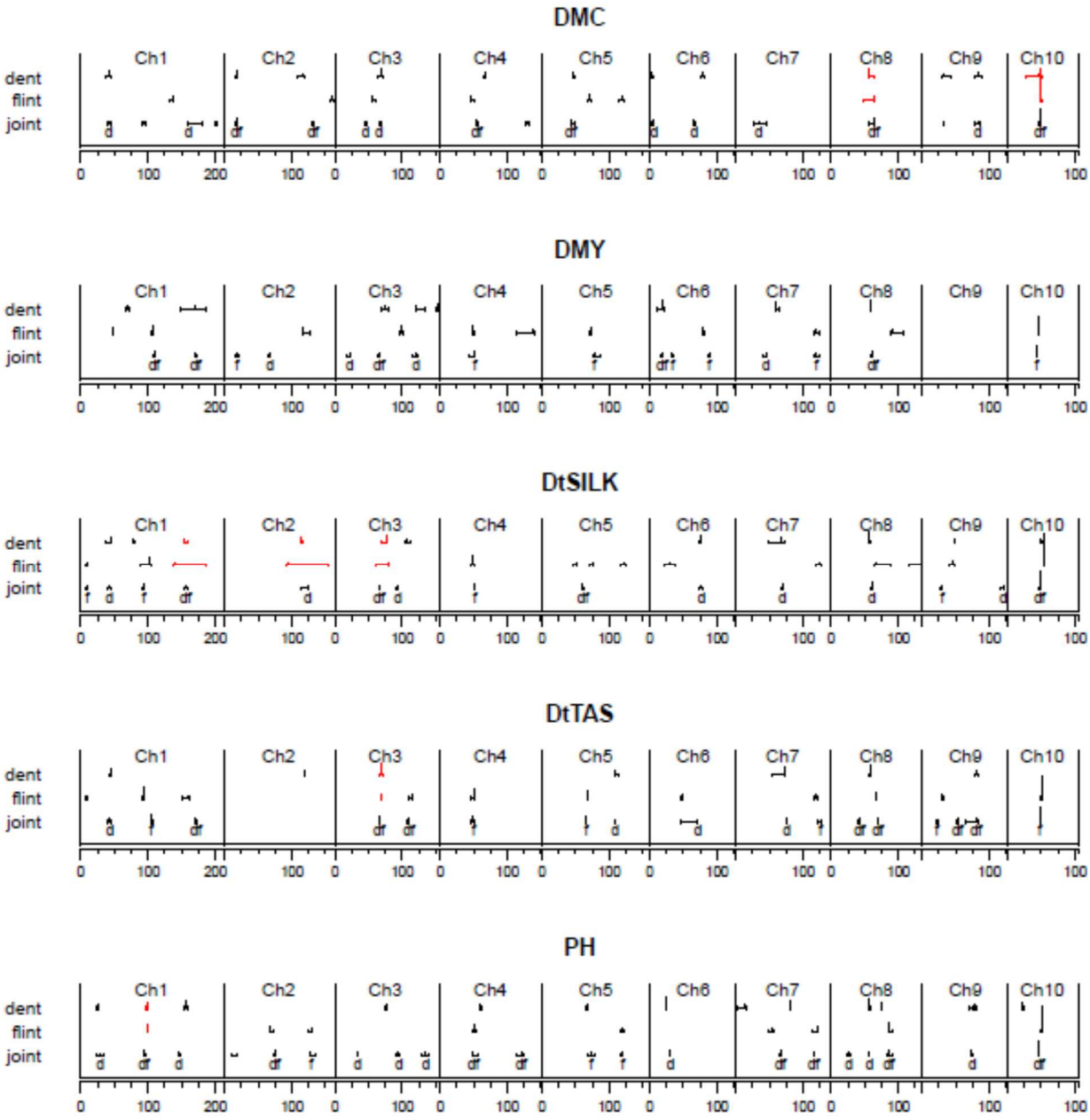

Figure 6 QTL projection on the flint-dent consensus map of the QTL detected in the dent dataset, the flint dataset and in the joint dataset for DMC, DMY, DtSILK, DtTAS and PH. Each QTL is displayed by one horizontal line bound by two vertical lines representing the confidence region and a vertical line proportional to the QTL adjusted R^2 symbolizing the QTL position. QTL common to dent and flint according to the overlap of their confidence region on the dent-flint consensus map are represented in red. For the QTL detected in the joint analysis, the letters d and f written below the QTL indicate that the QTL was significant when tested in the dent or flint dataset respectively.

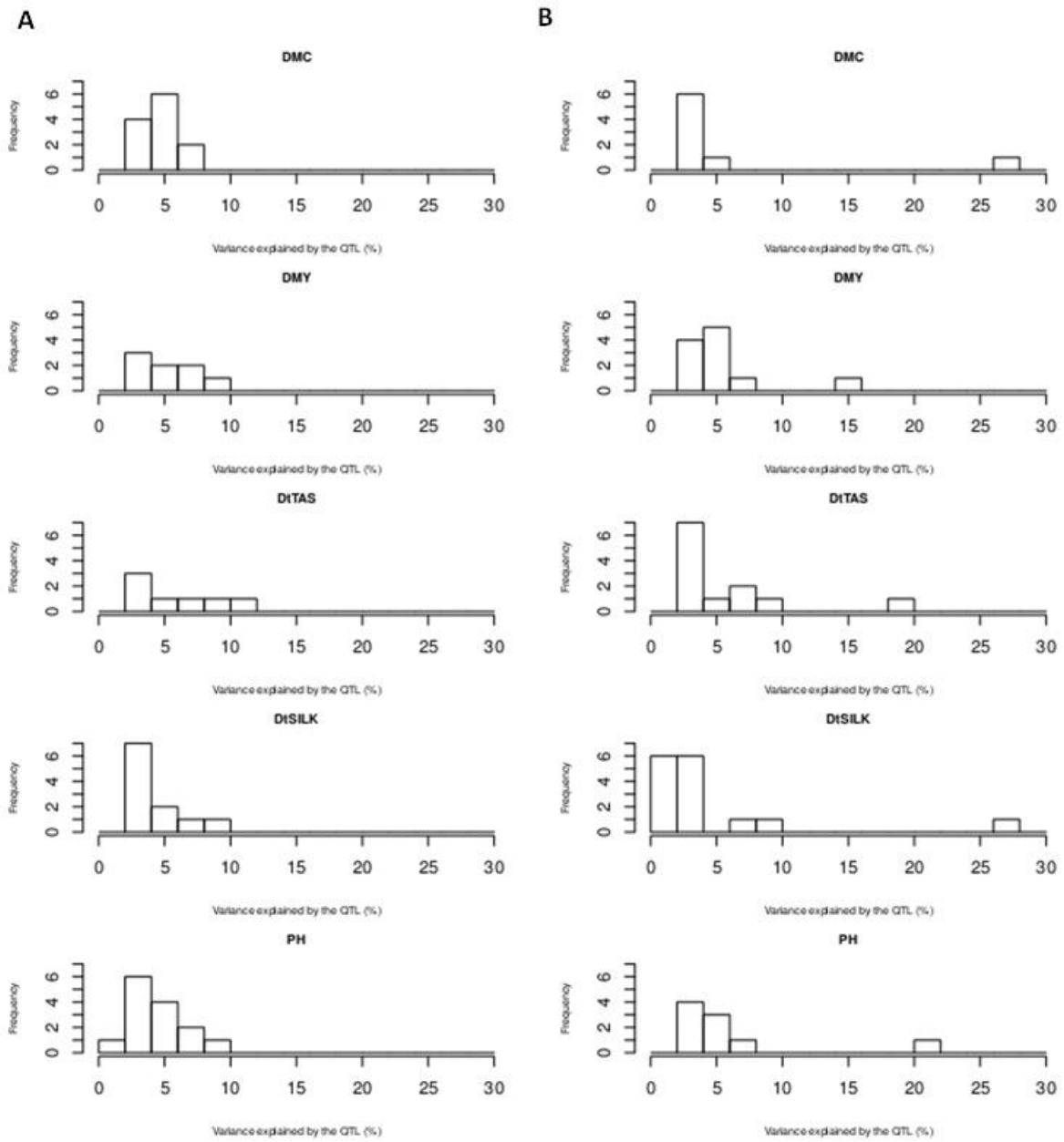

Figure 7 Distribution of the percentage of variance (R^2_{QTL}) explained by the QTL detected in A the dent design and B the flint design, with the connected model and for the five traits.

List of Tables

Table 1 Number of mapped markers, length of the genetic map and linkage disequilibrium decay modeled with the Hill and Weir model (1988) for a r2 = 0.2 for the two groups dent and flint for each chromosome and for the whole genome. 93

Table 2 Number of QTL detected (Nb) and adjusted percentage of variance explained by the detected QTL (R^2) for the five traits in the two separate datasets for each model and for the joint dataset for the connected model. We also indicated the total number of QTL detected over the traits and the average percentage of variance explained ("Total" column). 94

Table 3 Number of QTL detected for the five traits in the joint dataset for the connected model, in the whole dataset, in each separate dataset, in both separate datasets. 95

TABLES

Table 1 Number of mapped markers, length of the genetic map and linkage disequilibrium decay modeled with the Hill and Weir model (1988) for a r2 = 0.2 for the two groups dent and flint for each chromosome and for the whole genome.

	Dent (cM)			**Flint (cM)**		
	Markers	Length (cM)	LD decay (cM)	Markers	Length (cM)	LD decay (cM)
Chrm 1	3287	184.5	0.96	2892	237.2	0.76
Chrm 2	2402	137.9	2.51	2264	182.7	0.65
Chrm 3	2480	151.0	1.99	2410	156.4	0.45
Chrm 4	2528	134.6	1.47	2379	165.5	0.65
Chrm 5	2405	136.6	0.45	2322	180.6	0.35
Chrm 6	1695	119.9	1.47	1544	134.9	0.65
Chrm 7	1820	128.9	1.37	1709	149.7	0.76
Chrm 8	1992	125.6	1.47	1756	139.7	0.45
Chrm 9	1699	118.5	0.96	1610	133.5	0.76
Chrm 10	1570	105.8	1.89	1520	106.1	0.76
Genome	**21878**	**1343.3**	**1.2**	**20406**	**1586.3**	**0.65**

Table 2 Number of QTL detected (Nb) and adjusted percentage of variance explained by the detected QTL (R^2) for the five traits in the two separate datasets for each model and for the joint dataset for the connected model. We also indicated the total number of QTL detected over the traits and the average percentage of variance explained ("Total" column).

	DMC		DMY		DtSILK		DtTAS		PH		**Total**	
	Nb	R^2 (%)	Nb	R^2 (%)	Nb	R^2 (%)	Nb	R^2 (%)	Nb	R^2 (%)	**Nb**	**R^2 (%)**
Dent												
Connected	12	51.4	8	32.7	11	52.3	7	41.2	14	57.1	**52**	**46.9**
LDLA - 5cM	15	51.1	5	22.5	12	53.7	11	49.2	13	54.1	**56**	**46.1**
LDLA - 2cM	16	53.6	6	23.4	12	53.2	9	45.1	12	49.5	**55**	**45.0**
LDLA – 1-marker	12	37.4	5	18.6	11	43.2	7	33.3	10	36.4	**45**	**33.8**
Flint												
Connected	8	46.0	11	48.6	15	69.3	12	65.3	9	52.3	**55**	**56.3**
LDLA - 5cM	11	49.2	10	41.9	14	67.5	13	61.1	10	51.7	**58**	**54.3**
LDLA - 2cM	8	42.1	12	45.3	11	62.0	14	62.2	11	51.9	**56**	**52.7**
LDLA – 1-marker	7	36.1	11	39.0	16	61.7	16	58.0	9	41.9	**59**	**47.3**
Joint												
Connected	18	54.6	16	45.5	15	59.7	17	61.4	21	61.2	**87**	**56.5**

Table 3 Number of QTL detected for the five traits in the joint dataset for the connected model, in the whole dataset, in each separate dataset, in both separate datasets.

	DMC	DMY	DtSILK	DtTAS	PH	**Total**
Significant in the whole dataset (nb)	18	16	15	17	21	**87**
Significant in the dent dataset (nb)	14	9	11	11	17	**62**
Significant in the flint dataset (nb)	6	12	8	13	11	**50**
Significant in both datasets (nb)	6	5	4	7	8	**30**
Non significant in both datasets (nb)	4	0	0	0	1	**5**

Chapter 2

QTL detection on maize (*Zea mays* L.) hybrids derived from two multiparental designs allowed the simultaneous detection of QTLs for general and specific combining abilities for silage performances

Héloïse Giraud[*], Cyril Bauland[*], Matthieu Falque[*], Delphine Madur[*], Valérie Combes[*], Philippe Jamin[*], Cécile Monteil[*], Carine Palaffre[†], Antoine Gaillard[‡], Philippe Blanchard[§], Alain Charcosset[*], Laurence Moreau[*1]

[*] GQE-Le Moulon, INRA, Univ Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, F-91190, Gif-sur-Yvette, France

[†] UE0394 SMH Maïs, INRA, 2297 route INRA, F-40390 Saint-Martin-de-Hinx, France

[‡] Maïsadour Semences SA, Route de Saint Sever, BP27, F-440001 Mont de Marsan Cedex, France

[§] Euralis Semences, Domaine de Sandreau, 6 chemin de Panedautes, Mondonville, F-31700, France

[1]Corresponding author

Running title: QTL detection in reciprocal design

Key words: QTL detection; General Combining Ability (GCA); Specific Combining Ability (SCA); silage; maize

Corresponding author:
Dr. Laurence Moreau
GQE-Le Moulon, INRA, Univ Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay
F-91190
Gif-sur-Yvette
France
Phone: +33 1 69 33 23 37
Fax: +33 1 69 33 23 40
Email: moreau@moulon.inra.fr

ABSTRACT

Understanding genetic architecture of hybrid performances is of key importance for allogamous species such as maize (*Zea mays L.*). We developed two multiparental designs corresponding each to one of the main heterotic groups used for maize silage production in Northern Europe (the dent and flint groups). In each group, four founder lines were crossed to produce six connected biparental populations of segregating lines. These lines (821 and 801 for the dent and flint group, respectively) were genotyped for approximately 20k SNPs and were crossed according to an incomplete factorial design to produce 951 dent-flint hybrids, evaluated for silage performances in eight environments. Hybrid genetic variance decomposition showed a predominance of general (GCA) over specific (SCA) combining abilities. SCA explained between 13.8 and 22.6% of the within-population hybrid variance, depending on the trait. QTL detection was carried out for GCA and SCA using different models considering allelic effects transmitted from each founder lines (linkage analysis) or considering directly SNP alleles (linkage disequilibrium mapping) assuming equal or different effects in each group. In total, between 42 and 54 QTLs were detected depending on the model, among which 12 to 31% presented dominance/SCA effect significant at a 5% individual risk level. Only 16 QTLs were detected by all three models illustrating their complementary. Most of the QTLs (about 80%) were specific to one group, consistent with the long term divergence between the dent and the flint group. These results open interesting prospects for revisiting with markers the concept of reciprocal recurrent selection.

INTRODUCTION

Darwin, in 1876, observed that cross-pollination led to more vigorous plants than self-pollination for numerous species. This observation was later theorized by Shull (1908) and East (1908) who defined the concept of hybrid vigor or heterosis (Shull 1914). In maize, as in other allogamous species, heterosis is important for traits related to yield. In order to exploit this phenomenon, maize diversity was partitioned into heterotic groups and most of the varieties that are sold today correspond to hybrids between inbred lines belonging to complementary heterotic groups. When suitable heterotic groups have been established, the objective of breeders is to select new lines within each group and identify the best hybrid combinations between them. Several heterotic patterns are used in maize breeding depending on the region considered and on the breeding objectives. For instance, in Northern Europe, hybrids selected for silage generally issued from crosses between the dent and the flint groups.

Hybrid value is traditionally decomposed into two parts. The first one is the sum of the General Combining Abilities (GCA) of the two parental inbred lines, defined each as the average value of the hybrids that can be derived from this line when crossed to lines from the other group. The second one is the Specific Combining Ability (SCA) of the pair of parental lines, which is the difference between the hybrid value and the value predicted based on GCAs (Sprague and Tatum 1942). GCAs are statistically additive and involve additive, dominance and epistatic effects at quantitative trait loci (QTL). SCA only involve dominant and epistatic QTL effects. In breeding programs, due to practical considerations, all potential inter-group combinations cannot be evaluated phenotypically. Consequently, the selection is carried out in two stages. In the first stage, future potential inbred lines of each group are selected for their cross-value with one or few lines representative of the complementary group, called testers. In the second stage, a limited number of combinations between improved inbred lines of both groups are evaluated to identify the best hybrids. In this scheme, most of the selection is generally performed during the first stage. Due to the small number of testers considered, the selection of lines is based on a combination of their GCA and of the SCA with the tester(s). Understanding the relative magnitude of these components is therefore important to evaluate to which extent the choice of tester(s) may affect the estimation of the potential of new inbred lines. Reif *et al.* (2007) theorized that, in absence of epistasis, the ratio of SCA over GCA variance is expected to be low for hybrids between two divergent populations. This is consistent with the empirical expertise of the breeders, who observe that selection with a given tester generally leads to a correlative progress of hybrid performance with other partners. This is also consistent with

experimental studies (for instance Schrag *et al.* 2006 or Fischer *et al.* 2008 for grain yield, Argillier *et al.* 2000 or Grieder *et al.* 2012 for whole plant biomass yield).

Beyond global statistics of the relative magnitude of GCA and SCA, identification of Quantitative Trait Loci (QTL) involved in the genetic architecture of hybrid values and these two components is needed to better understand hybrid variation and improve the efficiency of hybrid breeding. Most of the QTL detection experiments conducted so far on maize yield related traits have been based on biparental populations evaluated with a single tester (Manicacci *et al.* 2011; Truntzler *et al.* 2012 for an overview of QTL detected for maize silage) which does not enable the detection of SCA effects. Few studies involved several testers. As expected, strong SCA effects were found for traits showing dominance (such as yield) when the testers used were related to the parents of the studied population, leading to hybrids with different levels of inbreeding (Lu *et al.* 2003; Frascaroli *et al.* 2007; Frascaroli *et al.* 2009; Larièpe *et al.* 2012). Several studies also reported poor consistency between the QTLs found with different testers for yield performances even in the cases when the testers were non-related to the studied population (Schön *et al.* 1994; Lübberstedt *et al.* 1997; Melchinger *et al.* 1998; Austin *et al.* 2000). This suggests that even in a context where small SCA is expected, tester choice may affect QTL detection results.

QTL detection in multiparental designs showed their interest for exploring a larger part of the diversity and for increasing power in comparison to biparental designs (Blanc *et al.* 2006; Kump *et al.* 2011; Bardol *et al.* 2013; Giraud *et al.* 2014; Foiada *et al.* 2015). Only few studies have investigated their interest for better understanding the genetic architecture of additivity – dominance or GCA and SCA. Larièpe *et al.* (2012) studied an advanced Northern Carolina III design between three connected RIL population and their three parental lines and found QTL with apparent overdominance for heterotic traits such as yield in the centromeric regions. The strong SCA effects were due to the presence in the studied design of hybrids between related lines.

Other QTL studies were conducted directly on hybrids between sets of lines selected in complementary genetic groups. This was carried out first with a limited number of SSR markers by Parisseaux and Bernardo (2004), then by van Eeujwick *et al.* (2010) who analyzed another factorial design genotyped with SNP markers. Both studies identified QTLs for GCA that were specific to each heterotic group. SCA effects were considered as negligible and were not included in these analyses. More recently Technow *et al.* (2014) also analyzed a factorial design corresponding to the last stages of a breeding program. The main objective of their study was to perform genomic predictions. The use of a Bayes B model led to the identification of a few

markers with sizable effects on GCA and SCA but they made only little comment on them. Thus, all these studies did not really considered SCA effects in the detection. They also relied on designs which may not be well adapted for QTL detection. Only hybrids between lines selected based on tester values were evaluated. These designs involved in each group a lot of parental lines more of less related to each other and that did not contribute equally to the hybrid population. All this complicated and possibly biased the estimation of SCA/GCA components and the identification of QTLs. To our knowledge, no QTL detection was carried out on hybrid design between inbred lines developed directly from segregating populations available at early selection stages in two complementary heterotic groups. With the development of double haploidization techniques, breeders can directly generate at each breeding cycle segregating populations composed of pure inbred lines. Instead of using a small number of testers from the opposite group to select among them the best ones, and evaluate inter-group hybrid combinations in a second stage, it might be relevant to directly evaluate hybrids between non-selected lines of the two groups. Such type of unselected hybrid population with known family structure is a priori ideal for detecting QTLs for GCA and SCA in a multi-allelic context. Once detected, QTL effects could enable the identification of the best hybrid combinations among all the potential single-cross (including the untested ones). They would also be useful for optimizing the selection of the future inbred lines by better taking into account GCA, SCA and the diversity of the material.

To evaluate this strategy, we developed one dent and one flint multiparental design. For each heterotic group, six biparental populations of inbred lines were developed from four founder lines and were crossed with the ones of the other group. Hybrids were phenotypically evaluated for silage performances. We first decomposed the genetic variance in its GCA and SCA components. We then performed QTL detection considering three different ways of coding the alleles. Predictive abilities of the different models were then compared based on cross-validation.

MATERIAL AND METHODS

Genetic material

The experimental material is constituted of 1044 dent – flint hybrids obtained by crossing inbred lines from two multiparental connected designs corresponding each to one heterotic group (Figure 1). To obtain each of these multiparental connected designs, four inbred lines, further

referred to as “founders”, were chosen: one for its good digestibility ability, the others for their good agronomical potential for yield. They were crossed according to a half diallel design in order to produce six different F_1. From these six F_1, six populations of on average 155.2 lines for the dent design (min 114, max 167) and 152.2 (min 126, max 185) for the flint design were derived. The dent lines were obtained by double haploidization and the flint lines were obtained by selfing independent F_2 individuals for five or six generations depending on the population. 931 dent lines and 913 flint lines, were obtained. From these lines, further called the “parental lines”, 863 dent lines and 879 flint lines were crossed according to an incomplete factorial design in order to produce 1044 experimental hybrids. Each population of one group was crossed with all the populations of the other group, with the objective to balance the contribution of all populations. The majority of lines (699 in the dent and 732 in the flint) contributed to only one hybrid, but some lines contributed twice (163 in the dent group and 146 in the flint group) or even three times (one dent parental line) or four times (one flint parental line). All founder lines of one group were crossed with the founder lines of the other group to create 16 hybrids that were used as checks.

Genotyping data

The founder lines were genotyped with a 50 K SNPs array (Ganal *et al.* 2011). The founder lines and the parental lines were genotyped with a 18 480 SNPs Affymetrix® array provided by Limagrain.

For the analyses we considered the Affymetrix® genotyping data for the founder lines, and when possible replaced missing data by the genotypes obtained with the 50 K SNP array. To avoid the ascertainment bias observed with the 50k SNP array, we only considered the PANZEA markers (Ganal *et al.* 2011) which were polymorphic among the founder lines. We restricted the analysis to loci which had less than 20% of missing values among the dent parental lines, the flint parental lines, and the dent and flint parental lines considered jointly. Markers with more than 5% of heterozygosity among the dent parental lines or in total, or more than 10% of heterozygosity among the flint parental lines were discarded. Markers with a Minor Allele Frequency (MAF) inferior to 5% among the dent parental lines or the flint parental lines or in total were discarded. Thus 9643 markers were considered.

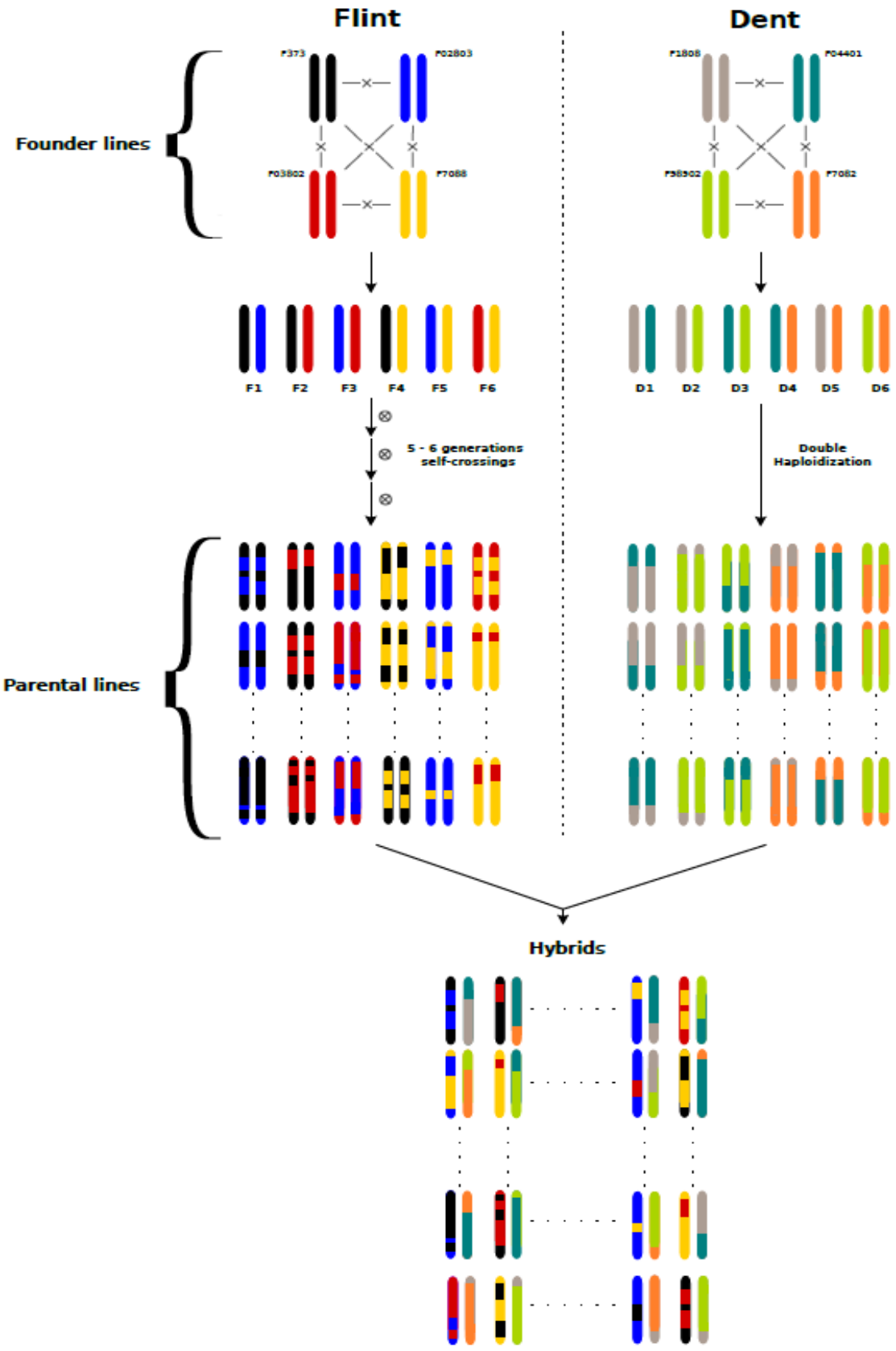

Figure 1: Schematic representation of the experimental design.

The genotyping failed for nine inbred lines. Two dent inbred lines were discarded as there was doubt on the DNA origin. Given the genotyping data, 8 inbred lines were represented under two different names and were thus renamed. Flint inbred lines with more than 25% of heterozygosity (5 lines) were not considered as well as dent lines with more than 10% of heterozygosity (20 lines) or dent lines with residual heterozygosity concentrated in some chromosome regions (15 lines). Consistency of genotypes between founder lines and parental lines was checked and off-type lines were excluded (18 flint lines and 9 dent lines). Thus only 875 dent lines and 883 flint lines were considered in further analyses.

Using the cleaned genotypic data, these 1758 inbred lines were considered for building the genetics maps: one genetic map for each of the 12 populations and one dent-flint consensus map. We followed for this the approach described in Giraud *et al.* (2014). The dent-flint consensus map was constituted of 9548 markers that were polymorphic in at least the dent or the flint design. This map had a total length of 1578.6cM and 5216 unique positions.

Field trial design and analysis

The hybrids were evaluated in a total of 8 different environments over two years (4 in 2013 and 4 in 2014) in the North of France and in Germany for four traits: silage yield (DMY in tons of dry matter per ha), dry matter content at harvest (DMC in %), plant height (6 environments) (PH in cm) and female flowering (DtSILK in days after January the first). Trials were conducted according to usual agricultural practice of the region.

The field experiments were laid out as an augmented p-rep design and were constituted of 1088 elementary plots, consisting each in two rows of 5 meters long. Most hybrids between the parental inbred lines were evaluated only once for a given environment. The founder hybrids and around 17 % of the experimental hybrids were evaluated twice. Trials were laid out in 68 incomplete blocks consisting of 16 elementary plots each with 5 to 6 plots used for repeated genotypes. 1044 hybrids were evaluated in total over the whole experimental design, out of which 951 hybrids were considered for further analyses (950 for PH and DMY), corresponding to hybrids for which both parents had correct genotypic data (821 dent parental lines and 801 flint parental lines). Outlying observations were deleted. For silage yield, data from one environment over eight were excluded as they were not correlated with the other environments.

Variance component analysis

Genetic variance decomposition was done on the single-plot performances using the ASReml-R package (Butler *et al.* 2007; R Core Team 2013). The objectives were to estimate the trait

heritabilities, evaluate the importance of the GCA and SCA components in the hybrid variance and the proportion of this variation that is due to the structure of the design in populations. The first model does not decompose the genetic value of the hybrid:

$$Y_{lii'jj'kk'txyzd} = \mu + \lambda_l + (\tau_{ii'} \times t) + H_{k(ij)k'(i'j')} \times (1-t) + \left(L_{x(l)} + C_{y(l)}\right) \times (1-d_l) + \left(B_{z(l)} \times d_l\right) + E_{liji'j'kk'txyzd} \quad (1)$$

Where $Y_{lii'jj'kk'txyzd}$ is the phenotypic value of the hybrid evaluated in the environment l at the plot located at the line x, the column y and in the block z. To distinguish between the checks and the experimental hybrids we used the parameter t. When the performance corresponds to a check hybrid between founder lines i and i', t was set to 1 and when the performance corresponds to an experimental hybrid issued from the cross between the flint parental line k (derived from the flint founder lines i and j), and the dent parental line k' (derived from the dent founder lines i' and j'), t was set to 0. In this model, μ is the intercept, λ_l is the fixed effect of the environment l, $\tau_{ii'}$ is the fixed effect of the check issued from the cross between the dent founder line i and the flint founder line i'. $H_{k(ij)k'(i'j')}$ is the genetic value of the hybrid issued from the cross between the flint parental line k (issued from the founder lines i and j) and the dent parental line k' (issued from the founder lines i' and j'). We assume that $H_{k(ij)k'(i'j')}$ are independent and identically distributed (iid) and follow a normal distribution: $H_{k(ij)k'(i'j')} \hookrightarrow N(0, \sigma_H^2), iid$. To correct for spatial heterogeneities we included in the model either a random block effect or random row column effects, depending on the environment and trait. The choice between the two models was done by analyzing independently each environment and by choosing the best correction model based on the likelihood and the repartition of the hybrid Best Linear Unbiased Predictors (BLUPs) in the field design. In the joint model, d_l is a parameter set to 1 if for the environment l the spatial effects correction chosen is a block correction, 0 else. $B_{z(l)}$ is the random effect associated to the block z in the environment l, $L_{x(l)}$ and $C_{y(l)}$are the random effects associated to the line x and the column y in the environment l, with $L_{x(l)} \hookrightarrow N\left(0, \sigma_{L_l}^2\right)$ which were assumed to be independent (id), $C_{y(l)} \hookrightarrow N\left(0, {\sigma^2}_{C_l}\right)$ id, and $B_{z(l)} \hookrightarrow N\left(0, {\sigma^2}_{B_l}\right)$ id. $E_{lii'jj'kk'txyzd}$ is the residual effect associated to the model for each environment with $E_{lii'jj'kk'txyzd} \hookrightarrow N(0, \sigma_{El}^2)$ id, and $L_{x(l)} \perp C_{y(l)} \perp B_{z(l)} \perp E_{lii'jj'kk'txyzd}$ where $\perp$ stands for independence between the random effects.

The second model considers the structure in populations of the design. It corresponds to model (1) except that $H_{k(ij)k'(i'j')}$ is decomposed into:

$$\alpha_{ij} + \beta_{i'j'} + (\alpha\beta)_{iji'j'} + H^*_{k(ij)k'(i'j')} \quad (2)$$

Where α_{ij} (respectively $\beta_{i'j'}$) is the fixed effect of the flint (respectively dent) population of origin of the flint (respectively dent) parental line k (respectively k'), $(\alpha\beta)_{iji'j'}$ is the fixed effect corresponding to the interaction between the flint and dent populations of origin of the parental lines. $H^*_{k(ij)k'(i'j')}$ is the within-population genetic value of the hybrid issued from the cross between the flint parental line k (issued from the founder lines i and j) and the dent parental line k' (issued from the founder lines i' and j') with $H^*_{k(ij)k'(i'j')} \hookrightarrow N(0, \sigma^2_{H^*})$ iid.

In the third model, the hybrid value is decomposed into GCA and SCA effects without considering the structure in populations. Thus $H_{k(ij)k'(i'j')}$ is decomposed into:

$$P_{k(ij)} + P'_{k'(i'j')} + (PP')_{k(ij)k'(i'j')} \quad (3)$$

Where $P_{k(ij)}$ (respectively $P'_{k'(i'j')}$) is the random effect of the inbred line k (respectively k') issued from the cross between the dent (respectively flint) founder lines i and j (respectively i' and j'), with $P_{k(ij)} \hookrightarrow N(0, \sigma^2_P)$ iid (respectively $P'_{k'(i'j')} \hookrightarrow N(0, \sigma^2_{P'})$ iid). These effects correspond to the dent (respectively flint) GCA of the parental lines. $(PP')_{k(ij)k'(i'j')}$ is the random effect of the interaction between the inbred line k and the inbred line k', with $(PP')_{k(ij)k'(i'j')} \hookrightarrow N\left(0, \sigma^2_{(PP')}\right)$ iid. This interaction corresponds to the SCA of the two parent lines.

In the fourth model, the hybrid value is decomposed into the population structure, the within-population GCA and the within-population SCA. Thus $H_{k(ij)k'(i'j')}$ is decomposed into:

$$\alpha_{ij} + \beta_{i'j'} + (\alpha\beta)_{iji'j'} + P^*_{k(ij)} + P^{*\prime}_{k'(i'j')} + (PP')^*_{k(ij)k'(i'j')} \quad (4)$$

Where α_{ij}, $\beta_{i'j'}$ and $(\alpha\beta)_{iji'j'}$ are defined as in model (2). $P^*_{k(i)}$, $P^{*\prime}_{k'(i'j')}$ and $(PP')^*_{k(ij)k'(i'j')}$ are the within-population equivalents of $P_{k(ij)}$, $P'_{k'(i'j')}$ and $(PP')_{k(ij)k'(i'j')}$ of model (3).

From model (1) we derived the heritability H^2 at the whole design level as: $H^2 = \frac{\sigma^2_H}{\sigma^2_H + (\frac{\sigma^2_E}{nrepH})}$

where σ^2_H is the genetic variance of the hybrids, $nrepH$ is the average number of times an experimental hybrid was evaluated in the whole design, σ^2_E is the average residual variance of the model over the different environments. The within-population heritability of the design was

calculated with a similar formula but considering the genetic variance of model (2) that takes into account the structure in populations. The percentage of within-population variance in the total genetic variance was calculated as $\%H^* = \frac{\sigma^2_{H^*}}{\sigma^2_H}$ with σ^2_H being the genetic variance of the hybrids in model (1) and $\sigma^2_{H^*}$ the within-population genetic variance of the hybrids, in model (2). The percentage of SCA in the genetic variance was calculated from model (3) as $\%ASC = \frac{\sigma^2_{(PP')}}{\sigma^2_P+\sigma^2_{P'}+\sigma^2_{(PP')}}$. The within-population percentage of SCA was calculated from model (4) using a similar formula but considering the within-population variances of the flint GCA, the dent GCA and the SCA effects.

Computation of adjusted means and correlations between traits

QTL detection was based on the least square-means (ls-means) of each experimental hybrid. To obtain these ls-means, we first corrected the individual single plots performances by the BLUPs of the field effects obtained with model (2). Then for each trait, least square means of hybrids were derived from the fixed effect model, considering jointly the experimental hybrids and the check hybrids between the founders: $Y^*_{hrl} = \mu + \lambda_l + \gamma_h + E_{hrl}$ where Y^*_{hrl} is the performance corrected for the field effects of the r-repetition of the hybrid h in the environment l, μ is the intercept, λ_l is the fixed effect of the environment l, γ_h is the hybrid genetic effect considered as fixed in this model. E_{hrl} is the residual of the model for each environment with $E_{hrl} \hookrightarrow N(0, \sigma^2_{El})$ id. Correlations between the different traits were calculated based on these ls-means.

QTL detection

Three models were used for QTL detection depending on the type of allele coding that was considered. In each model, we took into account the structure of the design in populations and included random genetic effects corresponding to the parents of the hybrids to take into account the fact that some of the hybrids derived from the same parental lines.

The “Founder alleles” model makes the assumption that each of the eight founder lines carried a different allele.

$$y = 1.\mu + A.\alpha + B.\beta + C.(\alpha\beta) + X_{F\_d}.\gamma_{F\_d} + X_{F\_f}.\theta_{F\_f} + X_{F\_d\,F\_f}.(\gamma\theta)_{F\_d\,F\_f} + Z_d.u_d + Z_f.u_f + e \quad \text{(Q1)}$$

Where y is a (N ×1) vector of the ls-means of the hybrids with N being the number of experimental hybrids phenotyped for the considered trait; μ is the intercept, 1 is a (N ×1) vector

of 1. α (respectively β) is a (6 ×1) vector of the fixed effects of the dent (respectively flint) populations of origin of the dent (flint) parental line, A (respectively B) is the (N ×6) design matrix of 0-1 that linked each hybrid to the dent (respectively flint) population of its dent (respectively flint) parental line, ($\alpha\beta$) is a (36 ×1) vector of the fixed interaction effects between the dent and flint populations of parental lines, C is the corresponding design matrix. u_d (respectively u_f) is a (N_d ×1) (respectively (N_f ×1)) vector of the random effects of the N_d dent (respectively N_f flint) parents, with $u_d \hookrightarrow N\ (0,\ \sigma^2_{u_d})$, iid (respectively $u_f \hookrightarrow N\ (0,\ \sigma^2_{u_f})$, iid). Z_d is the $(N\ \times\ N_d)$ design matrix that relates the N hybrids with the N_d different dent parents and Z_f is a $\left(N\ \times\ N_f\right)$ design matrix that relates the N hybrids to the N_f different flint parents. e is a (N ×1) column vector of the residuals of the model with $e\ \hookrightarrow N\ (0,\ \sigma^2_e)$, iid. The QTL effect was decomposed into three terms: γ_{F_d}, θ_{F_f} and $(\gamma\theta)_{F\_d\,F\_f}$. γ_{F_d} (respectively θ_{F_f}) is the (4 x 1) vector of the allelic effects at the marker associated with each dent (flint) founder line. These effects correspond to the GCA effects of the QTL. For each marker, X_{F_d} (respectively X_{F_f}) is a (N × 4) matrix of the probabilities that the hybrid received its dent (respectively flint) allele from each of the four dent (respectively flint) founder lines. $(\gamma\theta)_{F\_d\,F\_f}$ is the (16 x 1) line vector of the 16 levels of the interactions or SCA between the founder alleles, $X_{F\_d\,F\_f}$ is a (N ×16) matrix corresponding to the Hadamard product of X_{F_d} and X_{F_f}. As the sum of probabilities for each allele equals 1, this model has three degrees of freedom (df) for the additive effect of the founder alleles in each group and nine df for the interaction.

At each marker, probabilities of the four dent (respectively flint) founder lines were inferred using PlantImpute (Hickey *et al.* 2015) using 10 iterations for the 9548 mapped markers.

The "SNP within-group" model considered the observed alleles at SNP received from the parental inbred lines, assuming different effects in the two heterotic groups. This model assumes that two inbred lines from the same group that share the same allele at SNP are IBD.

$$y = 1.\mu + A.\alpha + B.\beta + C.(\alpha\beta) + X_{SNP\_d}.\gamma_{SNP\_d} + X_{SNP\_f}.\theta_{SNP\_f} + X_{SNP\_d\,SNP\_f}.(\gamma\theta)_{SNP\_d\,SNP\_f} + Z_d.u_d + Z_f.u_f + e \quad \text{(Q2)}$$

Where y, μ, , A, α, B, β, C, ($\alpha\beta$), Z_d, Z_f, u_d, u_f and e are defined as in model (Q1). The QTL effect was decomposed into three terms: γ_{SNP_d}, θ_{SNP_f} and $(\gamma\theta)_{SNP\_d\,SNP\_f}$. γ_{SNP_d}(respectively θ_{SNP_f}) is the GCA effect of the dent (respectively flint) minor allele, X_{SNP_d} (respectively X_{SNP_f}) is a (N ×1) vector of marker genotypes for the dent (respectively flint) parent of the hybrid, coded as 0 for homozygotes for the major allele, 1 for homozygotes

for the minor allele and 0.5 for heterozygotes. $(\gamma\theta)_{SNP\_d\ SNP\_f}$ is the SCA effect between the minor SNP marker alleles of each group, $X_{SNP\_d\ SNP\_f}$ is a (N ×1) column vector corresponding to the Hadamard product of X_{SNP_d} and X_{SNP_f}. This model as one df for the allelic effect of each group and one df for the interaction.

Missing genotypes at the parental level were imputed with Beagle v3.0. (Browning and Browning 2007). Imputations were done within each population after adding the founder lines in the dataset. Phasing of the flint lines and of the founder lines was done at the same time than missing genotypes imputation.

The "Hybrid genotype" model considered that the QTLs effects are the same in both heterotic groups. It decomposed the hybrid effect into terms of additivity and dominance.

$$y = 1.\mu + A.\alpha + B.\beta + C.(\alpha\beta) + X_a.\omega + X_d.\delta + Z_d.u_d + Z_f.u_f + e \quad \text{(Q3)}$$

Where y, μ, , A, α, B, β, C, $(\alpha\beta)$, Z_d, Z_f, u_d, u_f and e are defined as in model (Q1). The QTL effect was decomposed into two terms ω and δ which are respectively, the additive and dominance effect at the marker. X_a is a (N ×1) column vector coded in -1, -0.5, 0, 0.5, 1 indicating the genotype of the hybrid at the marker level. X_a equals -1 when the hybrid is homozygous for the major allele, 1 when the hybrid is homozygous for the minor allele, 0 if its parents are both homozygous but for a different allele, -0.5 (respectively 0.5) when the dent parent is homozygous for the major (respectively minor) allele and the flint parent is heterozygous. X_d is a (N ×1) column vector coded in 0, 0.5, 1. X_d equals 0 if both parents of the hybrid are homozygous for the same allele, 0.5 when the dent parent is homozygous and the flint parent is heterozygous or conversely, and 1 when both parents are homozygous but for different alleles. This model has one df for the additive effect and one df for the dominance effect.

For this model, as for the "SNP within-group" model missing marker data were replaced by imputed genotypes.

QTL detection was performed with ASReml-R (Butler *et al.* 2007). To avoid identifiability problems for the "SNP within-group" and the "Hybrid genotype" models, QTL detection was performed on the 4758 mapped markers which were polymorphic (MAF superior to 5%) in both heterotic groups whereas for "Founder alleles" model it was performed on the 9548 mapped markers. For each model, we considered a 5% genome-wide significance threshold based on the number of efficient markers (Gao *et al.* 2008).

The total effect at the marker was tested using the "group" function. After the first initial scan along the genome, a multi-marker procedure was implemented using a forward and backward

marker selection. In the forward stage, the most significant maker (based on the total locus effect) was added to the model until no-more marker had a significant total locus effect at the 5% genome-wide risk level. Then, markers with a GCA or SCA (or additivity or dominance) effect superior to the threshold were added to the model. Finally, in a backward stage, we removed step by step each effect that was not significant in the joint QTL model until we only kept markers for which the total effect or one of its components (GCA/SCA or additivity/dominance) was superior to the threshold.

The percentage of phenotypic variance explained by the population effects R^2_{pop} was calculated according to Nakagawa and Schielzeth (2013). To estimate the percentage of variance explained by the detected QTLs (R^2_{QTL}), we used a modified version of the R^2 presented by Nakagawa and Schielzeth (2013) with marker effects orthogonalized by population structure. From these two parameters we estimated the percentage of within-population phenotypic variance explained by the QTLs as $R^{2*}_{QTL} = \frac{R^2_{QTL}}{1- R^2_{pop}}$. We also estimated the individual R^2 of each QTL after orthogonalyzing its effect by the population structure and the effects of the other QTLs.

To evaluate the quality of prediction of these models, we also performed a cross-validation approach using four fifth of the data for estimating the population and the QTL effects, and predicting the values of the hybrids of the last fifth. Sampling was stratified by the structure in populations and was repeated 100 times. The squared correlation between the predicted and true hybrid values of the fifth fold were estimated. This procedure was conducted (i) without taking into account SCA/dominance at the QTL level and (ii) taking it into account for QTL for which it was significant at a 5% individual risk level. Percentage of variance explained by the population effects were estimated following the same cross-validation approach.

RESULTS

Genetic variance analysis

We observed large and significant hybrid variances for all traits (Table 1). Broad sense heritabilities at the design level were high for all traits: between 0.814 (DMY) and 0.892 (DMC) (Table 2).

For all the traits except DMC, the dent and flint population effects were both significant at a 5% level risk, whereas the interaction was not. For DMC, the effects of the flint population and

of the interaction between the dent and flint populations were significant whereas the effect of the dent population was not (result not shown). This is in agreement with the smaller variation of DMC performances among the dent founder lines compared to the flint ones. Even if the population effects were significant, the within-population hybrid genetic variance was large (Table 1). It represented between 63.1% (PH) and 86.7% (DtSILK) of the total hybrid genetic variance (Table 2).

Accordingly, for all traits, within-population heritabilities remained high and varied between 0.767 for grain yield and 0.876 for female flowering (Table 2).

The decomposition of the hybrid variance in GCA and SCA (models 3 and 4) showed that most part of the hybrid variation was due to GCA. When the structure in populations was not included in the model the flint and dent GCA variances were different for DMC, DMY and PH, whereas they were of similar magnitude for DtSILK (Table 1). After taking into account population structure, the within-population flint and dent GCA variances were of the same magnitude for all traits. For the model which did not take into account the structure in populations, SCA represented between 11.7% (DMC) and 17.4% (DMY) of the hybrid genetic variance (Table 2). The proportion of SCA was higher when considering the within-population hybrid variation (Table 2). It represented about 20% of the within-population genetic variance for all traits but PH for which it was lower (Table 2). It has to be noted that standard deviations for SCA variances were large (Table 1) certainly due to the small number of inbred lines that contributed to more than one hybrid in our design.

Table 1: Variance decomposition of the phenotypic data for the four variance decomposition models, for the different traits: dry matter content DMC, dry matter yield DMY, female flowering time DtSILK, plant height PH. For each genetic variance, precision corresponding to 1.96 x the standard error of the estimated variance is indicated

	Hybrid (1)	**Population structure + Hybrid (2)**	**Flint GCA + Dent GCA + SCA (3)**	**Population structure + Flint GCA + Dent GCA + SCA (4)**
DMC				
Hybrid variance	3.68 ± 0.37	2.46 ± 0.27		
Flint GCA variance			2.31 ± 0.51	0.98 ± 0.44
Dent GCA variance			0.92 ± 0.53	0.96 ± 0.44
SCA variance			0.43 ± 0.54	0.53 ± 0.49
Residual variances (range)	1.44 ± 0.18 : 6.15 ± 0.60	1.52 ± 0.19 : 6.05 ± 0.59	1.45 ± 0.18 : 6.13 ± 0.59	1.52 ± 0.18 : 6.04 ± 0.58
DMY				
Hybrid variance	1.24 ± 0.14	0.94 ± 0.11		
Flint GCA variance			0.28 ± 0.20	0.30 ± 0.19
Dent GCA variance			0.74 ± 0.19	0.44 ± 0.17
SCA variance			0.22 ± 0.22	0.20 ± 0.20
Residual variances (range)	1.01 ± 0.12 : 3.53 ± 0.35	1.02 ± 0.12 : 3.56 ± 0.34	1.00 ± 0.11 : 3.53 ± 0.35	1.02 ± 0.12 : 3.56 ± 0.35
DtSILK				
Hybrid variance	2.38 ± 0.24	2.06 ± 0.21		
Flint GCA variance			0.74 ± 0.38	0.51 ± 0.37
Dent GCA variance			1.26 ± 0.35	1.09 ± 0.33
SCA variance			0.39 ± 0.37	0.47 ± 0.37
Residual variances (range)	0.96 ± 0.11 : 6.12 ± 0.57	0.96 ± 0.11 : 6.11 ± 0.57	0.96 ± 0.11 : 6.13 ± 0.57	0.96 ± 0.11 : 6.12 ± 0.57
PH				
Hybrid variance	154.7 ± 15.7	97.6 ± 10.7		
Flint GCA variance			42.9 ± 22.9	38.4 ± 17.4
Dent GCA variance			92.8 ± 21.6	45.1 ± 16.0
SCA variance			19.3 ± 23.2	13.4 ± 17.3
Residual variances (range)	49.7 ± 6.5 : 244.2 ± 23.7	46.4 ± 6.0 : 243.6 ± 23.5	50.0 ± 6.5 : 243.6 ± 23.6	46.7 ± 6.0 : 243.2 ± 23.5

Table 2: Adjusted means of experimental hybrids and founder line hybrids and synthetic parameters on the experimental hybrid variation for the different traits (dry matter content DMC, dry matter yield DMY, female flowering time DtSILK, plant height, PH).

	DMC	**DMY**	**DtSILK**	**PH**
Adjusted means				
Experimental hybrids (average, min:max)	33.0	16.0	211.5	247.8
	(25.7 : 41.3)	(11.8 : 20.2)	(206.7 : 217.9)	(203.9 : 283.2)
Founder line hybrids (average, min:max)	33.5	16.5	211.6	251.3
	(30.2 : 36.1)	(14.8 : 18.4)	(209.0 : 213.6)	(228.8 : 275.1)
Within-population variance				
%Within_var	66.9	75.5	86.7	63.1
Heritabilities				
H^2	0.892	0.814	0.890	0.877
H^2* (within-population)	0.847	0.767	0.876	0.817
Part of SCA in the hybrid variance				
%SCA	11.7	17.4	16.3	12.4
% SCA* (within-population)	21.5	21.0	22.6	13.8

Adjusted means and correlations between traits

Adjusted means of the experimental hybrids were on average of 16.0t.ha^{-1} (min 11.8, max 20.2) for DMY, of 33.0% (min 25.7 max 41.3) for DMC, 211.5 days for DtSILK (min 206.7 max 217.9) and 247.8cm (min 203.9 max 283.2) for PH (Table 2). The experimental hybrids showed for all traits a variation that exceeded the one found for the hybrids between founder lines (Table 2). Adjusted means of hybrids between founder lines were on average 16.5t.ha^{-1} for DMY, 33.5% for DMC, 211.6 days for DtSILK and 251.3cm for PH (Table 2). These values were slightly higher than the average values of the experimental hybrids suggesting that recombination events broke some epistatic interactions. However, differences were small, suggesting a limited impact of epistasis. Founder line hybrids involving F98902 had higher yield, whereas those involving F7088 had higher DMC, consistent with known information on these founder lines. DMY was positively correlated to PH (0.64) and DtSILK (0.57) and negatively to DMC (-0.28). DMC was also negatively correlated to PH (-0.28) and DtSILK (-0.55). These correlations are consistent with those usually observed for these traits.

QTL detection

The thresholds at a 5% genome-wide level used for QTL detection were set to a –log(pvalue) equal to 4.53 for the "Hybrid genotype" model, 4.40 for the "SNP within-group" model and 3.84 for the "Founder alleles" model.

We detected QTLs for all trait × model combinations. For a given trait, the number of detected QTLs varied with the model. Between nine (DtSILK, "Founder alleles" model and DMY, "SNP within-group" model) and 16 QTLs were detected (DtSILK, "Hybrid genotype" model). In total for the four studied traits, the "SNP within-group" model and the "Hybrid genotype" model detected more QTLs, respectively 51 and 54, than the "Founder alleles" model which detected in total only 42 QTLs (Table 3). Nevertheless, the "Founder alleles" model detected more QTLs for DMY.

Table 3: QTL detection results with the different detection models for the different traits (dry matter content DMC, dry matter yield DMY, female flowering time DtSILK and plant height PH). For each method and trait we indicated the number of QTLs detected and between brackets the number of QTLs showing significant SCA effects at a 5% level risk, the proportion the of phenotypic variance (R^2_{QTL}, in %) and of the within-population phenotypic variance (R^{2*}_{QTL}, in %) explained by the detected QTLs (with and without including dominance/SCA effects in the model). The percentage of variance explained by the population effect are also indicated (R^2_{pop}). The total number of detected QTLs and the average percentages of variance explained over the different traits are also showed.

Trait	Model	Nb	R^2_{pop}	Without SCA		With SCA	
				R^2_{QTL}	R^{2*}_{QTL}	R^2_{QTL}	R^{2*}_{QTL}
DMC	Founder alleles	10 (4)	32.4	27.6	40.9	32.4	47.9
	SNP within-group	12 (2)	32.4	25.5	37.7	26.1	38.6
	Hybrid genotype	14 (1)	32.4	25.6	37.9	26.2	38.8
DMY	Founder alleles	12 (5)	21.9	27.7	35.5	34.2	43.9
	SNP within-group	9 (0)	21.9	20.3	26.0	20.5	26.3
	Hybrid genotype	11(3)	21.9	19.7	25.2	20.9	26.8
DtSILK	Founder alleles	9 (2)	15.0	31.4	36.9	36.7	43.2
	SNP within-group	15(0)	15.0	37.3	43.9	37.6	44.3
	Hybrid genotype	16(3)	15.0	34.1	40.2	35.6	41.9
PH	Founder alleles	11(2)	33.8	26.6	40.2	30.7	46.4
	SNP within-group	15(4)	33.8	24.7	37.3	26.6	40.2
	Hybrid genotype	13(2)	33.8	20.4	30.8	21.2	32.0
Total	Founder alleles	42 (13)	25.8	28.3	38.4	33.5	45.3
	SNP within-group	51 (6)	25.8	26.9	36.2	27.7	37.4
	Hybrid genotype	54 (9)	25.8	24.9	33.5	26.0	34.9

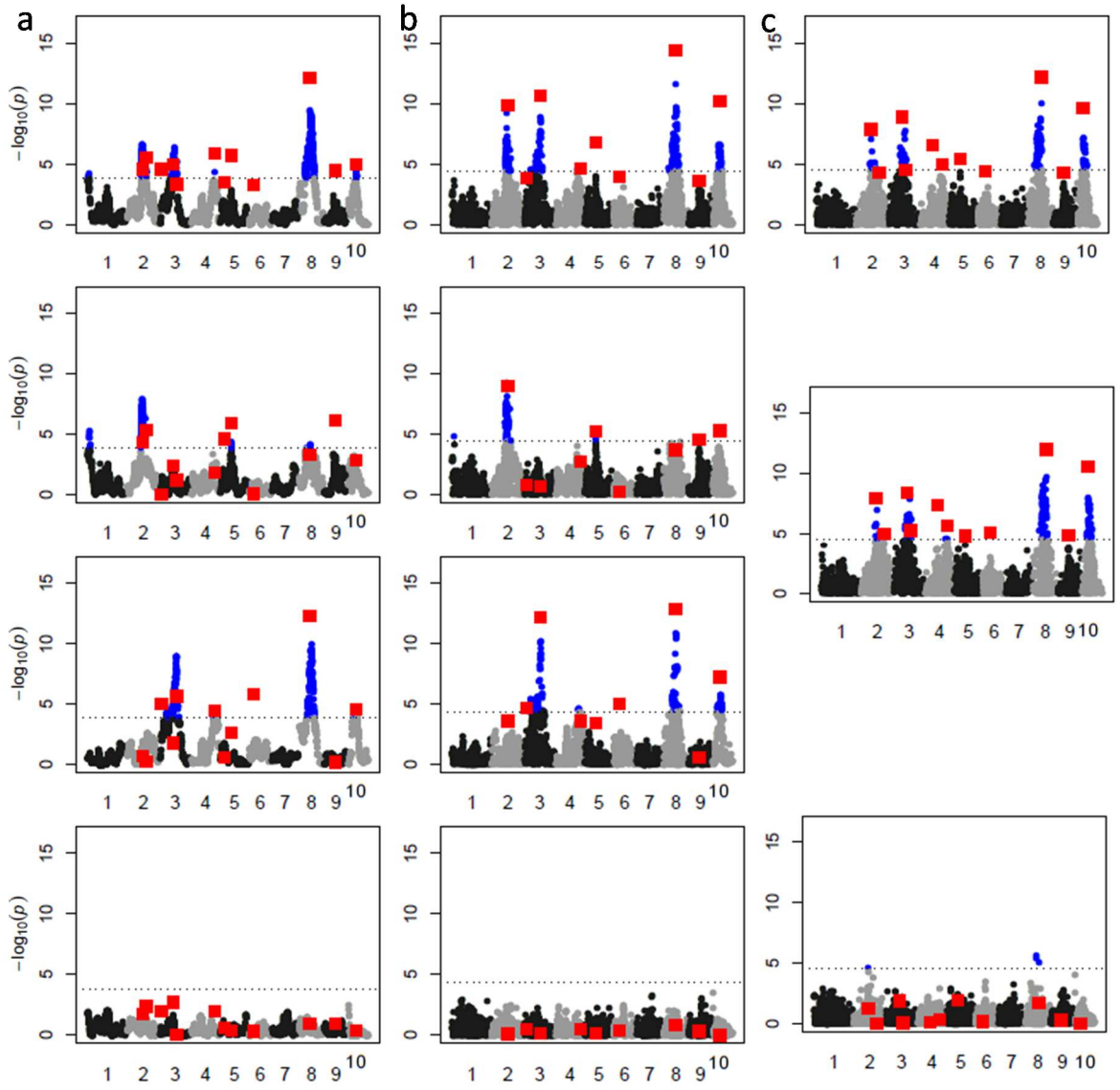

Figure 2: -log(p-value) curves of QTL detection for Dry Matter Yield (DMY) with (a) the "Founder alleles" model, (b) the "SNP within-group" model, (c) the "Hybrid genotype" model. The chromosome number is indicated on the abscissa. For the "Founder alleles" (a) and the "SNP within-group" (b) models, the graphic at the top corresponds to the global effects at the markers. The other graphics correspond to the different components of the decomposed effects: from the top to the bottom: the flint GCA, the dent GCA, the SCA. For the "Hybrid genotype" model, the graphic at the top corresponds to the global effect at the markers, the middle part to the additive effect of the markers and the bottom part to the dominance effect of the markers. The grey and black dots correspond to the significance levels of tests below the threshold at the first step of the forward procedure, the blue dots correspond to those that were above the threshold. The red squares correspond to the –log(p-value) of the QTLs that were included in the final multi-locus model.

Globally the different models detected QTLs in the same chromosome regions (Figure 2). To compare the QTL detected by the different models (Figure 2, Figure 3, Supplementary Tables S1, S2, S3), we considered arbitrarily that QTLs detected at positions closer than 10cM of each

other were identical. With this assumption, 16 QTLs were detected with the three models (for instance at 65 – 65.8cM on chromosome 5 for DMY). Some were detected with only two models: 11 with the "Hybrid genotype" and the "SNP within-group" models, (for instance at 113.4cM on chromosome 3 for DMC), 2 with the "Hybrid genotype" and the "Founder alleles" models (for instance at 86 – 89.3cM on chromosome 2 for DMY), 7 with the "SNP within-group" and the "Founder alleles" models (for instance at 25 – 29.9cM on chromosome 6 for DMY). The other QTLs were specific to one model for instance the DtSILK QTL detected with the "Founder alleles" model at 7.7cM on chromosome 1, the DMC QTL detected with the "SNP within-group" model at 70cM on chromosome 6, and the DMY QTL detected with the "Hybrid genotype" model at 74.5cM on chromosome 4.

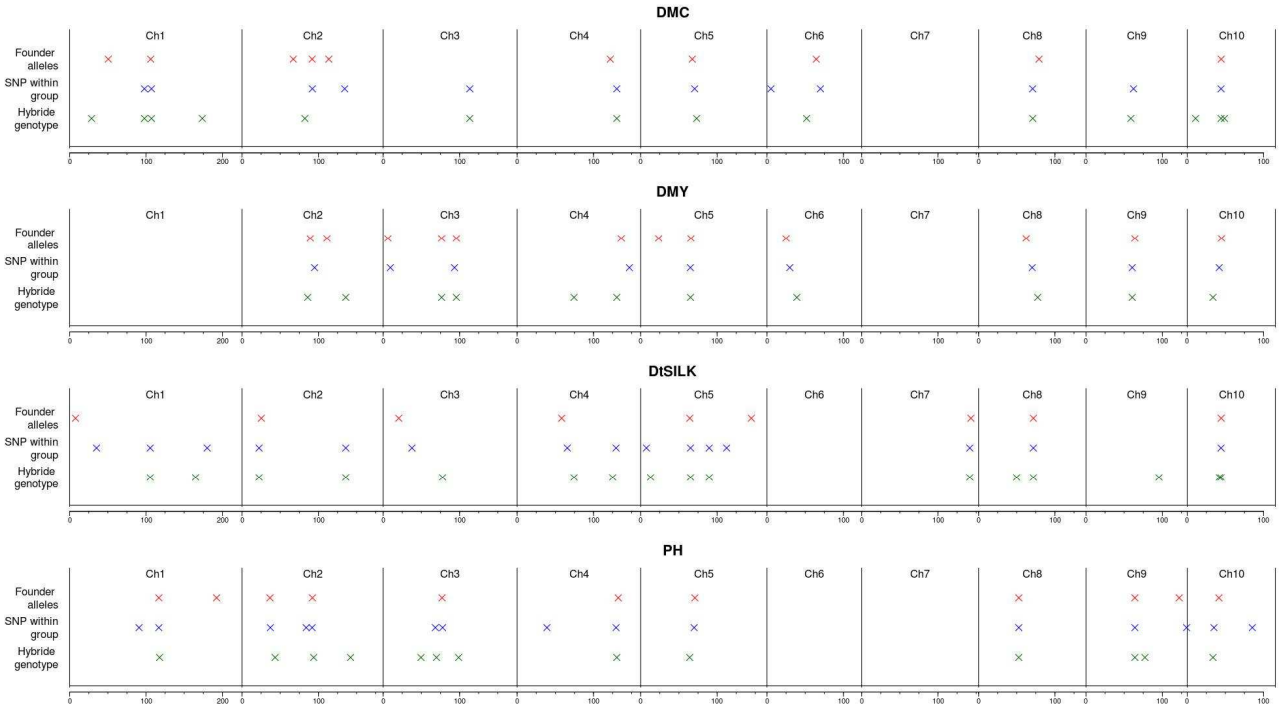

Figure 3: Synthesis of the positions of the detected QTLs for the four studied traits (DMC indicates dry matter content, DMC; dry matter yield, DMY; female flowering time DtSILK, and plant height, PH) and the different models. The QTLs detected with the "Founder alleles", "SNP within-group" and "Hybrid genotype" models are indicated respectively with red, blue and green crosses.

For all models, the majority of the QTLs had small effects (they explain less than 5% of the variation, see Supplementary Tables S1, S2, S3). The only exceptions are the PH QTL detected on chromosome 3, which explained 5.3% of the variance for the "SNP within-group" model and a QTL detected on chromosome 10 at 44.5cM that was detected by the three models and explained around 8% of the variance for DMC and 13% of the variance for DtSILK. This QTL

region was also involved but with a smaller effect on DMY and PH. Other QTL regions showed pleiotropic effects on different traits, for instance the one around 105.5 - 117.8cM on chromosome 1 which had an effect on DMC, DtSILK and PH. Other regions were specific to one trait (141.6 – 143.1cM on chromosome 7 for DtSILK, 25 – 39.1cM on chromosome 6 for DMY) (Figure 3, Supplementary Tables S1, S2, S3).

For each QTL we tested the level of significance of their GCA/SCA or additive/dominance components (Supplementary Tables S1, S2, S3). For the "Hybrid genotype" model, none of the detected QTL showed significant dominance effect at a 5% genome-wide level. Some QTLs had significant dominant effect in the initial QTL detection scan but their effects were no more significant in the final multilocus model (see Figure 2 for DMY). However, nine QTLs were significant for dominance with an individual risk at 5%: one for DMC, three for DMY and DtSILK (among which one was significant with an individual risk at 1%) and two for PH (Table 3 and Supplementary Table S1). For the "SNP within-group" model, considering the 5% genome-wide threshold, nine QTLs were significant for both GCA effects, 23 only for the dent GCA effect and 15 only for the flint one. No specific QTL for SCA was detected at a 5% genome-wide risk level, but six QTLs had a significant SCA effect at the 5% individual risk level: two for DMC and four for PH (Table 3, Supplementary Table S2). For the "Founder alleles" model, considering the 5% genome-wide threshold, seven QTLs were significant for both GCA effects, 20 only for the dent GCA effect and 12 only for the flint GCA effect. No specific QTL for SCA was detected at a 5% genome-wide risk level, but 13 QTLs had a significant SCA effect at the 5% individual risk level: four for DMC, five for DMY, two for DtSILK and two for PH. Among them three had significant SCA effects at a 1‰ risk level (Table 3, Supplementary Table S3). QTLs showing significant SCA were located all over the genome.

However, one region on chromosome 2, between 82.3cM and 135.8cM stands out for presenting SCA for both DMC and DMY (Supplementary Tables S1, S2, S3). All models detected QTLs in this region, and five over seven of the QTLs detected with the "Founder alleles" and the "SNP within-group" models had a significant effect on SCA at a 5% individual risk level.

For QTLs detected at close position with several models, GCA/additive QTL effects of the founder lines were consistent between models (result not shown). Interestingly, no founder line presented favorable alleles at all detected QTLs. For instance, considering the dent and flint GCA effects for DMY of the QTLs detected with the "Founder alleles" model showed that all founder lines presented positive and negative effects (Figure 4). This is consistent with the

transgressions observed in the experimental hybrid populations compared to the founder hybrids. It would be interesting to identify the best hybrid ideotype based on both GCA and SCA effects and to carry out in each group a marker-assisted selection scheme to obtain the parental lines corresponding to this ideotype.

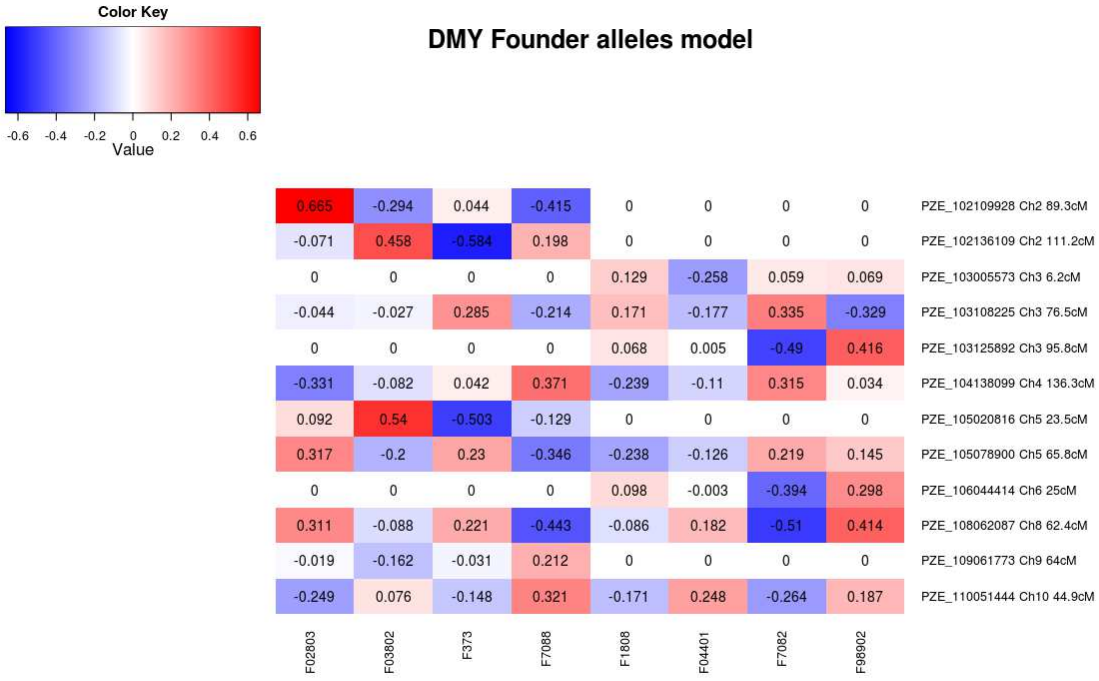

Figure 4: GCA effects for the founder lines for the QTLs detected with the "Founder alleles" model for DMY. Allelic effects are centered on zero for the dent founder lines (F1808, F04401, F7082 and F98902) and for the flint founder lines (F02803, F03802, F373 and F7088). QTLs presenting a dent (respectively flint) GCA effect not significant at a 5% individual risk level had their dent (respectively flint) GCA effects set to zero.

The detected QTLs explained jointly between 19.7% (DMY, "Hybrid genotype" model, without dominance) and 37.6% (DtSILK, "SNP within-group" model, with SCA) of the total phenotypic variance and between 26.8% and 47.1% of the within-population phenotypic variance. The model which explained the largest fraction of the phenotypic variance was the "Founder alleles" model for DMY, DMC and PH and the "SNP within-group" model for DtSILK. The increase in percentage of explained phenotypic variance when taking into account dominance/SCA is low for the "SNP within-group" model (+ 0.2 for DMY to +1.9 for PH) and for the "Hybrid genotype" model (+0.6 for DMC to +1.5 for DtSILK) whereas is it is more important for the "Founder alleles" model (+4.1 for PH to +6.5 for DMY) (Table 3).

Table 4: Cross-validation estimates of the quality of prediction of different models (average R^2 and its standard deviation, sd). For the different traits (dry matter content DMC, dry matter yield DMY, female flowering DtSILK and plant height PH) we considered model only including population effects or models including population effects and the QTL effects. For these later models prediction were based on GCA /additive effects only or on models considering also SCA/dominant effects significant at a 5% risk level. The number of QTLs detected with each model for each trait is indicated (Nb) as well as the number of QTLs significant for SCA/dominance with a 5% individual risk level (between brackets).

		DMC		**DMY**		**DtSILK**		**PH**	
		Nb	R^2 (%)	Nb	R^2 (%)	Nb	R^2 (%)	Nb	R^2 (%)
Population effects			28.4 sd 4.18		17.1 sd 4.16		10.4 sd 2.97		29.2 sd 4.35
Pop + QTLs "Founder alleles" model	GCA	10	53.4 sd 4.01	12	39.2 sd 4.84	9	37.7 sd 4.69	11	53.2 sd 4.37
	GCA+ SCA	(4)	54.0 sd 3.72	(5)	39.4 sd 4.83	(2)	39.1 sd 4.72	(2)	52.7 sd 4.64
Pop + QTLs "SNP within-group" model	GCA	12	54.4 sd 4.06	9	36.7 sd 4.64	15	47.6 sd 4.68	15	54.0 sd 4.93
	GCA+ SCA	(2)	54.7 sd 4.05	(0)	-	(0)	-	(4)	55.2 sd 4.87
Pop + QTLs "Hybrid genotype" model	Additivity	14	55.2 sd 3.98	11	37.3 sd 4.50	16	45.4 sd 4.88	13	50.6 sd 4.83
	Additivity+ dominance	(1)	55.3 sd 4.02	(3)	38.0 sd 4.52	(3)	46.4 sd 4.72	(2)	51.0 sd 4.78

Cross-validations were performed in order to evaluate the quality of prediction of the QTL models. This was done to eliminate potential bias in the R^2 values of Table 3 that were computed on the data also used to estimate QTL parameters, potentially advantaging models with high number of parameters. Population effects explained between 10.4% (DtSILK) and 29.2% (PH) (Table 4) of the phenotypic variance. Combination of the population effects and the detected QTLs without dominance/SCA, explained between 36.7% (DtSILK, "SNP within-group" model) and 55.2% (DMC, "Hybrid genotype" model) of the total phenotypic variance. Taking into account the dominance/SCA for the QTLs for which it was significant at a 5% individual risk had a small impact on the quality of prediction of the model: from negative (-0.5 for PH for the "Founder alleles" model) to low (+1.4 for DtSILK for the "Founder alleles" model). Considering or not dominance/SCA, the best predictions were obtained with the "SNP within-group" model for DMC and DtSILK, the "Founder alleles" model for DMY and the "Hybrid

genotype" model for DMC.

DISCUSSION

The objective of this work was to evaluate the interest of using a factorial design between two multiparental connected designs to detect QTLs for hybrid values. Using this design, we first decomposed the hybrid genetic variance into terms of population effects, GCAs and SCA components and then performed QTL detection using three different models.

Genetic variance components

We observed important variation among hybrids for all the observed traits with transgressions beyond the variation of the hybrids between founder lines. The fact that part of the parental inbred lines contributed to more than one hybrid allowed us to estimate SCA/ GCA variance components. Most of the hybrid variance was due to GCA but about 20% of the within-population genetic variance of our design was due to SCA for all traits except PH for which it was smaller. To our knowledge, few studies estimated SCA variances on European silage maize, so that the number of possible comparisons is limited. Argillier *et al.* (2000) found small but significant SCA for DMC and DMY between the dent and the flint group but did not estimate variances explained by each component. Grieder *et al.* (2012) evaluated GCA/SCA variances based on test-cross values of dent lines using flint hybrids as testers. SCA explained less than 10% of the hybrid variance for DMY and DMC but, as explained by the authors, these small values might be due to the use of a small set of related hybrid testers. Our study cannot be compared directly to the estimations of SCA/GCA components obtained for dent-flint hybrids designs evaluated for grain yield (Schrag *et al.* 2006; Fisher *et al.* 2008; Schrag *et al.* 2009; Schrag *et al.* 2010; and more recently Technow *et al.* 2014). In these studies authors consistently found that SCA usually explained less than 10% of the hybrid variation for grain yield and grain moisture at harvest.

Thus, other studies also clearly showed a predominance of GCA over SCA but they found in general less important SCA variance relatively to GCA variances than in our study. The predominance of GCA over SCA is in agreement with the fact that, in the hypothesis of absence of epistasis, the ratio of dominance (major component of SCA) over additive (major component of GCA) variance is expected to be low for hybrids obtained by crossing two divergent populations (Reif *et al.* 2007) such as the two heterotic groups considered in this study which

diverged 500 years ago (Tenaillon and Charcosset 2011). Reciprocal selection is expected to lead to the fixation in each group of alternative alleles for loci showing dominance effects that are thus absorbed in the mean value of the hybrids. This results in a decrease of the SCA variance compared to the GCA variance over time and consequently a difficulty for detecting dominance effects that become confounded with additive effects (for instance if one QTL segregates in one group but is fixed in the other one) (Technow *et al.* 2014). Most already published studies were based on factorials derived from inbred lines that passed through a selection stage based on their test-cross values on testers. This selection might have possibly retained lines with similar SCA pattern which might have contributed to lower SCA variance compared to GCA. In our design, the hybrid parental lines are derived without selection from the founder lines. They thus represent the whole allelic diversity available in each population, giving more opportunity to observe SCA. Even if the proportion of SCA is limited compared to GCA (20% versus 40% for each GCA), it is not negligible and might be sufficient to blur the estimation of GCA using only a small number (one or two) of tester lines from the opposite group as it is usually done in breeding programs.

QTL detection

One of the advantages of our design is that several sources of information can be used for QTL detection. The structure in biparental populations of the hybrid parents allowed us to trace founder alleles down to the hybrids and thus to perform a QTL detection based on linkage information (the “Founder alleles” model). This QTL detection model can be seen as an extension of the model used to detect QTLs in test-crossed connected multiparental designs (as done in Rebaï *et al.* 1997; Blanc *et al.* 2006 or Giraud *et al.* 2014) to the hybrid case. In addition to linkage analysis, we also used directly information provided by SNP markers to perform QTL detection (with the “SNP within-group” and “Hybrid genotype” models) with models close to the ones used for association mapping (LD mapping) except that we used the known structure in populations of the design instead of a kinship matrix to correct for false positives. The three models used for QTL detection performed differently depending on the trait and on the chromosome region considered. They make different assumptions in terms of allelic effects and are thus expected to perform differently depending on the actual distribution of QTL effects. The “Hybrid genotype” model considers only two degrees of freedom per marker and is thus expected to be more powerful that the other models which have more degrees of freedom to take into account per marker. However it makes strong hypotheses considering that the QTLs effects are biallelic, that they are the same in both heterotic groups and that the marker-QTL

phase is also conserved between groups. The other models have more degrees of freedom but make fewer hypotheses: the QTL effect is assumed different in the two heterotic groups for the "SNP within-group" model and each founder lines as a different allele at the QTL for the "Founder alleles" model.

Even if in total more QTLs were detected with the "Hybrid genotype" model than with the other models, this model almost never explained a larger part of the genetic variance (considering direct adjustment to the data or cross-validations). The strong hypotheses that it makes counterbalanced its potential interest in terms of power. This is consistent with the fact that the other models clearly detected QTLs specific to dent or to flint GCA. Thus the "Founder alleles" and the "SNP within-group" seem better adapted to QTL detection in such a design. This is in agreement with Giraud *et al.* (2014) who found different QTLs in each heterotic group. This confirms the long-term separation of the two heterotic groups (Tenaillon and Charcosset 2011) and shows that reciprocal selection structured the diversity of these two groups in order to fully benefit from heterosis by complementation of alleles in the hybrids. Van Eeuwijk *et al.* (2010), in a maize factorial obtained by crossing two other heterotic groups, also found different GCA QTLs for ear height in the two groups. Hence, complementarity of GCA effects in heterotic groups seems to be an important component of hybrid performance and an efficient hybrid breeding program should target both common and different QTL in the heterotic groups.

Even if SCA represented around 20% of the within-population genetic variance (except for PH), we did not detect QTLs specific to SCA at a 5% genome-wide risk level. We nevertheless detected dominance and/or SCA effects significant at a 5% individual risk level for some QTLs (at a 1‰ individual risk level for three of them). Cross-validation results showed that adding these SCA QTLs effects to the models only slightly improved in the best cases the quality of prediction of hybrid values, suggesting that these moderate QTLs SCA effects may not be well estimated. These results contrast strongly with those of Schön *et al.* (2010) and Larièpe *et al.* (2012) who found a majority of QTL with large dominance effects for grain yield. An important feature of these studies is that they involve hybrids with a high level of inbreeding, contrary to our present study in which all hybrids evaluated are issued from unrelated parents. This suggests that, in the absence of inbreeding, SCA is likely due too numerous small effects that are hardly detectable in our design and/or that SCA is due to epistatic effects, not included in our detection models.

Also, Larièpe *et al.* (2012) and Schön *et al.* (2010) detected a large proportion of QTLs with (pseudo-)overdominance in the pericentromeric regions, consistent with the observation of McMullen *et al.* (2009) that these regions show delayed fixation when developing recombinant

inbred lines. In our design, the QTLs presenting significant effect for dominance/SCA at a 5% individual risk level were not more specifically mapped in the pericentromeric regions. A similar observation was reported by Technow *et al.* (2014) for hybrids between the two heterotic groups. Altogether these observations concur to the hypothesis that reciprocal selection of heterotic groups has fixed complementary haplotypes in low recombinant centromeric regions involving linked dominant QTL. Such regions appear with large effects in populations that recombine different groups (*e.g.* Schön *et al.* 2010; Larièpe *et al.* 2012) and not in studies that only evaluate hybrids between groups (Technow *et al.* 2014; our present study).

Improvement of the QTL detection models

Our results showed the complementarity of the different QTL models depending on the trait considered. This is consistent with Bardol *et al.* (2013) and Giraud *et al.* (2014), who also found that the model considering that each founder line carried a different allele (the "Founder alleles" model) was more adapted to complex traits such as yield than to simpler trait such as flowering time. One of the main drawbacks of this QTL detection model is that it requires the estimation of a lot of parameters (25 df for the combinations between the dent and flint populations plus six df for the GCA and nine df for the SCA per QTL). For this reason it might become difficult to apply them to more complex situations, involving a larger number of founder lines and populations. Several alternative approaches could be explored in order to improve the power of the QTL detection models as for instance considering population structure as random. Another possibility would be to adapt a method developed by Rincent *et al.* (2014) for recovering power in association mapping panels by removing from the model the population effects and instead of considering the residual genetic variance of the parents as being independent, introducing the dent and flint parental lines kinships, estimated on all the chromosomes except the one which is scanned (Rincent *et al.* 2014). In this model, the random effects would take into account the structure in populations, the fact that some hybrid combinations derived from the same parents and will account for the genetic effects not located on the scanned chromosome. Another option for more complex cases would be to clusterize the parental alleles (using for instance the "clusthaplo" package, Leroux *et al.* 2014, as done in Bardol *et al.* 2013 and Giraud *et al.* 2014). Van Eeuwijk *et al.* (2010) performed QTL detection in a factorial design issued from a private breeding program that was derived by crossing a large number of parental lines (not structured in clear families as in our design). Their analyses were based on a Bayesian model that used both molecular markers and pedigree to trace back ancestral founder alleles

and reduce model parameterization. It would be interesting to evaluate this approach on our data set even if we have limited access to the pedigree of our founder lines.

Conclusions

Even if there is room for further model improvement, our results clearly showed the interest of our design. In the joint analysis of two NAM designs (one flint and one dent) evaluated for silage test-cross performances, Giraud *et al.* (2014) detected equal or slightly higher (up to six for PH) numbers of QTL than in our study. These differences are small in regards to the fact that these NAM designs involved a much broader diversity (11 and 12 different founder lines per group instead of four) and almost two times more hybrids (about 1650 hybrids). The two designs involved different founder lines which makes it difficult to directly compare the QTLs found in the two studies. Nevertheless it is interesting to note that both studies identified the same major QTL on chromosome 10 (for DtSILK, DMC, PH and DMY) close to the *ZmCCT* gene, which was fine mapped as a major flowering time QTL by Ducrocq *et al.* (2009), and validated by Coles *et al.* (2011).

The main interest of our design, is that each hybrid is informative on both heterotic groups, which allowed us to reduce the number of tested hybrids by a factor 2 in comparison to a test-crossed evaluation based on a single tester from the opposite group. By avoiding the use of testers, our design certainly leads to better estimations of GCA QTL effects and gives the possibility to detect QTL involved in SCA (even if in our case only small SCA effects were found).

We hypothesize that with development of double haploidization methods and the decreasing costs of molecular markers, it becomes conceivable to revisit the selection for hybrids by removing the need of using testers in early stages of the breeding scheme. Marker-based predictions of the hybrid performances calibrated on factorial designs could be used to perform selection in large sets of non phenotyped candidates. Such selection would enable to take into account complementarities of favorable alleles of each group (based on GCAs and also on SCA) and accelerate the development of superior hybrids. Breeders might nevertheless be reluctant to apply this strategy which necessitates a labor intensive step of hybrid production to create the factorial design. So clearly more work is needed to further optimize this design and compare its efficiency with conventional designs based on testers. The QTLs detected in this study open the way to the implementation of a marker-assisted selection of lines in order to produce superior hybrids. Our results are encouraging but the detected QTLs only partly explain the

hybrid variability. We are currently investigating the possibility to combine predictions based on QTLs with predictions based on genomic predictions.

Acknowledgements

H. Giraud was funded by Université Paris Sud through a Scholarship from the Ministry of Research given by the Doctoral School ABIES (Agriculture Food Biology Environment Health). We thank Caussade Semences, Euralis Semences, Limagrain Europe, Maïsadour Semences, Pioneer Genetics, R2n and Syngenta Seeds grouped in the frame of the ProMais SAM-MCR program for the funding, inbred lines development, hybrid production and phenotyping. We also are grateful to scientists from these companies for helpful discussion on the results. We thank the INRA metaprogram SelGen and specifically the X-Gen and SelDir projects coordinated respectively by Andrés Legarra and Hélène Gilbert.

Literature cited

Argillier, O., V. Méchin, and Y. Barrière, 2000 Inbred line evaluation and breeding for digestibility-related traits in forage maize. Crop Sci. 40: 1596-1600.

Austin, D. F., M. Lee, L. R. Veldboom, and A. R. Hallauer, 2000 Genetic mapping in maize with hybrid progeny across testers and generations: grain yield and grain moisture. Crop Sci. 40: 30–39.

Bardol, N., M. Ventelon, B. Mangin, S. Jasson, V. Loywick *et al.*, 2013 Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. Theor. Appl. Genet. 126: 2717–2736.

Blanc, G., A. Charcosset, B. Mangin, A. Gallais, and L. Moreau, 2006 Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. Theor. Appl. Genet. 113: 206–224.

Browning, B. L., and S. R. Browning, 2009 A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84: 210–223.

Butler, D. B. R. Cullis, A. R. Gilmour, and B. J. Gogel, 2007 ASReml-R reference manual. The State of Queensland, Department of Primary Industries and Fisheries.

Coles, N. D., C. T. Zila, and J. B. Holland, 2011 Allelic effect variation at key photoperiod response quantitative trait loci in maize. Crop Sci. 51: 1036–1049.

Darwin, C., 1876 The effects of cross and self fertilization in the vegetable kingdom. John Murray, London.

Ducrocq, S., C. Giauffret, D. Madur, V. Combes, F. Dumas *et al.*, 2009 Fine mapping and haplotype structure analysis of a major flowering time quantitative trait locus on maize chromosome 10. Genetics 183: 1555–1563.

East, E. M., 1908 Inbreeding in corn, pp.419-428 in *Reports of the Connecticut Agricultural Experimental Station for years 1907-1908*. Connecticut Agricultural Experiment Station, New Haven.

Fischer, S., J. Mohring, C. C. Schön, H.-P. Piepho, D. Klein *et al.*, 2008 Trends in genetic variance components during 30 years of hybrid maize breeding at the University of Hohenheim. Plant Breeding 127: 446-451.

Foiada F., P. Westermeier, B. Kessel, M. Ouzunova, V. Wimmer *et al.*, 2015 Improving resistance to the European corn borer: a comprehensive study in elite maize using QTL mapping and genome-wide prediction. Theor. Appl. Genet. 128: 875-891.

Frascaroli, E., M. A. Canè, P. Landi, G. Pea, L. Gianfranceschi *et al.*, 2007 Classical genetic and quantitatuve trait loci analyses of heterosis in a maize hybrid between two elite inbred lines. Genetics 176: 625-644.

Frascaroli, E., M. A. Canè, M. E. Pè, G. Pea, M. Morgante *et al.*, 2009 QTL detection in maize testcross progenies as affected by related and unrelated testers. Theor. Appl. Genet. 118: 993-1004.

Ganal, M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler *et al.*, 2011 A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS ONE 6: e28334.

Gao, X., J. Starmer, and E. R. Martin, 2008 A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet. Epidemiol. 32: 361-369.

Giraud, H., C. Lehermeier, E. Bauer, M. Falque, V. Segura *et al.*, 2014 Linkage disequilibrium with linkage analysis of multiline crosses reveals different multiallelic QTL for hybrid performance in the flint and dent heterotic groups of maize. Genetics 198: 1717-1734.

Grieder, C., B. S. Dhillon, W. Schipprack, and A. E. Melchinger, 2012 Breeding maize as biogas substrate in Central Europe: I. Quantitative-genetic parameters for testcross performance. Theor. Appl. Genet. 124: 971-980.

Hickey, J. M., G. Gorjanc, R. K. Varshney, and C. Nettelblad, 2015 Imputation of single nucleotide polymorphism genotypes in biparental, backcross, and topcross populations with a hidden Markov model. Crop Sci. 55: 1934-1946.

Kump, K. L., P. J. Bradbury, R. J. Wisser, E. S. Buckler, A. R. Belcher *et al.*, 2011 Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nat. Genet. 43: 163–168.

Larièpe, A., B. Mangin, S. Jasson, V. Combes, F. Dumas *et al.*, 2012 The genetic basis of heterosis: multiparental quantitative trait loci mapping reveals contrasted levels of apparent overdominance among traits of agronomical interest in maize (*Zea mays* L.). Genetics 190: 795-811.

Leroux, D., A. Rahmani, S. Jasson, M. Ventelon, F. Louis *et al.*, 2014 Clusthaplo: a plug-in for MCQTL to enhance QTL detection using ancestral alleles in multi-cross design. Theor. Appl. Genet. 127: 921-933.

Lu, H., J. Romero-Severson, and R. Bernardo, 2003 Genetic basis of heterosis explored by simple sequence repeat markers in a random-mated maize population. Theor. Appl. Genet. 107: 494-502.

Lübberstedt, T., A. E. Melchinger, D. Klein, H. Degenhardt, and C. Paul, 1997 QTL mapping in test crosses of European flint lines of maize: II. Comparisons of different testers for forage quality traits. Crop Sci. 37: 1913-1922.

Manicacci, D., L. Moreau, and A. Charcosset, 2011 Quantitative trait loci cartography, meta-analysis and association genetics. Chapter 2 in *Advances in Maize, Essential Reviews in Experimental Biology*, Volume 3. Edition Society for Experimental Biology.

McMullen, M. D., S. Kresovich, H. S. Villeda, P. Bradbury, H. Li *et al.*, 2009 Genetic properties of the maize nested association mapping population. Science 325: 737–740.

Melchinger, A. E., H. F. Utz, and C. C. Schön, 1998 Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL. Genetics 149: 383-403.

Nakagawa S., and H. Schielzeth, 2013 A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4: 133-142.

Parisseaux, B., and R. Bernardo, 2004 In silico mapping of quantitative trait loci in maize. Theor. Appl. Genet. 109: 508-514.

R Core Team, 2013 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rebaï, A., P. Blanchard, D. Perret, and P. Vincourt, 1997 Mapping quantitative trait loci controlling silking date in a diallel cross among four lines of maize. Theor. Appl. Genet. 95: 451-459.

Reif, J. C., F.-M. Gumpert, S. Fischer, and A. E. Melchinger, 2007 Impact of interpopulation divergence on additive and dominance variance in hybrid populations. Genetics 176: 1931-1934.

Rincent, R., L. Moreau, H. Monod, E. Kuhn, A. E. Melchinger *et al.*, 2014 Recovering power in association mapping panels with variable levels of linkage disequilibrium. Genetics 197: 375-387.

Schön, C. C., A. E. Melchinger, J. Boppenmaier, E. Brunklaus-Jung, R. G. Herrmann *et al.*, 1994 RFLP mapping in maize: quantitative trait loci affecting the testcross performance of elite European flint lines. Crop Sci. 34: 378-389.

Schön, C. C., B. S. Dhillon, H. F. Utz, and A. E. Melchinger, 2010 High congruency of QTL positions for heterosis of grain yield in three crosses of maize. Theor. Appl. Genet. 120: 321–332.

Schrag, T. A., A. E. Melchinger, A. P. Sørensen, and M. Frisch, 2006 Prediction of single-cross hybrid performance for grain yield and grain dry matter content in maize using AFLP markers associated with QTL. Theor. Appl. Genet. 113:1037-1047.

Schrag, T. A., J. Möhring, H. P. Maurer, B. S. Dilhon, A. E. Melchinger et al., 2009 Molecular marker-based prediction of hybrid performance in maize using unbalanced data from multiple experiments with factorial crosses. Theor. Appl. Genet. 118: 741-751.

Schrag, T. A., J. Möhring, A. E. Melchinger, B. Kusterer, B. S. Dhillon *et al.*, 2010 Prediction

of hybrid performance in maize using molecular markers and joint analyses of hybrids and parental inbreds. Theor. Appl. Genet. 120: 451-461.

Shull, G. H., 1908 The composition of a field of maize. Am. Breed. Assoc. Rep. 5: 51-59.

Shull, G. H., 1914 Duplicate genes for capsule-form in *Bursa bursa pastoris*. Zeitschrift ind. Abst. U. Verebsgl. 12: 97-149.

Sprague, G. F., and L. A. Tatum, 1942 General vs. specific combining ability in single crosses of corn. J. Am. Soc. Agron. 34: 923-932.

Technow, F., T. A. Schrag, W. Schipprack, E. Bauer, H. Simianer *et al.*, 2014 Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. Genetics 197: 1343-1355.

Tenaillon, M. I., and A. Charcosset, 2011 A European perspective on maize history. C. R. Biol. 334: 221–228.

Truntzler, M., Y. Barrière, M. C. Sawkins, D. Lespinasse, J. Bertran *et al.*, 2010 Meta-analysis of QTL involved in silage quality of maize and comparison with the position of candidate genes. Theor. Appl. Genet. 121: 1465-1482.

van Eeuwijk, F. A., M. Boer, L. R. Totir, M. Bink, D. Wright *et al.*, 2010 Mixed model approaches for the identification of QTLs within a maize hybrid breeding program. Theor. Appl. Genet. 120: 429-440.

Chapter 3

Can genomic selection deeply modify early stages of hybrid breeding programs? A case study on maize (*Zea mays* L.) silage performance of single-crosses between two multiparental populations

Héloïse Giraud[*], Cyril Bauland[*], Delphine Madur[*], Valérie Combes[*], Philippe Jamin[*], Cécile Monteil[*], Carine Palaffre[†], Antoine Gaillard[‡], Philippe Blanchard[§], Alain Charcosset[*], Laurence Moreau[*1]

[*] GQE-Le Moulon, INRA, Univ Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, F-91190, Gif-sur-Yvette, France

[†] UE0394 SMH Maïs, INRA, 2297 route INRA, F-40390 Saint-Martin-de-Hinx, France

[‡] Maïsadour Semences SA, Route de Saint Sever, BP27, F-440001 Mont de Marsan Cedex, France

[§] Euralis Semences, Domaine de Sandreau, 6 chemin de Panedautes, Mondonville, F-31700, France

[1]Corresponding author

Running title: Genomic prediction of hybrid value

Key words: Genomic prediction; General Combining Ability (GCA); Specific Combining Ability (SCA); hybrid; maize

Corresponding author:
Dr. Laurence Moreau
GQE-Le Moulon, INRA, Univ Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay
F-91190
Gif-sur-Yvette
France
Phone: +33 1 69 33 23 37
Fax: +33 1 69 33 23 40
Email: moreau@moulon.inra.fr

INTRODUCTION

Maize, as many allogamous species, presents a strong heterosis effect for many traits related to biomass production (Shull, 1914). To benefit from heterosis, maize diversity is partitioned into heterotic groups and most of maize varieties cultivated nowadays consist in F_1 hybrids between two inbred parental lines issued from different heterotic groups.

Since the beginning of the XX^{th} century, numerous hybrid maize breeding procedures have been proposed in order to optimize its selection. Sprague and Tatum proposed in 1942 the decomposition of the hybrid genetic value into General Combining Abilities (GCA) and Specific Combining Ability (SCA). The GCA of one parental line corresponds to the average performance of its descendants in hybrid combinations. The SCA corresponds to the difference between the hybrid observed value and its predicted value based on the GCAs of its parents. In 1949, Comstock *et al.* proposed the recurrent reciprocal selection which consists in improving simultaneously the two parental populations of the hybrids, by evaluating plants of one population for the value of their progeny when crossed to several plants of the other population. This approach allows selecting for both GCA and SCA of the hybrids and shows its superiority over other reciprocal selection methods (Coors 1999). Variations of this selection scheme were proposed later (see Hallauer *et al.* 2010, Chapter 12). In breeding companies, a modified recurrent reciprocal selection scheme is generally used due to practical considerations. Breeding procedure is generally divided into two stages. In the first stage, candidate lines of one heterotic group are crossed with one or a few “testers” (often inbred lines) from the complementary heterotic group. Hybrid progeny is then evaluated in a field network and only parents of the best hybrids are selected for the second stage. Tester choice is very important in this process. The testers often correspond to inbred lines of the reciprocal heterotic group presenting good agronomical qualities and which are likely to be used as one parent of the final hybrid. Thus, during early stages, not all potential hybrid combinations between the two heterotic groups are evaluated. It is only in the last stages that a larger set of hybrid combinations between the few inbred lines selected in each heterotic group is evaluated. Selection based on a few testers facilitates hybrid seed production for early testing trials but limits the variability considered at the second stage.

Genotyping developments in maize and other crops of agronomical interest should soon make it possible to genotype all potential candidates at a reasonable cost. In this context, genotyping based prediction of hybrid value is of major interest for limiting phenotyping effort to the most

promising hybrid combinations. Genetic factors involved in heterosis and hybrid value are still poorly known, but it is admitted that underlying mechanisms (dominance or epistasis) involved interactions between different alleles transmitted by the parents of the single-cross. The first attempts of hybrid value prediction consisted in using as predictor the marker-based genetic distance between the two parents of the single-cross (Lee *et al.* 1989). Theoretical studies were carried out (Charcosset *et al.* 1991 and Bernardo 1992) as well as numerous empirical studies, mostly based on Restricted Fragment Length Polymorphisms (RFLPs) data (Godhsalk *et al.* 1990; Melchinger *et al.* 1990; Dudley *et al.* 1991; Burstin *et al.* 1995). Results showed that genetic distance between parents failed at predicting inter-group hybrid value. One explanation is that marker-based distance between unrelated inbred lines only poorly reflects allelic differences at QTLs involved in hybrid value, due to weak and unstable marker-QTL associations over groups (Melchinger *et al.* 1992; Charcosset and Essioux 1994). Bernardo proposed in 1994 to use Best Linear Unbiased Predictor (BLUP) of the performance of single-crosses using performances of a related set of single crosses and genomic similarities between tested and untested single-crosses. This approach was applied to a factorial design between six lines from the Iowa Stiff Stalk Synthetic group and nine inbreds from another group. It is to our knowledge the first application of genomic prediction model that was later referred to as GBLUP model (Meuwissen *et al.* 2001). More recently support vector machine regression (Maenhout *et al.* 2007; Maenhout *et al.* 2010) or Bayesian approaches (Technow *et al.* 2014) were proposed as an alternative to the GBLUP model. In addition to studies on maize, exploratory studies on genomic prediction of hybrid value have been performed recently in other plant species such as wheat (Zhao *et al.* 2013), sunflower (Reif *et al.* 2013), triticale (Gowda *et al.* 2013), rye (Wang *et al.* 2014) and rice (Xu *et al.* 2014).

In maize, first genomic hybrid predictions were carried out in factorial designs between few inbred lines (Bernardo 1994) or by taking advantage of hybrids from advanced selection stages of breeding programs (Bernardo 1996a; Bernardo 1996b; Maenhout *et al.* 2010; Massman *et al.* 2013a; Technow *et al.* 2014). Most of these studies showed that the quality of prediction of a given hybrid mostly depended on the inclusion in the calibration set of hybrids issued from the same parental lines. Until recently, material screened through tester evaluation were unfixed individuals (F_2 or F_3 plants). In parallel to test-cross evaluation, selected plants were self-pollinated to produce inbred lines. With double haploid technology, it is now possible to obtain directly large segregating populations of inbred lines. This reduces the length of breeding cycles and offers new opportunities to revisit hybrid breeding schemes. Instead of using a few testers from one group for selecting inbred lines of the other one, before exploring larger set of hybrid

combinations, one can directly generate large sets of inter-group single-cross hybrids from unselected inbred lines. Then these data can be used to calibrate hybrid value prediction equation and such predictions can be used to identify the most promising hybrid combinations. The same model can also be used to predict the GCA of all candidate lines and select the most promising for creating breeding populations for the round of selection.

The objectives of this study were to evaluate the feasibility of this approach. We developed a design between the two main heterotic groups used for silage maize breeding in Northern Europe: the dent and the flint groups. In each heterotic group, six biparental populations of inbred lines were developed from four founder lines. These inbred lines were crossed according to an incomplete factorial design in order to obtain hybrids which were phenotypically evaluated for silage performances. QTL detection was performed in this design (cf. Chapter 2). We first evaluated the efficiency of genomic predictions in this design using GBLUP models considering different allele codings, population structure or not and decomposing hybrid value in GCAs or GCAs and SCA. We investigated the accuracy of our predictions on the different components of hybrid value (GCA / SCA). We then compared genomic prediction to QTL-based predictions. We also investigated the influence of the size and of the composition of the calibration set on the quality of hybrid value predictions to draw some conclusions for further optimization of this approach.

MATERIAL AND METHODS

Plant material

Two multiparental connected designs, each corresponding to one of the major heterotic groups used for silage maize breeding in Europe, the flint or the dent, were crossed in order to obtain F_1 hybrids. Each of these multiparental designs is constituted of six populations of parental lines derived by haplodiploidization (for the dent) or by five to six generations of selfing (for the flint) from a half diallel mating design between four founder lines. The 863 dent parental lines and the 879 flint parental lines were crossed according to an incomplete factorial design in order to obtain 1044 dent – flint hybrids which were phenotypically evaluated. The majority of lines (699 in the dent and 732 in the flint) contributed to a single hybrid. Some lines contributed to two hybrids (163 in the dent group and 146 in the flint group) and one dent parental line contributed to three hybrids and one flint to four hybrids. The experimental design is described in details in Chapter 2.

Genotyping data

The founder lines were genotyped with a 50 K SNP array (Ganal *et al.* 2011). The parental lines and the founder lines were genotyped with a 18 480 SNP Affymetrix® array provided by Limagrain. Inbred lines with unexpected high percentage of heterozygote loci as well as those presenting inconsistencies between founder lines and segregating lines were discarded. After quality control, 875 dent and 883 flint lines were considered in the analysis (see Chapter 2 for more details). We used the 9548 markers which were mapped at the parental lines level in the Chapter 2. For the analyses we considered the Affymetrix® genotyping data for the founder lines, and when needed replace missing data by the genotypes obtained with the 50 K chip when it was available.

Missing genotypes at the parental level were imputed with Beagle v3.0. (Browning and Browning 2007) by populations, on genetic maps, putting the founder lines in the dataset. Phasing of the flint lines and of the founder lines that presented residual heterozygosity was done at the same time than missing genotypes imputation.

Phenotypic evaluation and adjusted means

The hybrids between the parental lines were evaluated in 8 environments in Northern France and Germany over two years for four traits: silage yield (DMY in tons of dry matter per ha), dry matter content at harvest (DMC in %), plant height (6 environments) (PH in cm) and female flowering (DtSILK, days after January the first). The field experiments were conducted as augmented p-rep designs with 1088 two-row plots. They were laid out in 68 incomplete blocks consisting of 16 plots each with five to six plots used for repeated genotypes (hybrids between the founder lines plus around 17 % of the hybrids between the parental lines). Only 951 hybrids for which both parental lines had correct genotypic data were analyzed (950 for plant height and dry matter yield). All the dent and flint populations of parental lines contributed to the 951 hybrids that were considered in the analyses. As far as possible we tried to balance the contribution of each parental population to the hybrids but some populations contributed more than others (Figure 1). The number of analyzed hybrids derived from each flint population varied between 130 (for population F3) and 178 (for population F1) and the number of analyzed hybrids derived from each dent population varied between 126 (for population D6) and 168 (for population D3). For each dent-flint combination of populations between 15 (dent population D6 - flint population F3) and 34 hybrids (dent population D2 - flint population F1) were derived. Outlying observations were deleted. One environment over eight was excluded for dry matter yield as it was not correlated with the others. For each combination environment – trait, field

heterogeneities were corrected using a block or a row-column effect model as described in Chapter 2. Using the Asreml-R package (Butler *et al.* 2007; R Core Team 2013), individual single-plot performances were corrected by the BLUPs of the field effects and were used to compute for each hybrid adjusted means (ls-means) over the eight environments.

Genomic Best Linear Unbiased Prediction (GBLUP) models and estimation of variance components

Genomic predictions were done using as phenotypes the hybrid ls-means. We compared several GBLUP models that differed by the way population structure was considered, and the inclusion or not of SCA. All these models were implemented in Asreml-R (Butler *et al.* 2007; R Core Team 2013).

The following GBLUP model considered the structure in populations and the SCA:

$$y = 1\,.\,\mu + A\,.\,\alpha + B\,.\,\beta + C\,.\,(\alpha\beta) + Z_{GCA_d}\,.\,g_{GCA_d} + Z_{GCA_f}\,.\,g_{GCA_f} + Z_{SCA}\,.\,s_{SCA} + e$$

Where y is the $(N \times 1)$ vector of the adjusted means of the N phenotyped hybrids for the studied trait, μ is the intercept, 1 is a $(N \times 1)$ vector of 1. α (respectively β) is a (6×1) vector of the fixed effects of the dent (respectively flint) populations of origin of the dent (flint) parental lines, A (respectively B) is the $(N \times 6)$ design matrix of 0-1 that linked each hybrid to the dent (respectively flint) population of its dent (flint) parental line, $(\alpha\beta)$ is a (36×1) vector of the fixed interaction effects between the dent and flint populations of origin of the parental lines, C is the corresponding design matrix. g_{GCA_d} (respectively g_{GCA_f}) is the $(n_d \times 1)$ vector (respectively $(n_f \times 1)$) of the random effects corresponding to the n_d dent GCA (respectively n_f flint) of the parental dent (respectively flint) lines, with $g_{GCA_d} \hookrightarrow N\left(0, G_{GCA_d}\,\sigma^2_{GCA_d}\right)$ and $g_{GCA_f} \hookrightarrow N\left(0, G_{GCA_f}\,\sigma^2_{GCA_f}\right)$ where G_{GCA_d} (G_{GCA_f}) is the $(n_d \times n_d)$ (respectively $(n_f \times n_f)$) genomic relationship matrix between the dent (flint) lines, and $\sigma^2_{GCA_f}$ $(\sigma^2_{GCA_d})$ is the flint (dent) GCA variance component. s_{SCA} is the $(N \times 1)$ vector of the random effects corresponding to the SCA with $s_{SCA} \hookrightarrow N(0, S_{SCA}\sigma^2_{SCA})$ where S_{SCA} is a $(N \times N)$ matrix equals to the Hadamard product between the dent and the flint genomic relationship matrices $(S_{SCA} = G_{GCAf} \times G_{GCAd})$ and σ^2_{SCA} is the variance component associated with SCA effects, Z_{GCA_d}, Z_{GCA_f}, Z_{SCA} are the $(N \times n_d)$, $(N \times n_f)$, and $(N \times N)$ incidence matrices that relate y to g_{GCA_d}, g_{GCA_f} and s_{SCA}, e is the $(N \times 1)$ vector of the residual effects associated to the model with $e \hookrightarrow N(0, \sigma^2_E)$.

The genomic relationship coefficient of the G_{GCA_d} matrix between the dent parental lines i and j based on the observed SNPs was computed according to method 1 of Van Raden (2008) as

$G_{GCAd,i,j} = \frac{1}{L} \frac{\sum_{l=1}^{L}(x_{il} - p_l)(x_{jl} - p_l)}{\sum_{l=1}^{L} p_l(1 - p_l)}$, where p_l is the frequency calculated over the dent lines (respectively flint lines) of the allele coded 1 of marker l, x_{il} (respectively x_{jl}) is the genotype of the dent inbred line i (respectively j) at the marker l coded in 0, 0.5, 1, and L is the total number of markers. Genomic relationship matrix between the flint parental lines was calculated in a similar way. These kinship matrices are further called SNP kinship matrices.

Different variants of the GBLUP model presented above were considered: without the structure in populations, without the SCA component, without both. For each GBLUP model, we estimated the different variance components, in order to estimate the flint GCA, the dent GCA and the SCA variances captured by the markers as well as the percentage of genetic variance explained by SCA. For GBLUP models presenting the same fixed effects, Restricted Maximum Likelihood (REML) likelihood ratio tests were performed for evaluating the impact of incorporation of SCA considering a risk level of 1%. The variance components captured by the markers were compared with those estimated in Chapter 2 based on the analysis of the raw phenotypic data, without considering genotypic kinship matrices (cf. Chapter 2, Table 1). For each GBLUP model, we evaluated its adjustment to the ls-means (R^2) by the square of the correlation between predictions and ls-means.

GBLUP model using kinship matrices based on founder alleles

A second estimation of the genomic relationship matrices was used based on the founder alleles inherited by the parental lines. To do so, we considered the probability that a given dent (flint) founder line allele was transmitted to each hybrid at positions corresponding to the different markers. These probabilities for the 9548 mapped markers were inferred using PlantImpute (Hickey *et al.* 2015) using 10 iterations. Relationship matrix between the dent (flint) parental lines was calculated based on these probabilities as the expectation of the percentage of the marker positions at which two lines received the same founder allele. The genomic relationship coefficient of the F_{GCA_d} matrix between the dent parental lines i and j based on the founder alleles probabilities was computed as $F_{GCAd,i,j} = \frac{\sum_{l=1}^{L} \sum_{f=1}^{4} q_{ilf} \times q_{jlf}}{L}$ where q_{ilf} (respectively q_{jlf}) is the probability that the inbred line i (respectively j) received the founder allele f at the marker position l, and L is the total number of markers. Flint founder allele relationship matrix was computed in a similar way. Compared with the kinships estimated based on SNPs, kinships based on founder alleles considered as unrelated hybrids that did not share any of the founder lines. With this definition of alleles, hybrid value prediction only relied on the performances of

hybrids related by pedigree. These IBD coefficients are expected to capture differences between founder line alleles not captured by SNP polymorphisms (such as recent mutations, etc.).

Quality of prediction of the different GBLUP models. Comparison with prediction based on QTL

For comparing quality of prediction of the different GBLUP models, we performed a cross-validation approach using one fifth of the data (about 190 hybrids) for validation and four fifth for calibration. Sampling of the validation set was stratified by the structure in populations and was repeated 100 times. The calibration set consisted in the others individuals. The same samplings were used for all the GBLUP models to facilitate their comparison. For each replicate and each model we estimated the predictive ability of the model by the correlation between the predictions and the ls-means of the validation set. Results were then averaged over the 100 replications. As the same samplings were used for all models, we also evaluated the number of replications for which a given model led to a better predictive ability than another one.

Predictions based on GBLUP models were also compared with predictions based on the QTL detection results obtained in the Chapter 2, on the same experimental design. In this study, different QTL detection models were used (using different types of allele coding and considering equal or different effects in the two genetic groups). For the comparison with the GBLUP models, we considered for each trait the QTL detection model that gave the best predictive ability. In each replication, we considered a model that included the population structure and the QTL positions detected in the Chapter 2 as fixed effect. All fixed effects were estimated using the calibration set, then used to predict the value of each individual of the validation set. To make things comparable, the same samplings were used for the GBLUP and the QTL models. As for the GBLUP models, we then computed the predictive ability of the predictions by the correlation between predictions and ls-means of the validation set.

For the best GBLUP model considering SCA identified by cross-validation, we evaluated the quality of prediction of the flint and dent GCA components of the hybrid value. To do so, for each replication, we estimated the correlation between the dent (flint) GCA predicted using the GBLUP model and the observed phenotypes $r(GCA_{d_{mk}}, P)$. As this correlation depends on the percentage of the phenotypic variance explained by the GCA component, we divided it by the square root of the proportion of the corresponding GCA variance in the performances ($\sqrt{var(GCA_d)/var(P)}$) where $var(GCA_d)$ is the GCA variance estimated in the Chapter 2 (see Chapter 2, Table 1) and $var(P)$ is the variance of the ls-means of the studied trait. This

value $\rho_{GCA_d} = \frac{r(GCA_{d_{mk'}}, P)}{\sqrt{var(GCA_d)/var(P)}}$ corresponds to the observed accuracy of the dent GCA predictions. Similar estimation was done for the observed accuracy of flint GCA predictions, ρ_{GCA_f}.

Influence of the calibration set size on the hybrid predictions

For the best GBLUP model based on the result of previous cross validations, we evaluated the impact of the number of hybrids evaluated for each of the 36 dent-flint population combinations, considering or not SCA in the model. To do so, for constituting the calibration set, we sampled two, three, five, six, 10 and 12 hybrids in each of the 36 dent-flint hybrid populations: corresponding respectively to sampling in total of 72, 108, 180, 216, 360 and 432 hybrids. For each calibration set size, sampling was repeated 100 times and the remaining part of the population was used as validation set. For each replicate, the predictive ability was computed. Then, it was averaged over the 100 replications.

Influence of the calibration set composition on the hybrid predictions

Impact of including half-sib hybrids in calibration set

We also investigated the differences in quality of prediction of the hybrids of the validation set, depending if zero, one or both of their parents were also parents of hybrids included in the calibration set (this corresponds to the T0, T1 and T2 hybrid categories defined in Schrag *et al.* 2010 and Technow *et al.* 2012). We used a cross-validation approach using four fifth of the data for calibration and one fifth for validation. In the validation set, we included 30 hybrids that had at least one half-sib hybrid on the dent side and one on the flint side in the calibration set (T2 hybrids), 30 hybrids that had at least one half-sib hybrid on the dent side and no half-sib hybrid on the flint side in the calibration set (T1 Dent hybrids) and 30 hybrids that had at least one half-sib hybrid on the flint side and no half-sib on the dent side in the calibration set (T1 Flint hybrids). 100 T0 hybrids (hybrids with no half-sib in the calibration set) were used to complete the validation set. Sampling was repeated 100 times. For each replicate, the predictive ability of the hybrids of each category was computed and then averaged over the 100 replications.

Impact of including hybrids issued from the same founder lines in calibration set

Further, we evaluated the quality of prediction of a given dent-flint population using individuals from other populations. We did that in order to evaluate what would be the efficiency of

genomic selection for a new population of hybrids issued from inbred lines that did not contribute yet to the hybrids of the calibration set. For each dent-flint population to predict (in red on Figure 1), five types of calibration sets were tested (in blue, green, yellow, orange and purple on Figure 1) depending on their level of relatedness with the population to predict. In the first case, the populations of the calibration set had their two dent and one of their flint founders in common with the population to predict (four dent-flint populations in the calibration set, in blue on Figure 1). In this case, the target hybrid population was predicted using hybrids between (i) dent lines that were full-sibs (two dent founders in common) of their dent parents and (ii) flint lines that were half-sibs (one flint founder in common) of their flint parents. In the second case, the populations of the calibration set had their two flint and one of their dent founders in common with the population to predict (four dent-flint populations in the calibration set, in green on Figure 1). In the third case, they had only one dent and one flint founders in common with the population to predict (16 dent-flint populations in the calibration set, in yellow on Figure 1). In the fourth case, populations of the calibration set had only one of their dent founders in common with the population to predict and no flint founder (four dent-flint populations in the calibration set, in orange on Figure 1). In the last case, they had only one of their flint founders in common and no dent founder (four dent-flint populations, in purple on Figure 1). In all cases except the third one, the size of the populations to be included in the calibration set was too small to perform sampling. In the third case, sampling stratified by the structure in populations of a fourth of the hybrids of the calibration populations was done, in order to have a number of individuals equivalent to the ones of the others cases. Sampling was repeated 50 times and the predictive ability was averaged over the 50 replications for each population. For the five different cases, predictive abilities of the different dent-flint populations were averaged. It can be noted that we did not consider the case when calibration is done using populations which only have two dent founder lines or two flint founder lines in common with the target population or the case when calibration is done based on populations with no founder in common with the target population. In our design, considering these cases would have been equivalent to using only one of the 36 hybrid populations to predict another one which would not be feasible considering the size of each hybrid population (about 30 hybrids).

	F1 F373 x F02803	F2 F373 x F03802	F3 F02803 x F03802	F4 F373 x F7088	F5 F02803 x F7088	F6 F03802 x F7088
D1 F1808 x F04401	D1 x F1 To predict 31 hybrids	2D – 1F 27 hybrids	2D – 1F 22 hybrids	2D – 1F 30 hybrids	2D – 1F 25 hybrids	27 hybrids
D2 F1808 x F98902	1D – 2F 34 hybrids	1D – 1F 27 hybrids	1D – 1F 22 hybrids	1D – 1F 30 hybrids	1D – 1F 23 hybrids	1D – 0F 32 hybrids
D3 F04401 x F98902	1D – 2F 28 hybrids	1D – 1F 29 hybrids	1D – 1F 23 hybrids	1D – 1F 30 hybrids	1D – 1F 26 hybrids	1D – 0F 27 hybrids
D4 F1808 x F7082	1D – 2F 28 hybrids	1D – 1F 28 hybrids	1D – 1F 25 hybrids	1D – 1F 30 hybrids	1D – 1F 27 hybrids	1D – 0F 27 hybrids
D5 F04401 x F7082	1D – 2F 33 hybrids	1D – 1F 29 hybrids	1D – 1F 23 hybrids	1D – 1F 28 hybrids	1D – 1F 24 hybrids	1D – 0F 30 hybrids
D6 F98902 x F7082	24 hybrids	0D -1F 23 hybrids	0D – 1F 15 hybrids	0D – 1F 25 hybrids	0D – 1F 21 hybrids	18 hybrids

Figure 1: Repartition of the 951 hybrids in function of the population of origin of their dent and flint parent. Each inbred line population is denoted D1 to D6 for the dent ones and F1 to F6 for the flint ones and we indicated the names of the two founder lines of the population. This figure illustrates the different calibration sets categories considered for the prediction of a given dent-flint hybrid population (here the D1 x F1 population in red) using the other dent-flint populations as calibration set, in function of the number of founder lines in common with the target population.

RESULTS

Estimation of variance components and R^2 of the GBLUP models

For the different traits and models, a large part of the genetic variation was attributed to the GCAs components (Table 1). Variance estimations for the GCAs did not varied much from one model to the other. Including or not population structure in the model had a low impact on variance component estimation. Dent and flint GCA variances were of the same magnitude for

all traits except DtSILK for which flint GCA variance was equal to twice the dent GCA variance. Taking into account SCA in the model did not significantly improved the likelihoods of the model except for DMY. The SCA variance component was small compared to the GCAs variances. It explained between 0.00 (DMC) and 8.17% (DMY) of the total genetic variance when population structure was not included in the model and between 0.00 (DMC) and 11.24% (DMY) when population structure was included in the model. Based on the standard deviations, considering or not population structure, the SCA variance was significantly different from 0 only for DMY.

The proportion of SCA in the hybrid variance captured by the markers was much less important than the one obtained when using field plot data (Supplementary Table S1). It was more striking for DMC for which there was no SCA when considering genotypic information whereas it represented 11.7% of the genetic variance (21.5% of the within-population genetic variance) when performing variance component analysis using only filed plot performances. For the other traits, the percentage of SCA in the genetic variance was 2.1 (DMY) to 7.3 (DtSILK) times less important when considering genotypic information.

Adjustment of the model to the data (R^2) was high for all the trait x model combinations: between 0.717 (DMY, without SCA, with structure in populations) and 0.905 (PH with SCA, without the structure in populations). For all traits, considering structure in populations had a negative effect on the R^2. R^2 was higher when considering SCA than when not considering it except for DMC for which no difference was observed (Table 1).

Table 1: Variance components estimated using four different GBLUP models (all considering the SNP kinship matrices) and adjustment of the models to the data (measured by R^2) for the studied traits (Dry Matter Content DMC, Dry Matter Yield DMY, Silking date DtSILK, Plant Height PH). $\sigma^2_{GCA_d}$ corresponds to the dent General Combining Ability variance, $\sigma^2_{GCA_f}$ to the flint General Combining Ability variance, σ^2_{SCA} to the Specific Combining Ability variance, and σ^2_E to the residual variance. Confidence intervals of the variance components are indicated. For each trait, heritability is specified.

	Pop + GCA + SCA	Pop + GCA	GCA + SCA	GCA
DMC	*Heritability: 0.892*			
$\sigma^2_{GCA_d}$	1.368 ± 0.480	1.368 ± 0.480	1.413 ± 0.489	1.413 ± 0.489
$\sigma^2_{GCA_f}$	1.788 ± 0.543	1.788 ± 0.543	1.787 ± 0.542	1.787 ± 0.542
σ^2_{SCA}	0.000 ± 0.000	-	0.000 ± 0.000	-
σ^2_E	1.084 ± 0.154	1.084 ± 0.157	1.077 ± 0.156	1.077 ± 0.156
%SCA	0.00	-	0.00	-
R^2	0.833	0.833	0.848	0.848
DMY	*Heritability: 0.814*			
$\sigma^2_{GCA_d}$	0.618 ± 0.236	0.635 ± 0.244	0.622 ± 0.236	0.636 ± 0.243
$\sigma^2_{GCA_f}$	0.481 ± 0.188	0.442 ± 0.178	0.477 ± 0.186	0.442 ± 0.177
σ^2_{SCA}	0.139 ± 0.125	-	0.132 ± 0.121	-
σ^2_E	0.511 ± 0.124	0.643 ± 0.084	0.517 ± 0.122	0.643 ± 0.084
%SCA	11.24	-	8.17	-
R^2	0.829	0.717	0.834	0.726
DtSILK	*Heritability: 0.890*			
$\sigma^2_{GCA_d}$	1.180 ± 0.414	1.174 ± 0.413	1.208 ± 0.420	1.207 ± 0.420
$\sigma^2_{GCA_f}$	2.029 ± 0.570	2.007 ± 0.568	2.021 ± 0.567	1.998 ± 0.566
σ^2_{SCA}	0.089 ± 0.133	-	0.088 ± 0.130	-
σ^2_E	0.735 ± 0.167	0.823 ± 0.131	0.734 ± 0.166	0.821 ± 0.130
%SCA	2.70	-	2.23	-
R^2	0.862	0.828	0.882	0.848
PH	*Heritability: 0.877*			
$\sigma^2_{GCA_d}$	70.396 ± 21.813	69.395 ± 22.747	70.635 ± 22.962	70.013 ± 22.878
$\sigma^2_{GCA_f}$	68.642 ± 22.105	66.705 ± 21.814	68.749 ± 22.100	67.157 ± 21.893
σ^2_{SCA}	6.868 ± 8.904	-	6.576 ± 8.496	-
σ^2_E	32.923 ± 9.306	39.434 ± 6.169	33.196 ± 9.079	39.381 ± 6.170
%SCA	4.71	-	3.81	-
R^2	0.884	0.841	0.905	0.862

Predictive abilities of the GBLUP models. Comparison with predictions based on QTL

Predictive abilities of the validation set were high for all the models and traits and only slightly varied according to the trait. When considering the best model for each trait, they varied between 0.652 for DMY (SNP kinship – no population structure – SCA) and 0.771 for PH (SNP kinship – no population structure – SCA) (Table 2).

For all traits, with the SNP kinship matrices, considering population structure had a negative effect on the prediction accuracy (Table 2). Without SCA in the model, not considering structure in populations gave better predictive abilities for 74% (DtSILK) to 98% (PH) of the replications. When considering SCA in the model, between 60% (DMY and DtSILK) and 63%

(DMC and PH) of the replications gave better predictive abilities without population structure than with population structure. Without SCA in the model, not considering population structure was better for 54% (DMY) to 67% (DtSILK) of the replications. With SCA in the model, not considering population structure was better for 54% (DMY) to 68% (DtSILK) of the replications.

Adding SCA in the GBLUP model with the SNP kinship IBD matrices had a really low effect, either positive or negative, on the predictive ability (Table 2). Only, between 44% (PH) and 52% (DMY, DtSILK) of the replications gave better predictive abilities when considering SCA than when not considering it in the GBLUP model without population structure included. Similar results were observed when considering the founder allele kinship matrices.

For all trait x model combinations, considering the founder allele kinship matrices decreased predictive abilities compared to using the SNP kinship matrices (Table 2). As for the GBLUP model using SNP kinship, we observed that incorporating the population structure in the model had a negative effect and that adding SCA had a small impact on the predictive abilities.

For all traits, best predictive abilities obtained from GBLUP were higher than the ones obtained based on detected QTLs, which varied between 0.626 for DMY and 0.742 for PH (Table 2). For genomic predictions and QTL based predictions, DMC and PH showed close predictive abilities and were the best predicted traits. DMY was the worst predicted one. The best GBLUP model, gave better predictive abilities than the QTL-based prediction model for around 75% of the replications (between 71% for DtSILK and 76% for PH).

Table 2: Predictive ability for the different GBLUP models using a cross validation procedure with four fifth of the data for calibration and one fifth for the validation for the studied traits (Dry Matter Content DMC, Dry Matter Yield DMY, Silking date DtSILK, Plant Height PH). Average values over 100 replications and standard deviations (sd) are indicated. Predictive abilities of predictions based on QTL (Pop+ QTL) obtained in Chapter 2 are also provided.

Model	DMC	DMY	DtSILK	PH
GBLUP SNP kinship				
Pop + GCA + SCA	0.757 sd 0.024	0.636 sd 0.036	0.687 sd 0.031	0.755 sd 0.032
Pop + GCA	0.757 sd 0.025	0.633 sd 0.036	0.689 sd 0.031	0.754 sd 0.032
GCA + SCA	0.768 sd 0.024	0.652 sd 0.042	0.697 sd 0.035	0.771 sd 0.026
GCA	0.768 sd 0.024	0.644 sd 0.035	0.700 sd 0.029	0.769 sd 0.031
GBLUP Founder allele kinship				
Pop + GCA + SCA	0.755 sd 0.029	0.623 sd 0.037	0.674 sd 0.034	0.751 sd 0.028
Pop + GCA	0.756 sd 0.029	0.623 sd 0.037	0.676 sd 0.034	0.751 sd 0.028
GCA + SCA	0.763 sd 0.025	0.631 sd 0.037	0.688 sd 0.031	0.762 sd 0.030
GCA	0.763 sd 0.025	0.628 sd 0.036	0.688 sd 0.032	0.762 sd 0.030
Pop + QTL				
	0.743 sd 0.027	0.626 sd 0.039	0.689 sd 0.034	0.742 sd 0.034

Quality of prediction of GCA

Quality of prediction of GCA

Based on the result of previous section, we will further only consider results obtained with the GBLUP model based on the SNP kinship and without including the population structure. For the GBLUP model considering SCA, averaged correlations between observed ls-means and predicted dent GCA (respectively flint GCA) of the hybrids of the validation set were calculated over the 100 replications. For the dent GCA, they varied between 0.386 (DMC) and 0.616 (PH). For the flint GCA, they varied between 0.328 (DMY) and 0.671 (DMC) (Table 3).

These correlations should be put in relation with the percentage of dent GCA variance (respectively flint GCA variance) in the total phenotypic variance, which varied depending on the trait. The dent GCA variance represented between 25.15% (DMC) and 59.97% (DMY) of the total genetic variance. (Table 3). The flint GCA variance represented between 22.63% (DMY) and 63.14% (DMC) of the total genetic variance. For DMY, DtSILK and PH, the proportion of dent GCA variance in the genetic variance was more important than the proportion of flint GCA variance whereas it was the reverse for DMC.

Thus, based on predictive abilities we computed observed accuracies of the GCA predictions. The observed accuracies were high and varied between 0.739 (DtSILK) and 0.827 (DMC) for the dent GCA and between 0.830 (DMY) and 0.938 (DtSILK) for the flint GCA (Table 3). For all traits, flint GCAs were better predicted than dent GCAs.

Table 3: Quality of prediction of dent and flint GCAs for the best GBLUP model considering SCA for the four studied traits (Dry Matter Content DMC, Dry Matter Yield DMY, Silking date DtSILK, Plant Height PH). For the correlations, the standard deviation is indicated.

	DMC	DMY	DtSILK	PH
Average correlation between predicted dent GCA and calculated adjusted means	0.386 sd 0.060	0.539 sd 0.049	0.490 sd 0.046	0.616 sd 0.039
Average correlation between predicted flint GCA and calculated adjusted means	0.675 sd 0.030	0.350 sd 0.056	0.476 sd 0.047	0.452 sd 0.046
%GCA dent variance in the total genetic variance	25.15	59.97	52.82	59.89
%GCA flint variance in the total genetic variance	63.14	22.63	30.93	27.67
Observed accuracy of the dent GCA	0.827	0.786	0.739	0.819
Observed accuracy of the flint GCA	0.913	0.830	0.938	0.884

Effect of the size of the calibration set

Predictive ability increased with the size of the calibration set, especially for the small sizes of calibration set (Figure 2). Similar trends were observed for the different traits. Prediction accuracies reached plateau for calibration sets including more than 12 individuals per dent-flint populations (432 hybrids). No difference was observed between the model with SCA and

without SCA except for the small sizes of calibration set (72 and 108 hybrids) where the model without SCA performed better than the one with SCA.

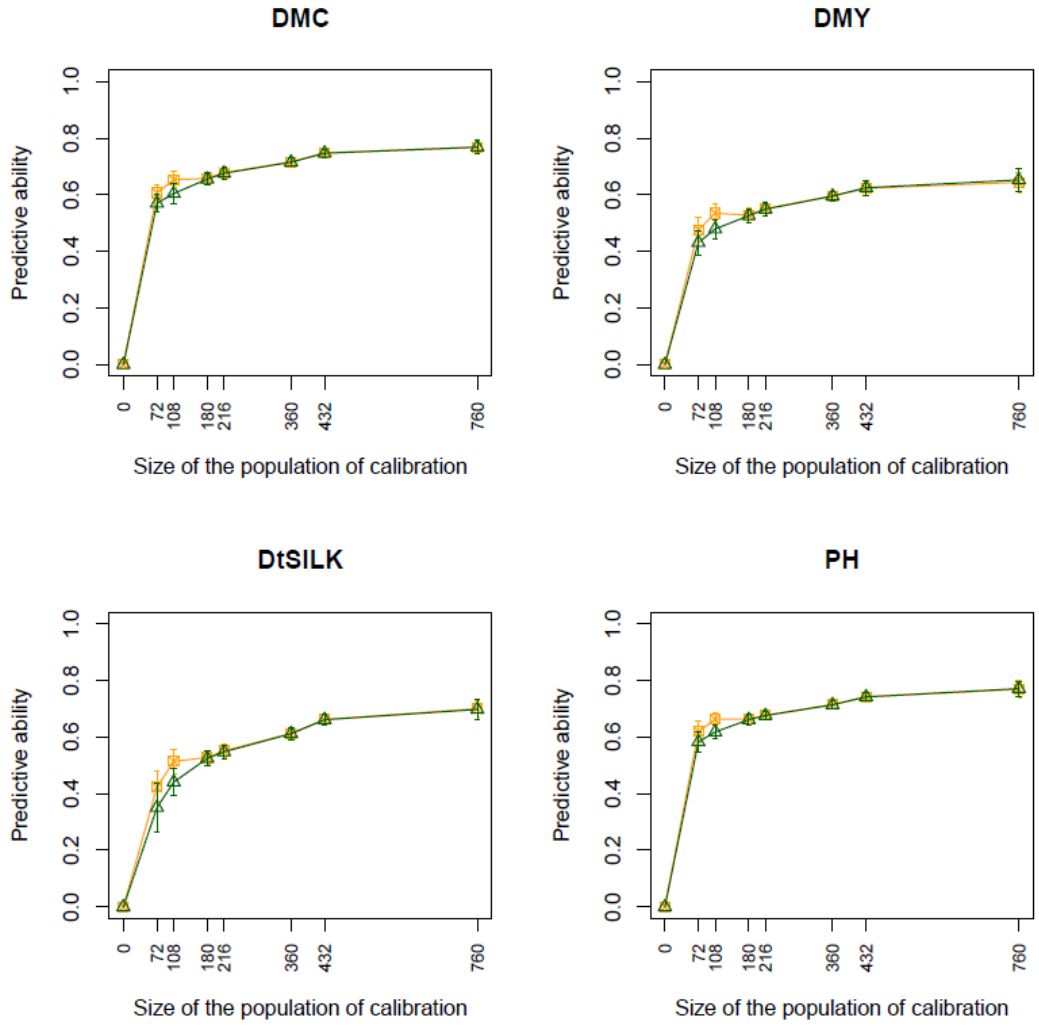

Figure 2: Predictive ability as a function of the calibration set size for the GBLUP models without population structure and with the SNP kinship matrices for the studied traits (Dry Matter Content DMC, Dry Matter Yield DMY, Silking date DtSILK, Plant Height PH). The green curve corresponds to the model taking into account the SCA, the yellow curve corresponds to the model without SCA.

Calibration set composition

Impact of including half-sibs hybrids in calibration set

For all traits, no special trend was observed for the predictive ability of the hybrids whether zero parent (T0), their dent parent (T1 Dent), their flint parent (T1 Flint) or both of their parents (T2) were parents of hybrids of the calibration set (Table 4). Considering SCA in our model did not change this result. Depending on the trait, the best predictive abilities were obtained for the

T2 (PH), T1D (DMC, DtSILK) or T1C (DMY) calibrations. T0 hybrids were never the best predicted ones.

Table 4: Predictive ability in function of the presence of half-sibs in the calibration set for the studied traits (Dry Matter Content DMC, Dry Matter Yield DMY, Silking date DtSILK, Plant Height PH), for the GBLUP models considering the SNP kinship matrices and no population structure. Standard deviations are indicated.

	DMC	DMY	DtSILK	PH
GCA model				
T0 hybrids	0.759 sd 0.034	0.613 sd 0.061	0.704 sd 0.043	0.760 sd 0.042
T1 Flint hybrids	0.743 sd 0.072	0.720 sd 0.084	0.708 sd 0.090	0.730 sd 0.084
T1 Dent hybrids	0.821 sd 0.049	0.688 sd 0.082	0.724 sd 0.079	0.784 sd 0.060
T2 hybrids	0.765 sd 0.066	0.692 sd 0.079	0.660 sd 0.109	0.822 sd 0.052
GCA + SCA model				
T0 hybrids	0.758 sd 0.034	0.615 sd 0.059	0.704 sd 0.042	0.760 sd 0.042
T1 Flint hybrids	0.743 sd 0.072	0.727 sd 0.085	0.713 sd 0.091	0.730 sd 0.083
T1 Dent hybrids	0.821 sd 0.049	0.685 sd 0.080	0.726 sd 0.079	0.784 sd 0.060
T2 hybrids	0.764 sd 0.067	0.700 sd 0.080	0.657 sd 0.111	0.821 sd 0.052

Impact of including hybrids issued from the same founder lines in calibration set

Predictive abilities increased with the level of pedigree relatedness between the validation set and the calibration set (Table 5). They varied between 0.123 (DMC – case4 1D) and 0.539 (DMC – case 1 2D-1F) when not including SCA in the model and between 0.0651 (DMC – case 4 1D) and 0.542 (DMC – case 1 2D-1F) when including SCA. Including SCA in the model only slightly changed the predictive abilities. With and without SCA, calibration with populations which shared only one founder line with the target population always gave the worst predictive abilities. Using in the calibration set populations which shared three founder lines with the target population was better than when using populations which shared only two founder lines, except for DMY (predictive ability with the 1D-2F case was lower than in the 1D-1F case). Standard deviations of the estimations were important, illustrating the differences in the quality of prediction that we observed depending on the target population considered.

Table 5: Average of the predictive abilities of a targeted population in function of the composition of the calibration set for the studied traits (Dry Matter Content DMC, Dry Matter Yield DMY, Silking date DtSILK, Plant Height PH), for the GBLUP models considering the SNP kinship matrices and no population structure

	DMC	**DMY**	**DtSILK**	**PH**
GCA model				
Case 1: 2 dent – 1 flint	0.539 sd 0.132	0.423 sd 0.167	0.436 sd 0.184	0.426 sd 0.197
Case 2: 1 dent – 2 flint	0.453 sd 0.179	0.341 sd 0.211	0.464 sd 0.139	0.430 sd 0.196
Case 3: 1 dent – 1 flint	0.438 sd 0.186	0.372 sd 0.210	0.407 sd 0.211	0.417 sd 0.202
Case 4: 1 dent	0.123 sd 0.201	0.201 sd 0.219	0.184 sd 0.211	0.261 sd 0.176
Case 5: 1 flint	0.309 sd 0.199	0.226 sd 0.164	0.295 sd 0.205	0.240 sd 0.240
GCA + SCA model				
Case 1: 2 dent – 1 flint	0.542 sd 0.130	0.426 sd 0.169	0.437 sd 0.184	0.427 sd 0.193
Case 2: 1 dent – 2 flint	0.456 sd 0.178	0.338 sd 0.208	0.458 sd 0.138	0.422 sd 0.195
Case3: 1 dent – 1 flint	0.436 sd 0.194	0.376 sd 0.215	0.407 sd 0.217	0.418 sd 0.205
Case 4: 1 dent	0.0651 sd 0.215	0.143 sd 0.226	0.0882 sd 0.266	0.256 sd 0.175
Case 5: 1flint	0.306 sd 0.199	0.241 sd 0.168	0.274 sd 0.229	0.238 sd 0.253

DISCUSSION

GCA/SCA Variance components

The lack of gain in efficiency of our prediction models including SCA effects is consistent with the small proportion of the SCA variance component in our analyses. To our knowledge, no study estimated SCA/GCA components on European silage maize using marker-based genomic relationships whereas several estimations are available for grain maize. Even if silage and grain yield are correlated, they correspond to different traits measured on plants harvested at different biological stages. Keeping this is mind, we found for DMY a percentage of SCA in the total genetic variance of 8.17% close to the one found on grain yield by Schrag *et al.* (2010) and Technow *et al.* (2014) but lower than the one found by Massman *et al.* (2013a). The same authors estimated the proportion of SCA in the genetic variance between 6.2 and 10% for grain moisture content whereas it was null in our design for DMC. Note that these studies involved inbred lines that had been selected before on their combining ability which may have increased the relative importance of SCA in the hybrids.

Proportion of SCA in the genetic variance was much lower when estimated on ls-means using

genotypic information than when estimated directly on field plot data (Supplementary Table S1). One explanation could be that the number of markers was not sufficient for capturing well SCA effects through kinship matrices. However, our design is structured in biparental families, with a large within-family LD extent and one can reasonably assume that the number of markers was *a priori* high enough to estimate kinship between lines at the within family level. Another explanation could therefore be that an important part of SCA in our design can be due to epistasis and was not reflected through our kinship matrices.

Comparison of the GBLUP models

We compared several GBLUP models. Considering structure in populations as a fixed effect in the GBLUP models decreased the predictive abilities. This suggests that the structure in population is well captured by the markers and that there is no need to add it in the model. It would be interesting to see if it would be the same for more structured traits, such as digestibility traits where the variation in each heterotic group is mainly due to the use of one founder line carrying favorable alleles. Considering founder allele kinship matrices and not SNP kinship matrices also decreased the predictive abilities. Founder allele matrices set the relatedness between two individuals from populations with no founder in common to zero. This shows that it is beneficial to take into account co-ancestry between founder lines. It may also indicate that causal mutations are not recent, and so are well captured by SNP.

Adding SCA in the GBLUP model did not increase the predictive abilities and could even decrease it when using a small training population size. This observation is consistent with the one of Zhao *et al.* (2013) on rice, who observed higher prediction accuracies when ignoring dominance effects than when including them. They related it among other things to the small size of their dataset. Using simulations, BayesB models were shown to slightly outperform GBLUP models when incorporating dominance (Technow *et al.* 2012) but this result was not consistent with empirical studies (Technow *et al.* 2014). Reproductive kernel Hilbert space regression may be an option for including non-additive effects such as dominance or epistasis (Gianola *et al.* 2006; Gianola and van Kaam 2008). Wellmann and Bennewitz (2012) proposed a Bayesian linear regression model called BayesD (D standing for dominance) which allowed on simulated data an improvement of the prediction accuracies compared to GBLUP.

Comparison with QTL - Combining the two approaches

Genomic predictions gave better results than QTL based predictions for all traits, but to a smaller extent for DtSILK. The small differences between the two models is a bit surprising. It

contrasts with the clear advantage of genomic selection over QTL-based marker-assisted selection schemes found by Bernardo and Yu (2007) on simulations and validated experimentally by Massman *et al.* (2013b) It suggests that including QTLs information in GBLUP models may improve the predictions accuracy by combining the two sources of information, giving more weight in prediction to the most important QTLs than when considering only relatedness matrices. Brøndum *et al.* (2015), on dairy cows, included in their GBLUP model a second component with its own variance corresponding to markers identified as linked to QTL from association mapping on sequence data. It led to an increase in the predictive abilities, especially for production traits presenting major QTLs. Zhao *et al.* (2014) on hybrid wheat proposed a weighted best linear unbiased prediction (W-BLUP) model for treating the effects of known functional markers by considering their additive and dominance effects as fixed. This model allowed an improvement of the prediction accuracies in comparison to marker-based predictions and to genomic predictions using RR-BLUP and BayesCπ models. Brard and Ricard (2015), on French trotters, also observed an increase in prediction accuracies when including in the GBLUP model as fixed effect genotype at a SNP strongly associated with the studied traits.

In our case, including the markers associated to QTL with marker effects identified previously in Chapter 2, especially the ones presenting the strongest effects, would probably increase the predictive abilities. We would have much less markers to add in our model than Brøndum *et al.* (2015) and would add them (or at least the most significant ones) as fixed effects as in Zhao *et al.* (2014) and Brard and Ricard (2015). Adding the QTL effects estimated with the hypothesis that all the founder lines carry a different allele, would be probably be the best option as it can account for multiallelic QTLs (cf. Chapter 2).

Optimization of the calibration set for hybrid prediction

Increasing the size of the calibration set allowed to increase the prediction ability but a plateau was reached for about 10-12 hybrids per dent-flint hybrid populations that is to say for 360 to 432 hybrids in total. Stagnation of the predictive abilities when reaching a certain size of calibration set was already observed on hybrids from advanced breeding selection stages predicted by GBLUP by Technow *et al.* (2014) or predicted by support vector machine regression and GBLUP by Maenhout *et al.* (2010). Same observation was done in studies willing to predict maize test-cross values in connected biparental populations as in Riedelsheimer *et al.* (2013) and Lehermeier *et al.* (2014). It is interesting from a practical point of view to note that based on this result it seems possible to consider only 10 hybrids per dent-

flint population without decreasing predictive abilities.

Few differences in terms of prediction accuracies were observed between T0, T1 Dent, T1 Flint and T2 hybrids contrary to observations made by Shrag *et al.* (2010), Massman *et al.* (2013a) and Technow *et al.* (2012, 2014) on simulations and real datasets. This may due to differences in design characteristics between studies. In our case, T1 hybrids in the validation set have generally a single half-sib hybrid in the training set whereas this number can be much more important in other studies (*e.g.* Figure S1 in Technow *et al.* 2014). Conversely, for a T0 hybrid, there will be in the calibration set many hybrids whose parents will be half-sibs or full-sibs of its parents. We showed the importance of hybrids issued from half-sib and full-sibs lines when predicting new populations, presenting one, two or three founder lines in common with the hybrids of the calibration set. The worst predictive abilities were obtained when calibration set included hybrids that only had one founder line in common with the population in the validation set. Increasing the number of founders in common between calibration and validation set had a strong positive effect of the predictive abilities. This is consistent with results obtained on test-cross hybrids issued from a multi-parental connected design by Riedelsheimer *et al.* 2013.

Reconsidering hybrid breeding selection schemes?

Usually, at early selection stages, parental lines are selected based on their test-cross value with one or a few testers of the complementary heterotic group. Until recently this screening was made in parallel to self-fecundation for production of homozygous lines. Double haploid technology makes it possible to obtain directly large segregating populations of inbred lines. This facilitates the production and evaluation of large sets of inter-group single-cross hybrids between new unselected inbred lines. We obtained for such material prediction accuracies close or superior to the ones obtained on test-cross designs. Indeed, Lehermeier *et al.* (2014) on silage test-cross hybrids obtained on average predictive abilities equal to 0.41 for their dent biparental populations and 0.48 for their flint ones for DMY, to be compared to 0.65 in our study. For DMC, they were equal to 0.58 for the dent and 0.52 for the flint, to be compared to 0.77 in our study. This most likely relates to a higher genetic variance in our design due to the contribution of the two genetic groups to the total variance, instead of a single group when a tester is used. A next question is therefore to which extent the quality of GCA prediction with our approach can be compared with that of test-cross value. The decomposition we used highlights high accuracies of GCA for both flints and dents (0.83 and 0.79 for DMY, respectively).

Further studies are needed to compare predictions accuracies in our design with those that could be obtained in a tester-based design. If the advantage of our design, or at least its equivalence,

in terms of prediction accuracies compared to test-cross design, is confirmed, it opens new perspectives for maize breeding. In test-cross evaluation, the dent lines should be evaluated in cross with at least one tester and similarly for the flint lines. This necessitates at least two times more phenotyping efforts than in our approach, which therefore could lead to a substantial economy in phenotyping. However, hand-made pollination is needed for creating our hybrids contrary to test-cross hybrids, which increases the production costs of a given quantity of hybrid seed. Further studies are needed to evaluate if the diminution of field plots for the hybrid phenotyping compensates this increased hybrid production cost. However, as a first approximation this increase will be about 100%, considering that a single hybrid is produced instead of two. Moreover, the GCA estimates using testers are biased by the SCA with the tester, whose choice is really important. Selection of the lines at early selection stages is conditioned by their specific combining abilities with the tester. This restricts the field of possible combinations at the second stage of hybrid testing. Our design makes it possible to circumvent these limits.

Conclusion

Our design, based on hybrids between two heterotic groups, allowed us to obtain good predictive abilities for the hybrid value and specifically for the dent and flint GCA. It highlights that genomic selection has the potential to uncorrelate the GCAs of the parents of hybrids with very economical designs in which most inbreds are only represented in a single hybrid. Further design optimization is needed as well as comparative studies with tester designs but we already obtained encouraging results for revisiting maize hybrid breeding selection scheme. Hybrids between unselected biparental populations of two reciprocal heterotic groups could be evaluated in the first steps of the selection scheme and used for estimating values of all the potential hybrids that could be derived from the available inbred lines, even if not involved in phenotyped hybrids. It would allow to explore much more combinations and diversity and will not restrain the field of possible parental lines combinations. So, the use of genomic selection expands the possibility of classical recurrent reciprocal selections schemes, which have been shown to be the most efficient among phenotype-based selection schemes (Coors, 1999).

Acknowledgements

H. Giraud was funded by Université Paris Sud through a Scholarship from the Ministry of Research given by the Doctoral School ABIES (Agriculture Food Biology Environment

Health). We are grateful to Caussade Semences, Euralis Semences, Limagrain Europe, Maïsadour Semences, Pioneer Genetics, R2n and Syngenta seeds grouped in the frame of the ProMais SAM-MCR program for the funding, inbred lines development, hybrid production and phenotyping. We also thank scientists from these companies for helpful discussion on the results. We thank the INRA metaprogram SelGen and specifically the X-Gen and SelDir projects coordinated respectively by Andrés Legarra and Hélène Gilbert.

Literature cited

Bernardo, R., 1992 Relationship between single-cross performance and molecular marker heterozygosity. Theor. Appl. Genet. 83: 628-634.

Bernardo, R., 1994 Prediction of maize single-cross performance using RFLPs and information from related hybrids. Crop Sci. 34: 20–25.

Bernardo, R., 1996a Best linear unbiased prediction of the performance of crosses between untested maize inbreds. Crop Sci. 36: 50–56.

Bernardo, R., 1996b Best linear unbiased prediction of maize single-cross performance. Crop Sci. 36: 872–876.

Bernardo R, and J. Yu, 2007 Prospects for genomewide selection for quantitative traits in maize. Crop Sci. 47: 1082–1090.

Brard, S., and A. Ricard, 2015 Should we use the single nucleotide polymorphism linked to *DMRT3* in genomic evaluation of French trotter? Journal of Animal science 93: 4651-4659.

Brøndum, R. F., G. Su, L. Janss, G. Sahana, B. Guldbrandtsen *et al.*, 2015 Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. J. Dairy Sci. 98: 4107-4116.

Browning, B. L., and S. R. Browning, 2009 A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84: 210–223.

Burstin, J., A. Charcosset, Y. Barrière, Y. Hébert, D. de Vienne *et al.*, 1995 Molecular markers and protein quantities as descriptors in maize. II. Prediction of performance of hybrids for forage traits. Plant Breeding 114: 427-433.

Butler, D. B. R. Cullis, A. R. Gilmour, and B. J. Gogel, 2007 ASReml-R reference manual. The State of Queensland, Department of Primary Industries and Fisheries.

Charcosset, A., M. Lefort-Buson, and A. Gallais, 1991 Relationship between heterosis and heterozygosity at marker loci: a theoretical computation. Theor. Appl. Genet. 81: 571-

575.

Charcosset, A., and L. Essioux, 1994 The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. Theor. Appl. Genet. 89: 336-343.

Comstock, R.E., H.F. Robinson, and P.H. Harvey, 1949 A breeding procedure designed to make maximum use of both general and specific combining ability. J. Am. Soc. Agron. 41:360–367.

Coors, J. G., 1999 Selection methodology and heterosis. In *The genetics and exploitation of heterosis in crops. Proceedings of an international symposium, CIMMYT, Mexico City, Mexico, 17—22 August 1997.*

Dudley, J. W., M. A. Saghai-Maroof, and G. K. Rufener, 1991 Molecular markers and grouping of parents in maize breeding programs. Crop Sci. 31: 718-723.

Ganal, M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler *et al.*, 2011 A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS ONE 6: e28334.

Gianola, D., R. L. Fernando, and A. Stella, 2006. Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics 173: 1761–1776.

Gianola, D., and J. B. C. H. M. van Kaam, 2008 Reproducing Kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. Genetics 178: 2289–2303.

Godhsalk, E. B., M. Lee, and K. R. Lamkey, 1990 Relationship of restriction fragment length polymorphisms to single-cross hybrid performance in maize. Theor. Appl. Genet. 80: 273-280.

Gowda, M., Y. Zhao, H. P. Maurer, E. A. Weissman, T. Würschum *et al.*, 2013 Best linear unbiased prediction of triticale hybrid performance. Euphytica 191: 223-230.

Hallauer, A. R., M. J. Carena, and J. B. Miranda Filho, 2010 *Quantitative genetics in maize breeding.* Handbook of plant breeding, 6. Springer.

Hickey, J. M., G. Gorjanc, R. K. Varshney, and C. Nettelblad, 2015 Imputation of single nucleotide polymorphism genotypes in biparental, backcross, and topcross populations

with a hidden Markov model. Crop Sci. 55: 1934-1946.

Lee, M., E. B. Golshalk, K. R. Lamkey, and W. L. Woodman, 1989 Association of restriction length polymorphism among maize inbreds with agronomic performance of their crosses. Crop Sci. 29: 1067-1071.

Lehermeier, C., N. Krämer, E. Bauer, C. Bauland, C. Camisan *et al.*, 2014 Usefulness of multiparental populations of maize (*Zea mays* L.) for genome-based prediction. Genetics 198: 3-16.

Maenhout, S., B. De Baets, G. Haesaert, and E. Van Bockstaele, 2007. Support vector machine regression for the prediction of maize hybrid performance. Theor. Appl. Genet. 115: 1003-1013.

Maenhout, S., B. De Baets, and G. Haesaert, 2010 Prediction of maize single-cross hybrid performance: support vector machine regression versus best linear prediction. Theor. Appl. Genet. 120: 415-427.

Massman, J. M., A. Gordillo, R. E. Lorenzana, and R. Bernardo, 2013a Genomewide predictions from maize single-cross data. Theor. Appl. Genet. 126: 13-22.

Massman, J. M., H.-J. G. Jung, and R. Bernardo, 2013b Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. Crop Sci. 53: 58-66.

Melchinger, A. E., M. Lee, K. R. Lamkey, and W. L. Woodman, 1990 Genetic diversity for restriction fragment length polymorphisms: Relation to estimated genetic effects in maize inbreds. Crop Sci. 30: 1033-1040.

Melchinger, A. E., J. Boppenmaier, B. S. Dilhon, W. G. Pollmer, and R. G. Hermann, 1992 Genetic diversity for RFLPs in European maize inbreds: II. Relation to performance of hybrids within versus between heterotic groups for forage traits. Theor. Appl. Genet. 84: 672-681.

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

R Core Team, 2013 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reif, J. C., Y. Zhao, T. Würschum, M. Gowda, and V. Hahn, 2013 Genomic prediction of sunflower hybrid performance. Plant Breeding 132: 107-114.

Riedelsheimer, C., J. B. Endelman, M. Stange, M. E. Sorrells, J.-L. Jannink *et al.*, 2013 Genomic predictability of interconnected bi-parental maize populations. Genetics 194: 493–503.

Schrag, T. A., J. Möhring, A. E. Melchinger, B. Kusterer, B. S. Dhillon *et al.*, 2010 Prediction of hybrid performance in maize using molecular markers and joint analyses of hybrids and parental inbreds. Theor. Appl. Genet. 120: 451-461.

Shull, G. H. 1914 Duplicate genes for capsule-form in Bursa bursa pastoris. Zeitschrift ind. Abst. U. Verebsgl. 12: 97-149.

Sprague, G. F., and L. A. Tatum, 1942 General vs. specific combining ability in single crosses of corn. J. Am. Soc. Agron. 34: 923-932.

Technow, F., C. Riedelsheimer, T. A. Schrag, and A. E. Melchinger, 2012 Genomic prediction of hybrid performance in maize with models incorporating dominance and population specific marker effects. Theor. Appl. Genet. 125: 1181-1194.

Technow, F., T. A. Schrag, W. Schipprack, E. Bauer, H. Simianer *et al.*, 2014 Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. Genetics 197: 1343-1355.

Van Raden, P. M., 2008 Efficient methods to compute genomic predictions. J. Dairy Sci. 91: 4414-4423.

Wang, Y., M. F. Mette, T. Miedaner, M. Gottwald, P. Wilde *et al.*, 2014 The accuracy of prediction of genomic selection in elite hybrid rye population surpasses the accuracy of marker-assisted selection and is equally augmented by multiple field evaluation locations and test years. BMC Genomics 15: 556.

Wellmann, R., and J. Bennewitz, 2012 Bayesian models with dominance effects for genomic evaluation of quantitative traits. Genet. Res. 94: 21-37.

Xu, S., D. Zhu, and Q. Zhang, 2014 Predicting hybrid performance in rice using genomic best linear unbiased prediction. Proc. Natl. Acad. Sci. U. S. A. 111: 12456-12461.

Zhao, Y., J. Zeng, R. Fernando, and J. C. Reif, 2013 Genomic prediction of hybrid wheat performance. Crop Sci. 53: 802-810.

Zhao, Y., M. F. Mette, M. Gowda, C. F. H. Longin, and J. C. Reif, 2014 Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. Heredity 112: 638-645.

General discussion

Maize is a crop of major importance in the world presenting strong heterosis for traits related to biomass. Silage maize represents around 44% of the total surface grown with maize in France. To benefit from heterosis, F_1 hybrids are cultivated in majority, although, to diminish seed production costs, one observed a significant return to three way hybrids involving a female hybrid between related lines (Lorgeou, Pers. Com.). Maize hybrid selection scheme is mostly based on the reciprocal recurrent selection scheme initially proposed by Comstock *et al.* (1949) that was adapted in the breeding companies. In the first step, parental lines of each heterotic group are evaluated and selected for their value when crossed to one or several testers of the complementary heterotic group. It is only in the last stages that a larger set of combinations between a few selected lines of each heterotic group are evaluated. Hybrid genetic value can be decomposed in General Combining Abilities (GCA) of the parental lines and Specific Combining Ability (SCA) (Sprague and Tatum 1942). Understanding the genetic architecture underlying hybrid value for traits of interest for breeding is of main importance from both theoretical and applied points of view. We studied two multiparental designs of hybrids between the dent and flint heterotic groups, mainly used for silage breeding in Northern Europe. The first design (Chapter 1) consisted in two Nested Association Mapping (NAM) designs. Parental lines of one NAM design were evaluated for their test-cross value with the central line of the complementary design used as a tester. In the second design or multiparental reciprocal design (Chapters 2 and 3), we evaluated hybrids issued from a factorial design between two multiparental connected designs. QTL detection was performed in both cases. The potential of genomic selection was evaluated in the NAM design by Lehermeier *et al.* (2014) and we performed genomic predictions in the second one. Our study shed new lights on the genetic architecture below hybrid value for silage maize and opened prospects for the use of genomic information in new hybrid selection schemes.

QTL detection in multiparental designs

Multiallelic QTLs heterotic group dependent

Numerous QTLs were detected in the two multiparental studied designs allowing a better understanding of the genetics of hybrid value of silage maize.

QTL detection on the test-cross hybrids of the NAM designs could be performed with MCQTL_LD (Jourjon e*t al.* 2005) whereas for studying the hybrids of the factorial design, we had to use more complex mixed QTL detection models and implement them in ASReml-R (Butler *et al.*. 2007). Different ways of coding genotypes were used and LA and LDLA models were applied. Some of the QTL detection models used allowed us to detect allelic series in both

designs, which showed the importance of not directly considering the observed genotype but a haplotypic or a founder information, especially for complex traits such as yield. Compared to previous studies (Bardol *et al.* 2013), LDLA models using ancestral haplotypic information (modelled using "clusthaplo", Leroux *et al.* 2014) yielded closer results to those of the parental allele model, probably because of the limited relatedness between founder lines.

Both the NAMs and the factorial design highlighted different QTLs for dent and flint lines, consistent with the long term divergence between groups (Tenaillon and Charcosset 2011). Interestingly no QTL showed a major effect for DMY and DMC. A notable exception was a major QTL for flowering time, with pleiotropic effects on the other traits, which appeared on chromosome 10 in both studies in the flint group and most likely corresponds to the *ZmCCT* gene previously reported by Ducrocq *et al.* (2009). It would be interesting to compare more carefully in terms of number and position, the QTLs detected in our designs with the ones detected on silage on dent and flint diversity panels using association mapping methods by Rincent *et al.* (2014) and other results published on silage yield. Comparison of the detection power in the two designs is not straightforward as the diversity addressed in the factorial design and the number of hybrids evaluated were much lower than in the NAMs. It is nevertheless interesting to note the number of QTL where only slightly lower in the factorial.

Finally, it can be noted that only very few QTL were detected for dominance effects with the factorial design. The contrast with the numerous dominance QTL reported by Larièpe *et al.* (2012) is most likely due to the presence of strong consanguinity in tested hybrids in her study *vs.* absence of consanguinity in our factorial design.

Prospects for QTL detection

In the reciprocal multiparental design, one founder out of four in each heterotic group was chosen with the objective to bring favorable alleles for digestibility. Hybrids were phenotyped for various digestibility traits such as DINAG (digestibility of non-starch and insoluble carbohydrate under the hypothesis that starch and soluble sugars are completely digestible, Argillier and Barrière 1996), DINAGZ (extension of the DINAG to nitrogen compounds, Barrière and Emile 2000), Milk Forage Unit (MFU), digestibility of the Neutral Detergent Fiber (NDF), of the Acid Detergent Fiber (ADF) and of the Acid Detergent Lignin (ADL). Thus, phenotypic data for traits directly related to silage maize performances (DINAG, DINAGZ, MFU, NDF) as well as data concerning cell wall composition (content in lignin, hemicellulose and cellulose derived from NDF, ADF and ADL) are available. Digestibility traits are expected to be mostly additive which was confirmed by the first variance component estimation results.

Except for MFU (7.5%) and NDF (10.9%) the percentage of SCA in the genetic variance was close to zero for all other traits. QTL detection for all these trait is currently in progress. It could be interesting to study in more details these traits and potentially fine-map some interesting QTLs. Fine-mapping of some QTLs detected in the NAM designs could also be considered even though the major detected QTLs seem to correspond to already known and fine-mapped QTLs.

In this thesis, QTL detection was carried out on adjusted means obtained over various environments. It would be interesting to perform QTL detection on data from each single environment or preferably to perform multi-environment QTL mapping. In his association study, Rincent *et al.* (2014) detected very few QTLs for silage yield when considering ls-means over environments but much more associations when considering separately each field trial. This suggests that Genotype x Environment (G x E) interactions are important to consider for silage traits. One option for performing multi-environment QTL detection would be to add environmental variables in our QTL detection models (Boer *et al.* 2007). In this approach, genetic variances for each environment and genetic correlations between environments are first modelled. Then QTL detection is performed decomposing the QTL effect in a main QTL effect and a QTL x Environment interaction effect. In the final step, both genotypic and environmental variables are included in order to model QTL responses on specific environmental variables. This approach can be extended to multi-trait multi-environment QTL detection models (Malosetti *et al.* 2008), with or without including environmental variables. It would be interesting to evaluate if this approach could be extended to our QTL detection model decomposing the genetic effect in dent GCA, flint GCA, and SCA. Such study would be interesting to evaluate the stability of SCA effects in different environments.

Genomic prediction in multiparental reciprocal designs

The idea of using genomic selection at the level of factorial designs traces to Bernardo (1994). To our knowledge it has been applied efficiently to phases of hybrid breeding programs where lines have passed a first selection for hybrid value and are used in mating designs with a number of other lines. We showed in our study that genomic selection on hybrids obtained through a highly incomplete factorial mating design between two multiparental unselected populations gave good predictive abilities even though the majority of the inbred lines were parents of only one hybrid. Prediction accuracies obtained in our reciprocal multiparental design on inter-group single-cross hybrids between unselected lines were superior to the ones obtained on test-cross hybrids for silage maize by Lehermeier *et al.* (2014). Our higher prediction accuracies may be

due to a higher genetic variance due to the contributions of the two genetics groups to the total variance and to the fact that both inter and intra-population variations are considered. To further investigate prediction accuracy, we estimated accuracies of GCA and obtained high values for both flints and dents (0.83 and 0.79 for DMY, respectively). This shows that it was possible to uncorrelate the two GCAs and opens interesting prospects for revisiting inbred line selection in early phases of hybrid selection programs. Further studies are however needed for comparing the quality of GCA prediction of our approach with those based on test-cross hybrids. To do so, it will be necessary to evaluate jointly (i) single-crosses between inbred lines of each multiparental design like in our design and (ii) test-crosses of the same lines using as testers with few inbred lines of the other heterotic group.

Whatever the approach used to evaluate GCA, genomic selection offers very promising application when coupled with double-haploid (DH) production technology. Development of double haploid technology now makes it possible to obtain directly large segregating populations of maize inbred lines without requiring several generations of self-fertilization. All maize populations do not respond in the same way to double haploid technology and for some, more particularly in the flint heterotic group, success rate may be low. Another issue is that DH lines do not always produce enough seeds to produce single-cross hybrids and another selfing generation is necessary. Genomic prediction model can be trained using single-crosses for the DH lines with enough seeds and used to predict GCAs of DH lines which did not have enough seeds available at time t. Best inbred lines from all populations will thus be chosen for entering in the next step of the breeding process. We also showed that on such material, we could predict with relative high accuracy one population which shared founders with populations of the training set. This opens prospects for at least a prescreening of inbred lines in new populations, as providing the relatedness of these with previous populations is high enough to guarantee prediction efficiency.

If our study highlights a good ability to predict GCA, it also shows very strong limits with respect to SCA. These may first relate to intrinsic genetics features. The estimation of SCA variance components that we obtained using the GBLUP model were much lower than estimates given directly by the analysis of field plot data which suggests that epistasis might be a major component of SCA. One however may not preclude that out inability to predict SCA can be due in part to the experimental design that we used. When designing the experiment, choice was made to evaluate a high number of parental lines in a limited number of hybrid combinations (generally one), which may have limited the possibility to calibrate the SCA component of the model.

Applied prospects of the reciprocal multiparental design experiment

Short-term applied applications are started or intended following QTL detection and genomic selection results, based on the genetic material already created.

Some hybrids phenotyped in the reciprocal multiparental design showed good potential to be registered in the French catalogue of varieties. They were included in the 2014 and 2015 trial networks for further evaluation and comparison to known varieties. A few hybrids confirmed their potential and will be incorporated in the evaluation official trials network in view of a potential registration. It has to be noted that more inbred lines were created for the project than the ones phenotyped in hybrid combinations. Genomic prediction of the GCA of these lines was performed for various traits as dry matter yield but also MFU and an index trait used as a criteria for the registration to the French Catalogue of silage varieties. It allowed us to choose the best hybrid combinations considering the whole set of available inbred lines. Hybrid seeds will be produced in winter 2015 and the hybrids will be evaluated in field trials in summer 2016. In parallel to these promising hybrids, other hybrids will be created and evaluated, chosen at random between all the possible combinations of inbred lines. They will allow us to study the quality of the genomic and QTL-based predictions.

Improving the parental inbred lines through Marker-Assisted Selection using the QTLs detected is also an interesting option. Choice of the crosses between the already existing parental lines for creating new material could be optimized in order to combine at the detected QTLs the maximum of favorable alleles. To do so, software such as OptiMAS (Valente *et al.* 2013) could be used. Combining information on QTLs detected for dry matter yield and dry matter content with information on QTLs for digestibility traits will be particularly interesting. Indeed, favorable alleles for digestibility traits are mostly bring by one dent and one flint founder lines which do not necessarily present the most favorable alleles for dry matter yield. It would be interesting in this context to also consider the genomic predictions of the lines selected based on their alleles at QTL, to take advantage of the two approaches and capture genetic variation not associated with detected QTL.

Prospects for enhancing genomic prediction efficiency

An alternative to our factorial design, at constant number of field single-plots, would have been to evaluate a more limited number of parental lines but in more hybrid combinations. Increasing the number of contribution of parental lines may help to better estimate the SCA of these lines and by extension of the lines to predict. This may however be at the cost of the number of lines evaluated for their GCA, so that a global optimization considering the different steps of the

breeding program is needed. Besides from the number of hybrids evaluated per parental line, another important aspect is the choice, among all possible hybrids of the one to phenotype. Criteria like CDmean (Rincent *et al.* 2012) which allowed optimization of the calibration sets in a diversity panel could be used for optimizing the choice of individuals in other types of designs. First studies are currently carried out on the NAM designs for optimizing in each group the calibration set in order to improve predictions made by Lehermeier *et al.* (2014) (Rincent, Moreau and Charcosset, unpublished). The CDmean approach was initially developed for genomic selection using additive models but it should be possible to extend it to the two GCA components (dent and flint in our design) and the SCA component of the hybrid value.

One option to explore would be to perform prediction models combining genomic and QTL information, for instance adding some of the detected QTLs as fixed effects in a GBLUP model. Such models proved their efficiency for improving prediction accuracies (Zhao *et al.* 2014; Brard and Ricard 2015). SCA was not well taken into account in our GBLUP model. Other models should be explored which may probably better take into account SCA: reproductive kernel Hilbert (Gianola *et al.* 2006; Gianola and van Kaam 2008), BayesD (Wellmann and Bennewitz 2012) among others.

Contribution of approaches similar to the Single Step GBLUP models (Legarra *et al.* 2014) should be explored. Indeed, breeding companies carry out multiple breeding programs in which many parental lines are evaluated but not necessarily genotyped. Huge amount of phenotypic data corresponding to different years and types of material is available. If reciprocal multiparental designs such as ours are implemented, using information from other individuals may improve the accuracy of the predictions. Indeed, inbred lines of heterotic groups of private breeding programs are strongly related and lines related to the ones we want to predict may have already been evaluated as single-crosses in former advanced breeding stages or as test-crosses in former early stages of breeding programs. These individuals were phenotyped but not necessarily genotyped. Thus they cannot be included as such in GBLUP models. However, if their pedigree is known they could be included in Single Step GBLUP models. Indeed this method considers an extended relationship matrix H between all individuals, combining pedigree and genomic kinship. Single Step GBLUP showed its interest in animal breeding but it has to be noted that in plant breeding, pedigree are less reliable than in animal breeding. In addition, G x E interactions are much more important making it difficult and possibly inefficient to assemble data from different years, different environments in a single analysis. Information from specific QTLs can be added in Single Step GBLUP models by giving a different weigh to some SNPs (Legarra *et al.* 2011; Wang *et al.* 2012). Future mentioned developments of Single

Step GBLUP models are extension to crosses and to fit dominance effects (Legarra *et al.* 2014) and one may hope theses aspects will be derived soon.

Reconsidering early steps of maize hybrid breeding selection programs?

Further studies are needed to evaluate the prediction accuracies on our design compared to those of tester-based designs. If advantage or at least equivalence of our reciprocal multiparental design is confirmed, it opens new prospects for reconsidering maize breeding. Indeed test-cross evaluation could be replaced by single-cross evaluation between the two heterotic groups according to an incomplete factorial design. Test-cross evaluation, by evaluating lines of each heterotic group using one or a few testers of the other one, needs at least two times more phenotyping effort than our design. This may lead to substantial economies. However, for creating our single-cross hybrids, hand-made pollination is necessary, which is not the case for test-cross hybrid production. Thus it increases the cost of hybrid seed production for an individual hybrid. Further studies are needed to evaluate if this increase in production cost is compensated by the diminution of field plots needed for phenotypic evaluation.

For a first evaluation and comparison of the cost of the two methods, we could consider the evaluation of 500 dent lines and 500 flint lines, each one observed in eight environments. For the test-cross evaluation, cost of seed production will be around 22€ per line (including the catch-up costs, considering that 10% of the pollinations failed). Considering that the dent and flint lines are evaluated in combination with only one tester, the seed production cost will be equal to (500 x 22 + 500 x 22) = 22 000€. For simplifying the cost evaluation, we will consider that all lines are observed only once in all environments and that there is no checks. 16 trials are needed for test-cross evaluation of all the lines (eight for the 500 dent lines and eight for the 500 flint lines). Considering a field cost evaluation of one hybrid in one environment of 20€, we have a total cost for the field trials of (8 x 20 x 500 + 8 x 20 x 500) = 160 000€. Thus evaluation of the lines based on their test-cross value will cost 182 000€, considering that each line should be seen in one environment. For direct evaluation of the hybrids between the 500 dent and 500 flint lines, we will consider that each line is parent of only one hybrid and we will thus need to produce 500 hybrids. Hand-made pollination will be needed for seed production and we will consider a cost of 50€ per line (including the catch-up costs, considering that 40% of the pollinations failed). Thus, the total cost for seed production is equal to (500 x 50) = 25 000€. Only 8 trials will be needed for evaluating all hybrids. Considering a field cost evaluation of one hybrid in one environment, (8 x 20 x 500) = 80 000€ will be needed for phenotypic evaluation. Thus, evaluation of the lines based on a reciprocal design as studied

during this phD will cost 105 000€. Considering eight environments and two heterotic groups, evaluation of 500 lines in each heterotic groups based on their cross value according to an incomplete factorial allows a decrease of the costs of around 40% in comparison to their evaluation in test-cross. This evaluation does not take into account the differences between the two designs in prediction accuracies for the GCA and SCA of the lines. If prediction accuracies in reciprocal designs are higher than in test-cross designs, it strengthens the advantages of the reciprocal design compared to the test-cross one. However, private breeding companies may be reluctant to apply such designs as they will need to reorganize their breeding selection schemes (currently improvement of both heterotic groups is not completely simultaneous) and as they may find risky to create by hand-made pollination hybrids when both parents are not fully characterized for their flowering date.

To conclude, even if further investigations are needed for comparing accuracies of reciprocal multiparental designs to test-cross designs and emphasize their advantages, our study shown its potential interest for maize hybrid selection. Maize hybrid selection schemes should be reconsidered in the light of the results presented in this phD manuscript.

Literature cited

Argillier, O., and Y. Barriere, 1996 *DINAG, une estimation de la qualité de la partie non grain du maïs ensilage sur des échantillons de plante entière, Colloque maïs ensilage*, Nantes, 17-18 sept. 1996.

Bardol, N., M. Ventelon, B. Mangin, S. Jasson, V. Loywick *et al.*, 2013 Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. Theor. Appl. Genet. 126: 2717–2736.

Barrière, Y., and J. C. Emile, 2000 Le maïs fourrage. III. Évaluation et perspectives de progrès génétique sur les caractères de valeur alimentaire. Fourrages 163: 221–238.

Bernardo, R., 1994 Prediction of maize single-cross performance using RFLPs and information from related hybrids. Crop Sci. 34: 20–25.

Boer, M. P., D. Wright, L. Z. Feng, D. W. Podlich, L. Luo *et al.*, A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. Genetics 177: 1801-1813.

Brard, S., and A. Ricard, 2015 Should we use the single nucleotide polymorphism linked to *DMRT3* in genomic evaluation of French trotter? Journal of Animal science 93: 4651-4659.

Butler, D. B. R. Cullis, A. R. Gilmour, and B. J. Gogel, 2007 ASReml-R reference manual. The State of Queensland, Department of Primary Industries and Fisheries.

Comstock, R.E., H.F. Robinson, and P.H. Harvey, 1949 A breeding procedure designed to make maximum use of both general and specific combining ability. J. Am. Soc. Agron. 41:360–367.

Ducrocq, S., C. Giauffret, D. Madur, V. Combes, F. Dumas *et al.*, 2009 Fine mapping and haplotype structure analysis of a major flowering time quantitative trait locus on maize chromosome 10. Genetics 183: 1555–1563.

Gianola, D., R. L. Fernando, and A. Stella, 2006. Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics 173: 1761–1776.

Gianola, D., and J. B. C. H. M. van Kaam, 2008 Reproducing Kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. Genetics 178: 2289–2303.

Jourjon, M.-F., S. Jasson, J. Marcel, B. Ngom, and B. Mangin, 2005 MCQTL: multi-allelic QTL mapping in multi-cross design. Bioinformatics 21: 128–130.

Larièpe, A., B. Mangin, S. Jasson, V. Combes, F. Dumas *et al.*, 2012 The genetic basis of heterosis: multiparental quantitative trait loci mapping reveals contrasted levels of apparent overdominance among traits of agronomical interest in maize (*Zea mays* L.). Genetics 190: 795-811.

Legarra, A., F. Calenge, P. Mariani, P. Velge, and C. Beaumont, 2011 Use of a reduced set of single nucleotide polymorphisms for genetic evaluation of resistance to *Salmonella* carrier state in laying hens. Poultry Sci. 90: 731-736.

Legarra, A., O. F. Christensen, I. Aguilar, and I. Misztal, 2014 Single Step, a general approach for genomic selection. Livest. Sci. 166: 54-65

Lehermeier, C., N. Krämer, E. Bauer, C. Bauland, C. Camisan *et al.*, 2014 Usefulness of multiparental populations of maize (*Zea mays* L.) for genome-based prediction. Genetics 198: 3-16.

Leroux, D., A. Rahmani, S. Jasson, M. Ventelon, F. Louis *et al.*, 2014 Clusthaplo: a plug-in for MCQTL to enhance QTL detection using ancestral alleles in multi-cross design. Theor. Appl. Genet. 127: 921-933.

Malosetti, M., J. M. Ribaut, M. Vargas, J. Crossa, and F. A. van Eeuwijk, A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). Euphytica 161: 241-257.

Rincent, R., D. Laloë, S. Nicolas, T. Altmann, D. Brunel *et al.*, 2012 Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). Genetics 192: 715-28.

Rincent, R., S. Nicolas, S. Bouchet, T. Altmann, D. Brunel *et al.*, 2014 Dent and flint maize diversity panels reveals important genetic potential for increasing biomass production. Theor. Appl. Genet. 127: 2313-2331.

Sprague, G. F., and L. A. Tatum, 1942 General *vs.* specific combining ability in single crosses of corn. J. Am. Soc. Agron. 34: 923-932.

Tenaillon, M. I., and A. Charcosset, 2011 A European perspective on maize history. C. R. Biol. 334: 221–228.

Valente, F., F. Gauthier, N. Bardol, G. Blanc, J. Joets, et al., 2013 OptiMAS: a decision support tool for marker-assisted assembly of diverse alleles. J. Heredity 104: 586-590.

Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. Muir, 2012 Genome-wide association mapping including phenotypes from relatives without genotypes. Genet. Res. 94: 73-83.

Wellmann, R., and J. Bennewitz, 2012 Bayesian models with dominance effects for genomic evaluation of quantitative traits. Genet. Res. 94: 21-37.

Zhao, Y., M. F. Mette, M. Gowda, C. F. H. Longin, and J. C. Reif, 2014 Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. Heredity 112: 638-645.

General appendices

Appendix I: Supporting information of Chapter 1

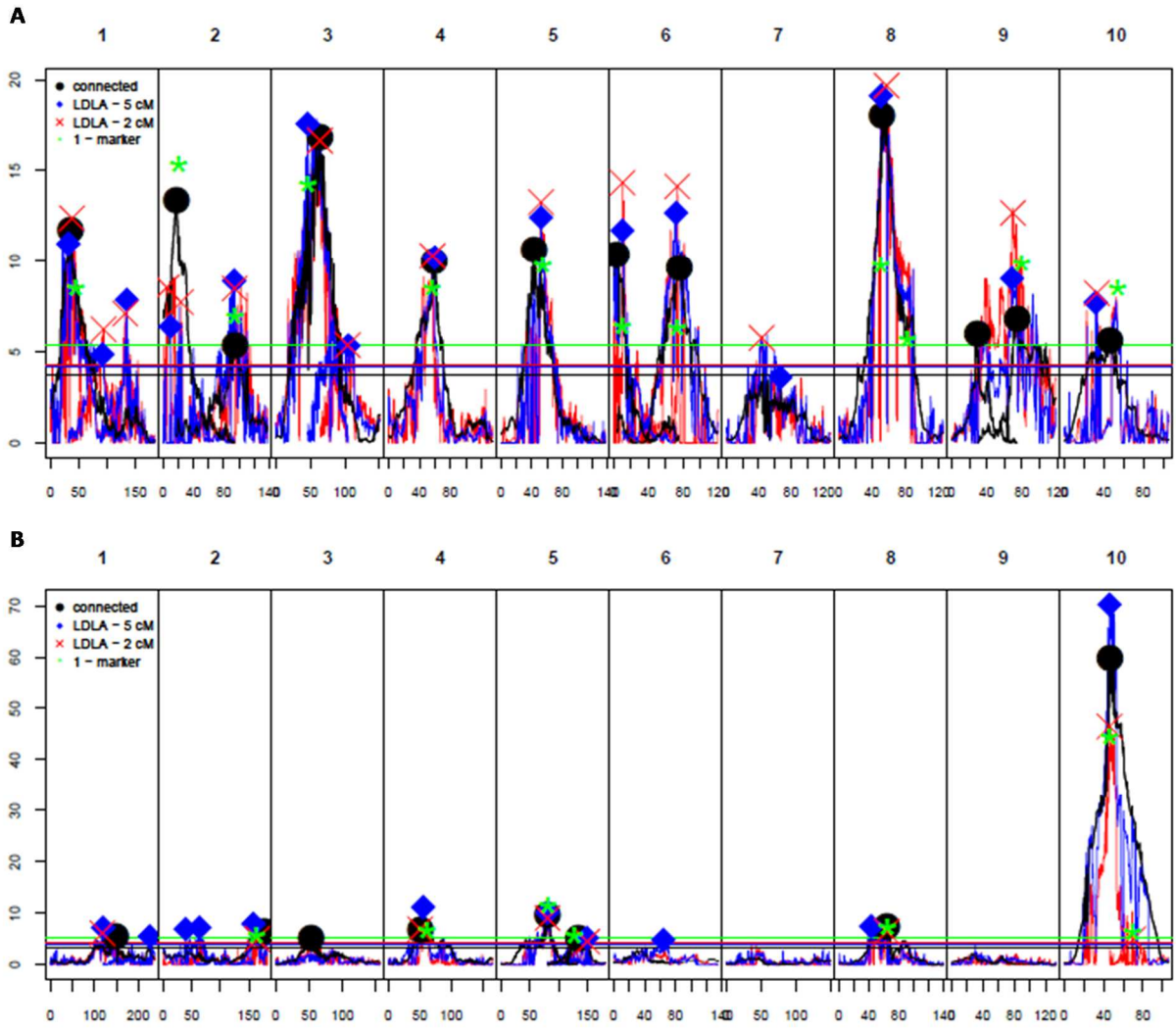

Figure S1 Results of the QTL detection with each model for DMC for **(A)** the dent design and **(B)** the flint design. The -log10(p-values) of the connected model are represented by black lines, the QTL positions of the connected models by black dots. The -log10(p-values) of the LDLA – 5 cM model are represented by blue lines and the QTL positions by blue diamonds. The -log10(p-values) of the LDLA – 2 cM model are represented by red lines and the QTL positions by red crosses. The -log10(p-values) of the QTL detected by the LDLA – 1-marker model are represented by green stars. Horizontal lines correspond to the threshold values of the different models.

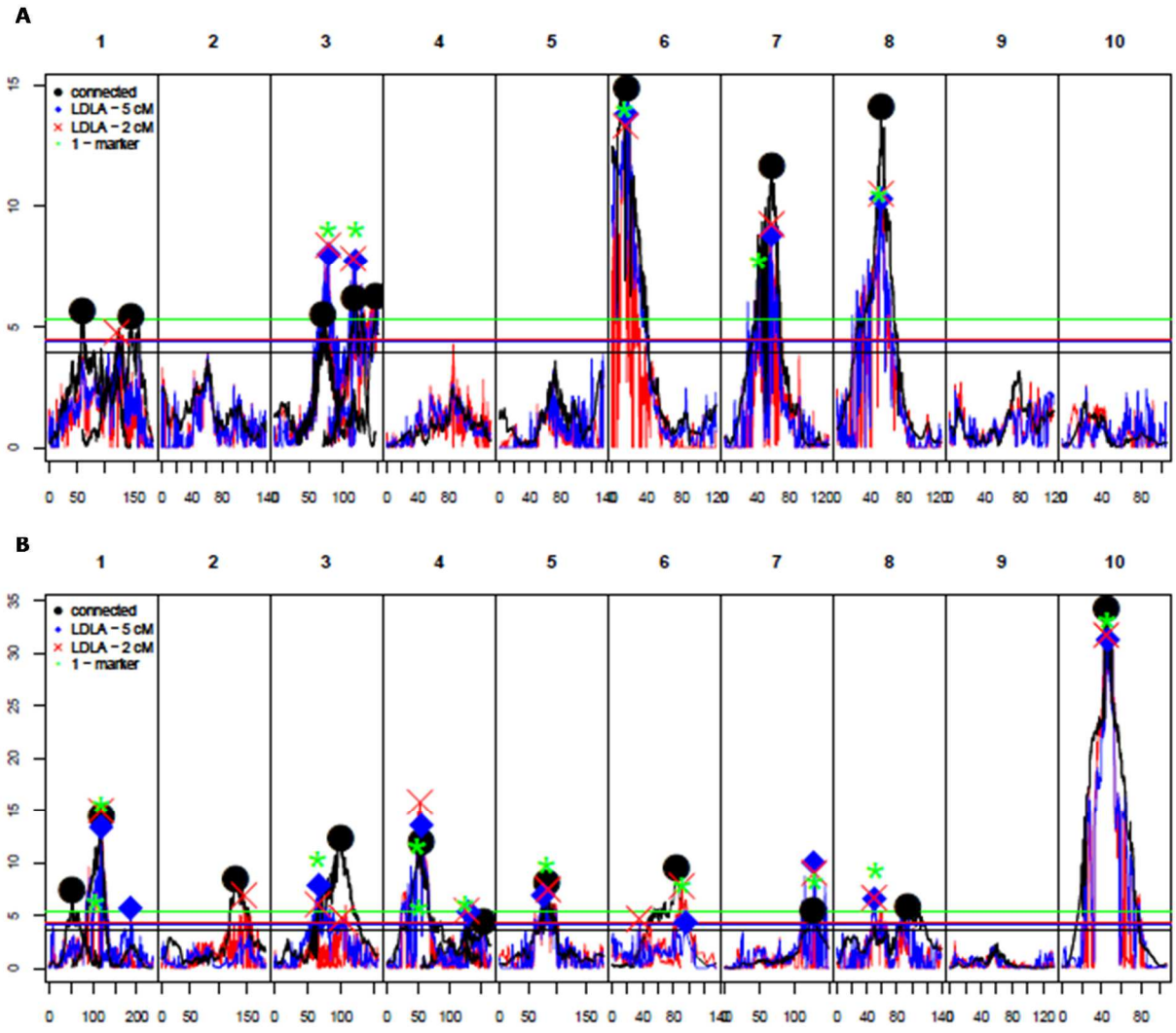

Figure S2 Results of the QTL detection with each model for DMY for **(A)** the dent design and **(B) the** flint design. The results for the dent design are in the superior part, flint in the inferior part. The -log10(p-values) of the connected model are represented by black lines, the QTL positions of the connected models by black dots. The -log10(p-values) of the LDLA – 5 cM model are represented by blue lines and the QTL positions by blue diamonds. The - log10(p-values) of the LDLA – 2 cM model are represented by red lines and the QTL positions by red crosses. The - log10(p-values) of the QTL detected by the LDLA – 1-marker model are represented by green stars. Horizontal lines correspond to the threshold values of the different models.

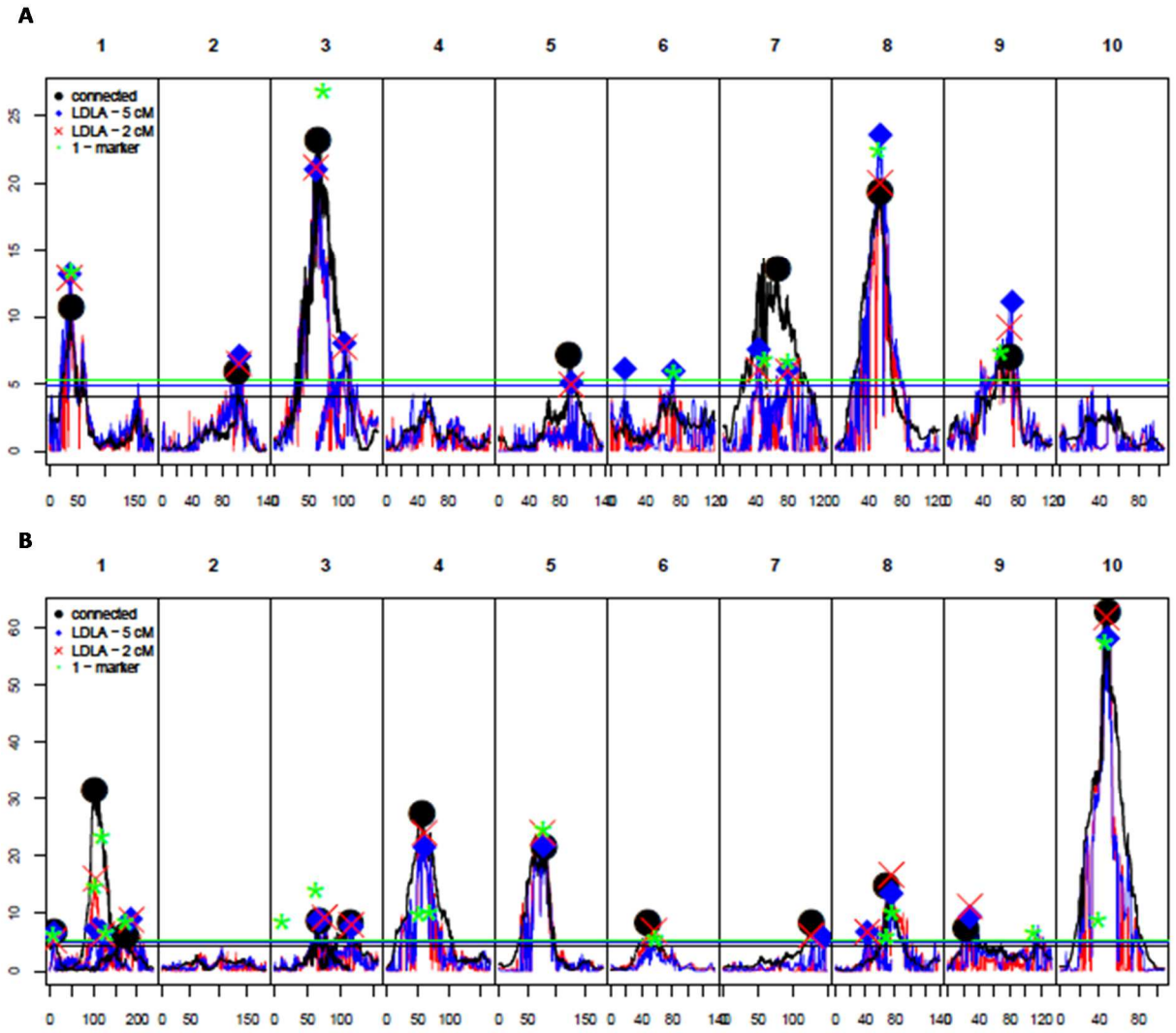

Figure S3 Results of the QTL detection with each model for DtTAS for **(A)** the dent design and **(B) the** flint design. The -log10(p-values) of the connected model are represented by black lines, the QTL positions of the connected models by black dots. The -log10(p-values) of the LDLA – 5 cM model are represented by blue lines and the QTL positions by blue diamonds. The -log10(p-values) of the LDLA – 2 cM model are represented by red lines and the QTL positions by red crosses. The -log10(p-values) of the QTL detected by the LDLA – 1-marker model are represented by green stars. Horizontal lines correspond to the threshold values of the different models.

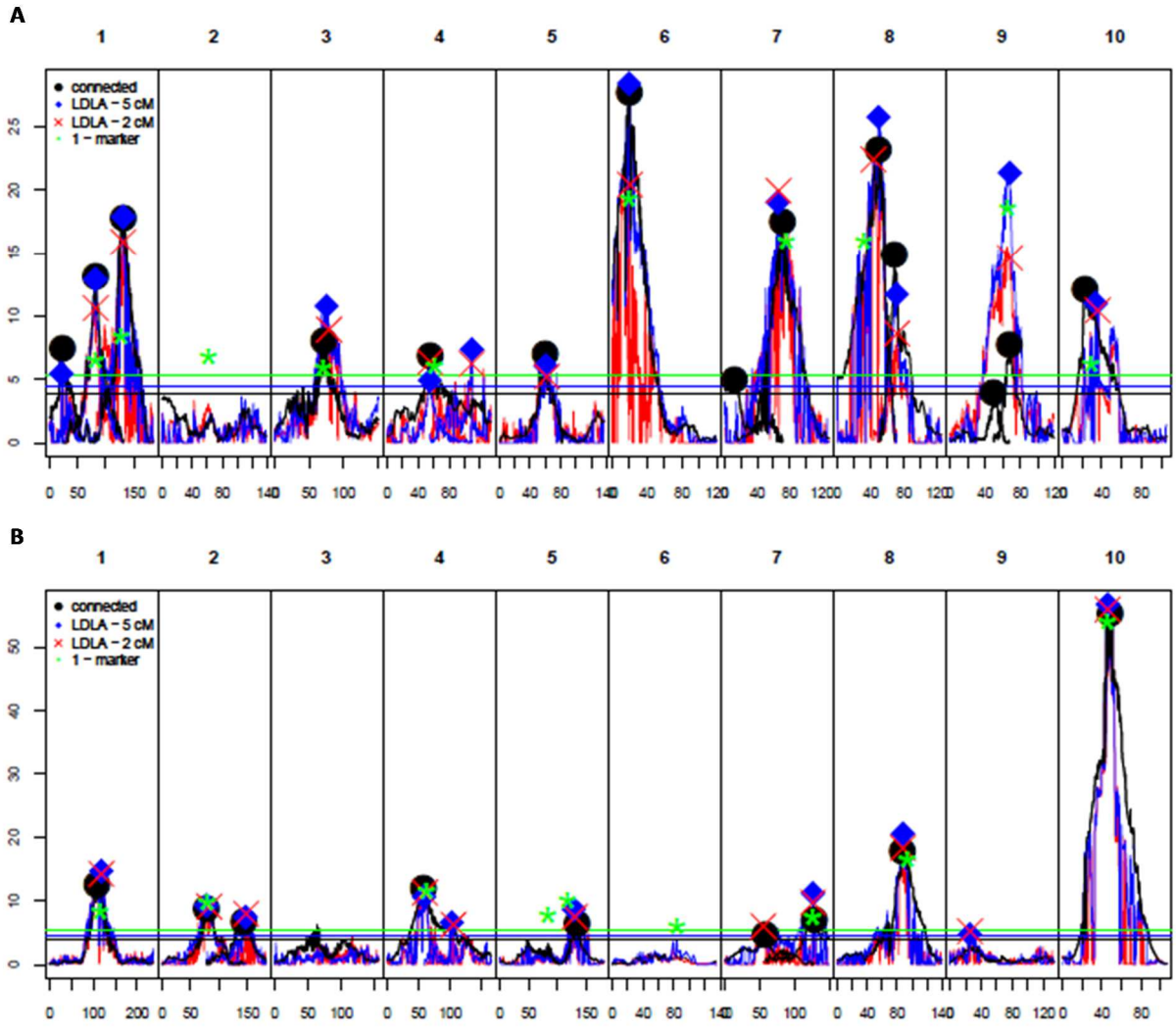

Figure S4 Results of the QTL detection with each model for PH for **(A)** the dent design and **(B) the** flint design. The -log10(p-values) of the connected model are represented by black lines, the QTL positions of the connected models by black dots. The -log10(p-values) of the LDLA – 5 cM model are represented by blue lines and the QTL positions by blue diamonds. The -log10(p-values) of the LDLA – 2 cM model are represented by red lines and the QTL positions by red crosses. The -log10(p-values) of the QTL detected by the LDLA – 1-marker model are represented by green stars. Horizontal lines correspond to the threshold values of the different models.

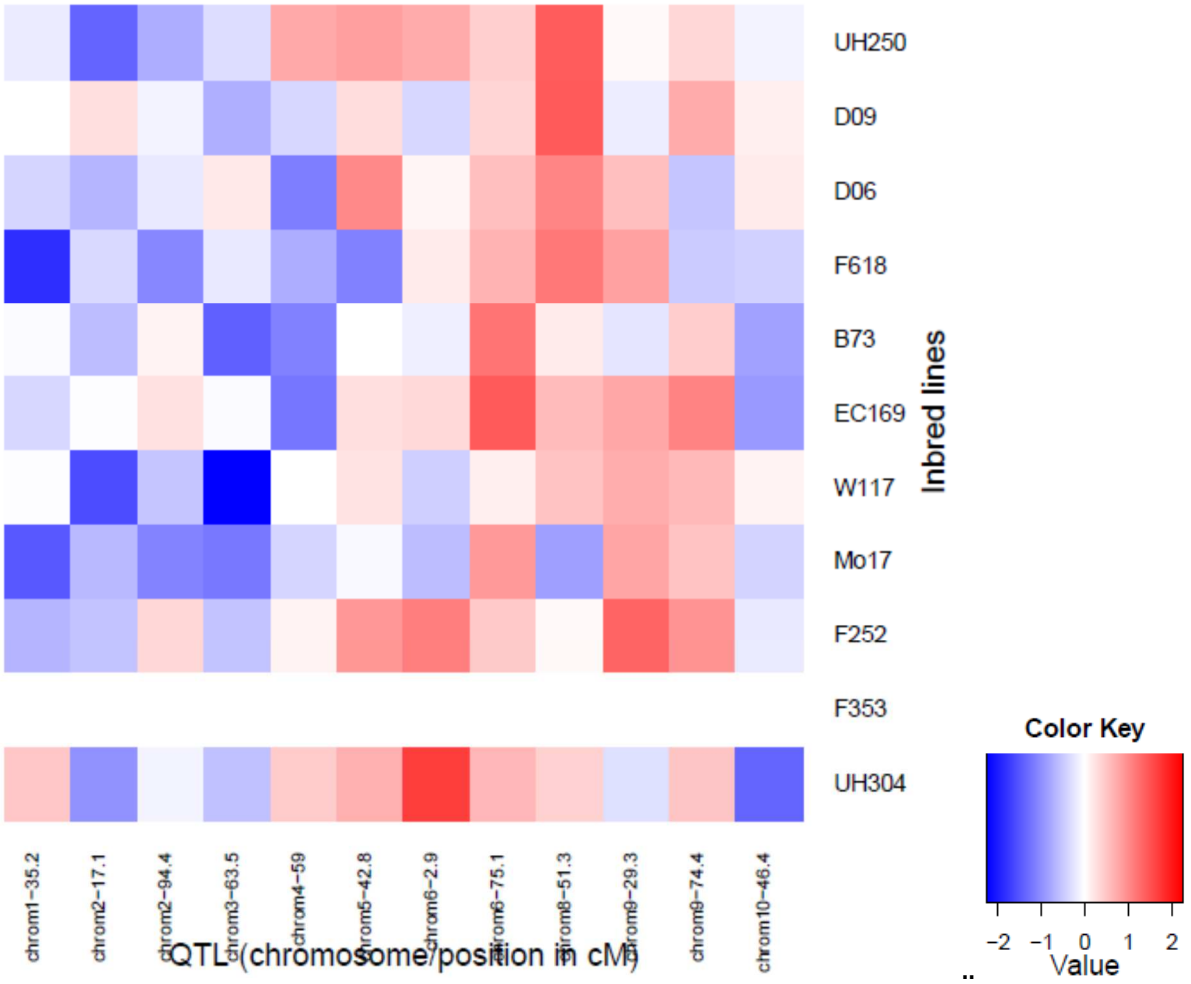

Figure S5 Allelic effects for the different dent lines for the QTL detected for DMC with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero.

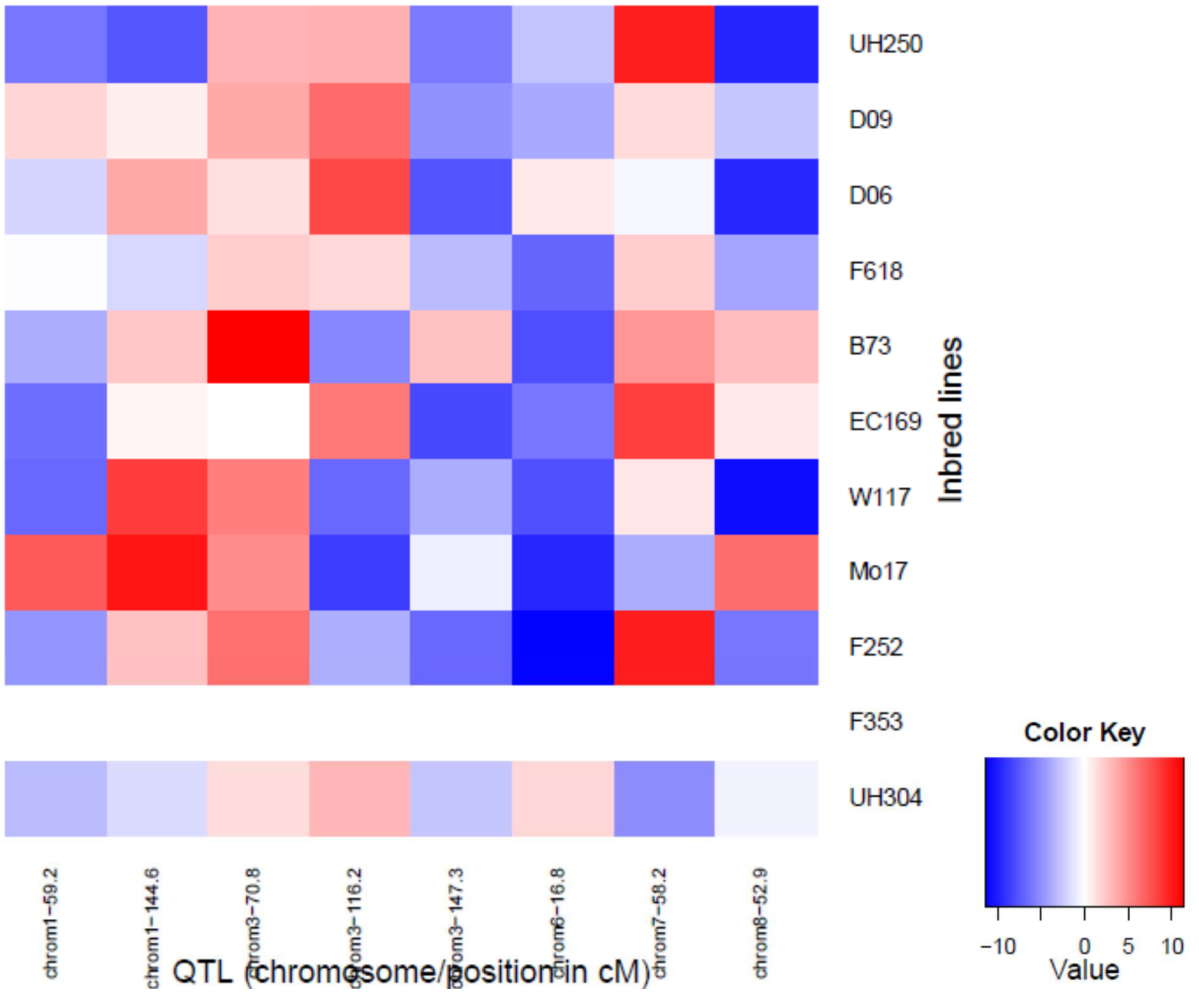

Figure S6 Allelic effects for the different dent lines for the QTL detected for DMY with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero.

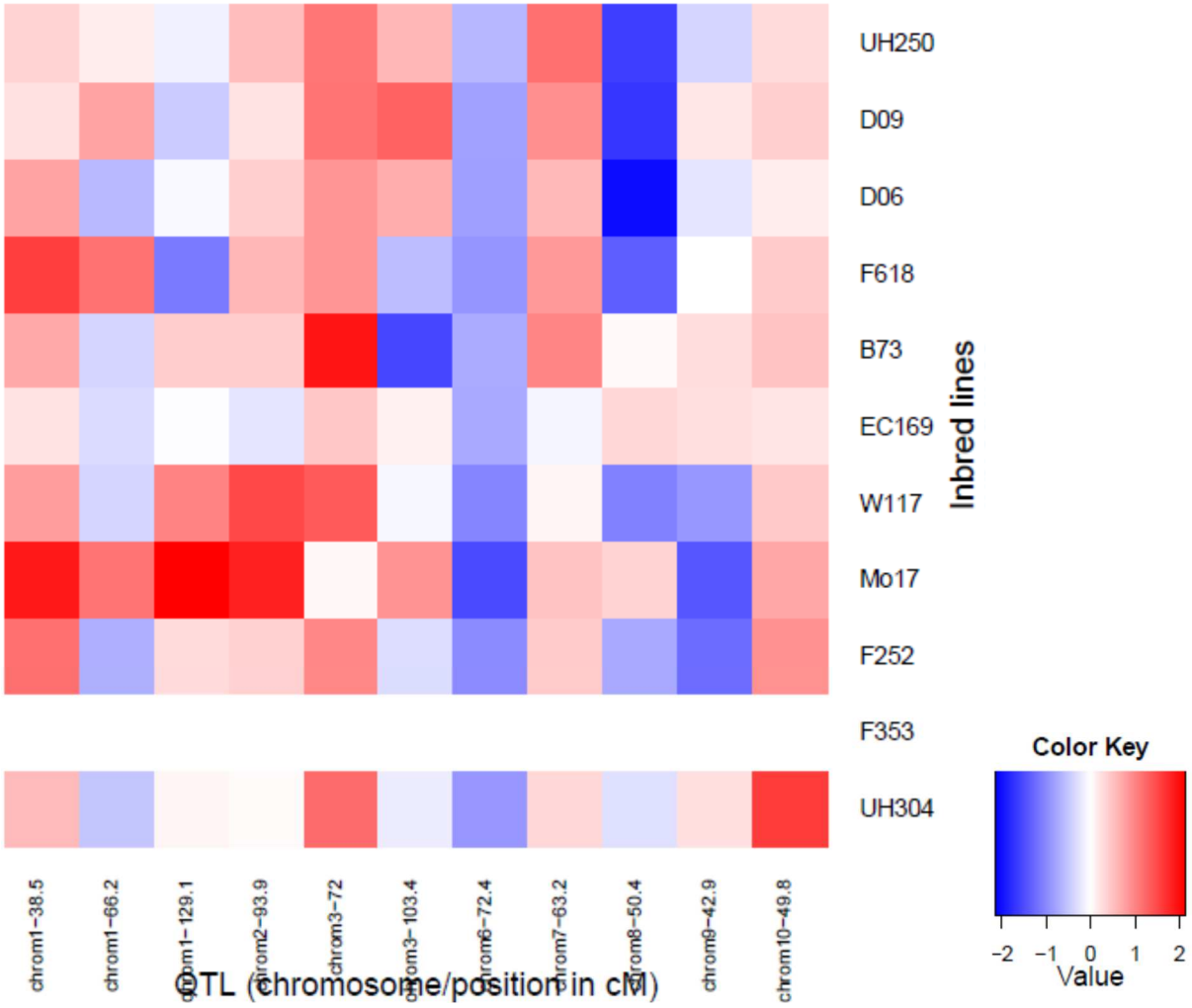

Figure S7 Allelic effects for the different dent lines for the QTL detected for DtSILK with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero.

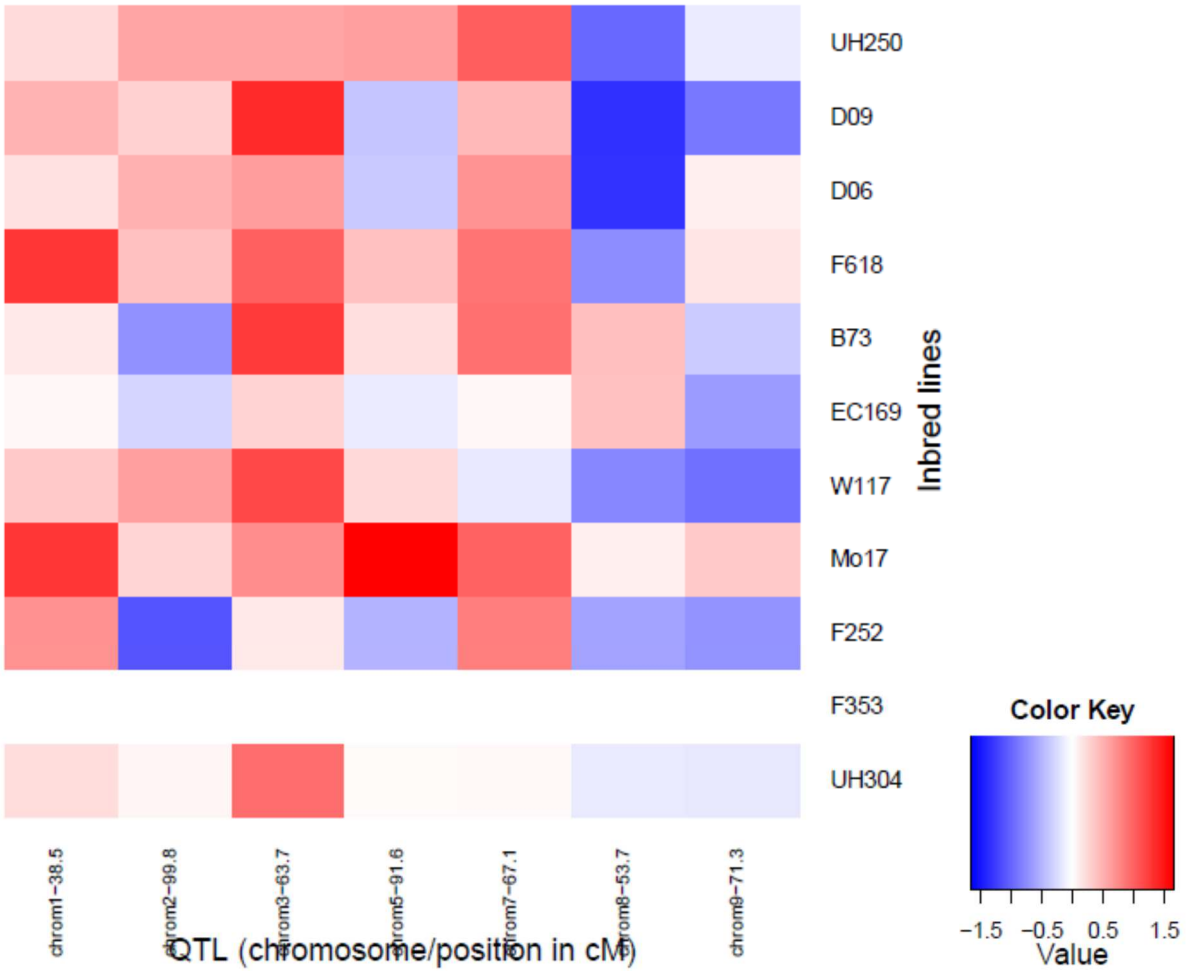

Figure S8 Allelic effects for the different dent lines for the QTL detected for DtTAS with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero.

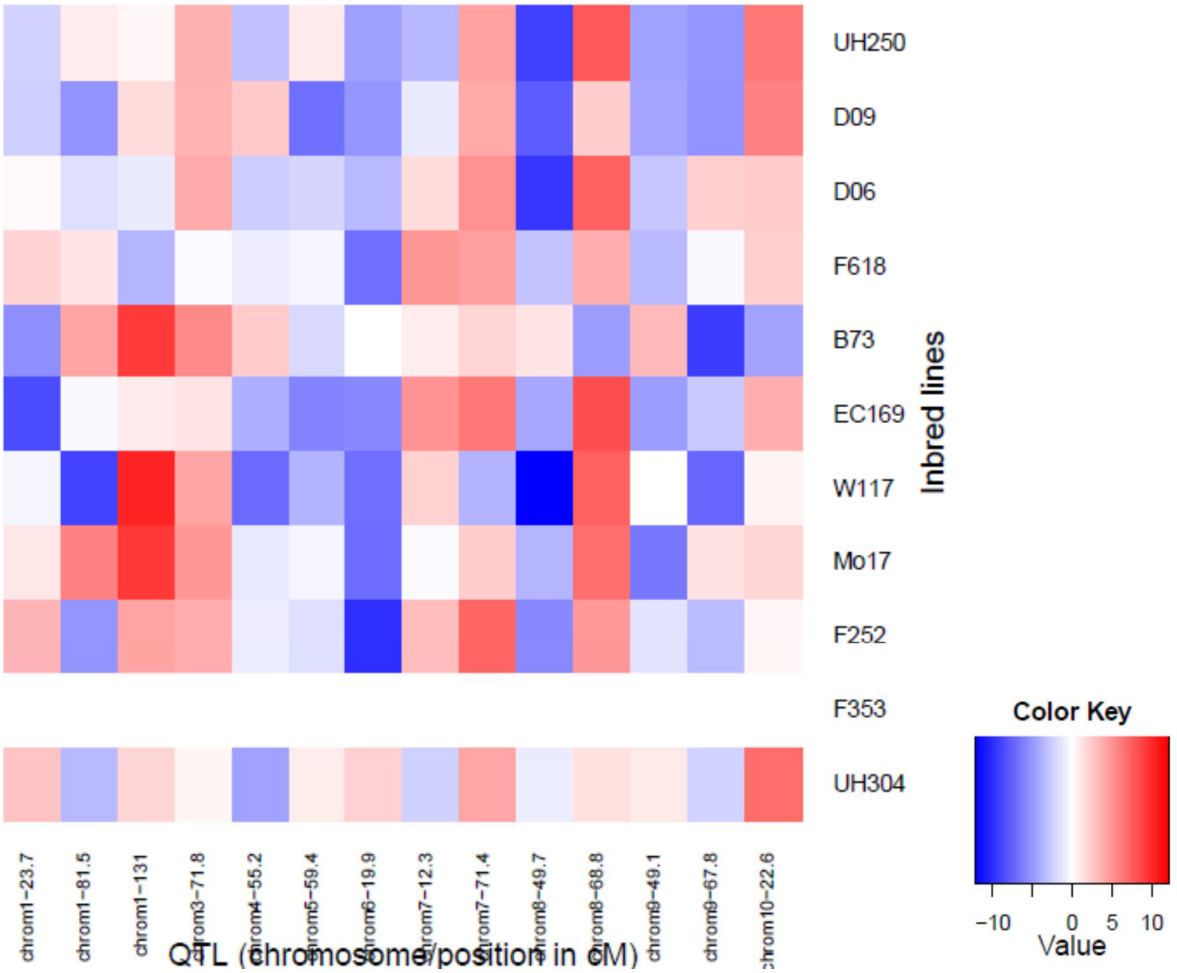

Figure S9 Allelic effects for the different dent lines for the QTL detected for PH with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero.

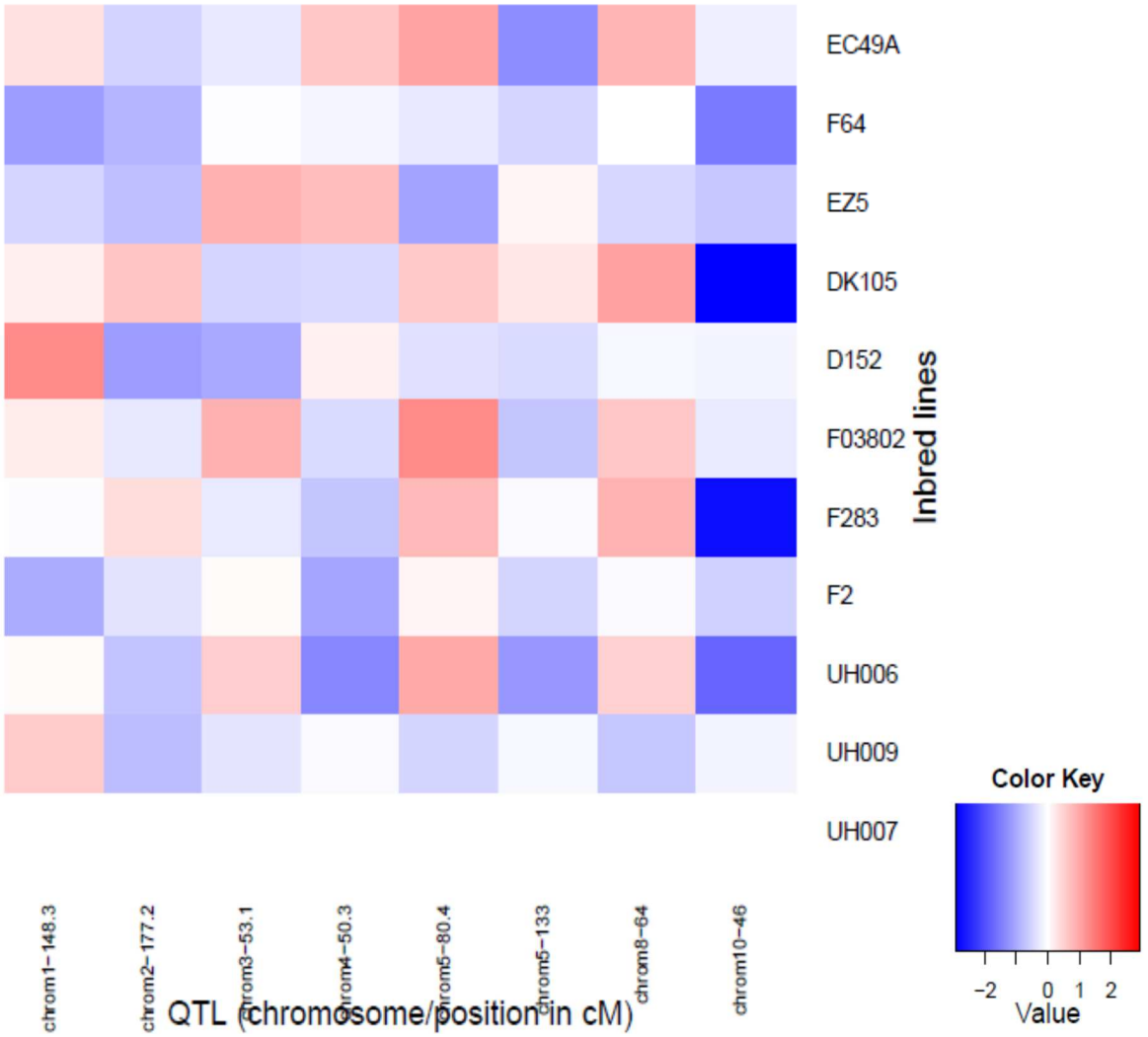

Figure S10 Allelic effects for the different flint lines for the QTL detected for DMC with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero. Allelic effects estimated for EP44 were not shown because the population where it segregates was too small (17 individuals) to obtain a reliable estimation.

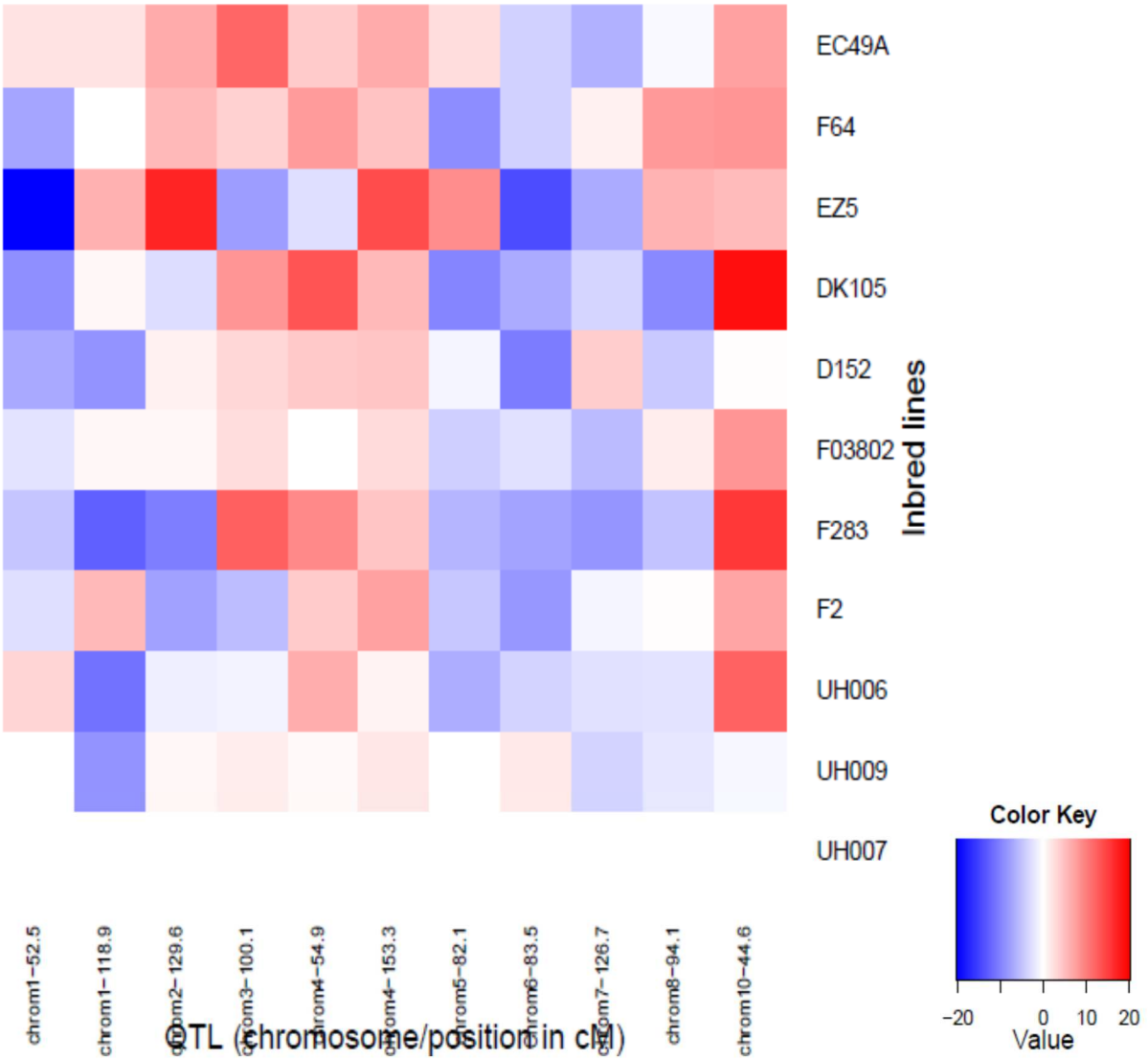

Figure S11 Allelic effects for the different flint lines for the QTL detected for DMY with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero. Allelic effects estimated for EP44 were not shown because the population where it segregates was too small (17 individuals) to obtain a reliable estimation.

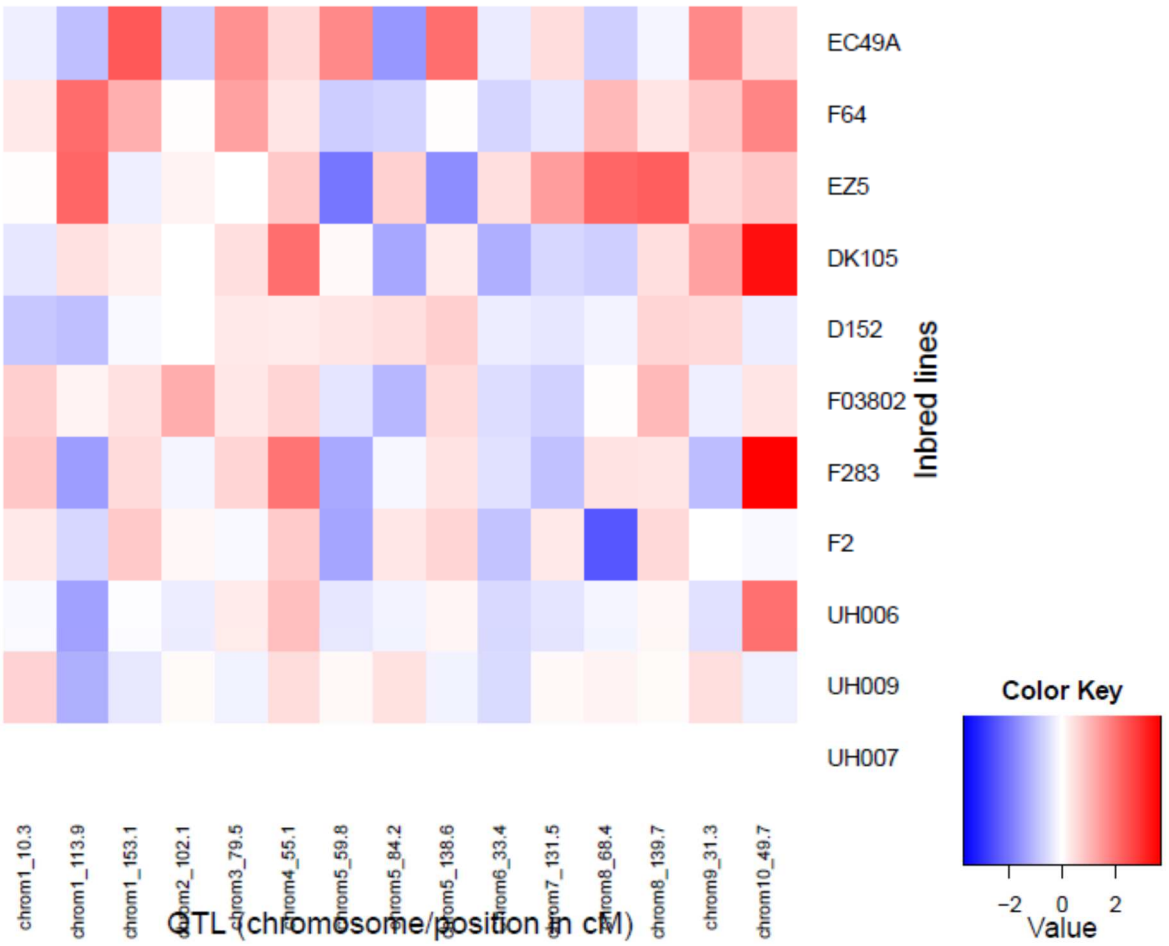

Figure S12 Allelic effects for the different flint lines for the QTL detected for DtSILK with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero. Allelic effects estimated for EP44 were not shown because the population where it segregates was too small (17 individuals) to obtain a reliable estimation.

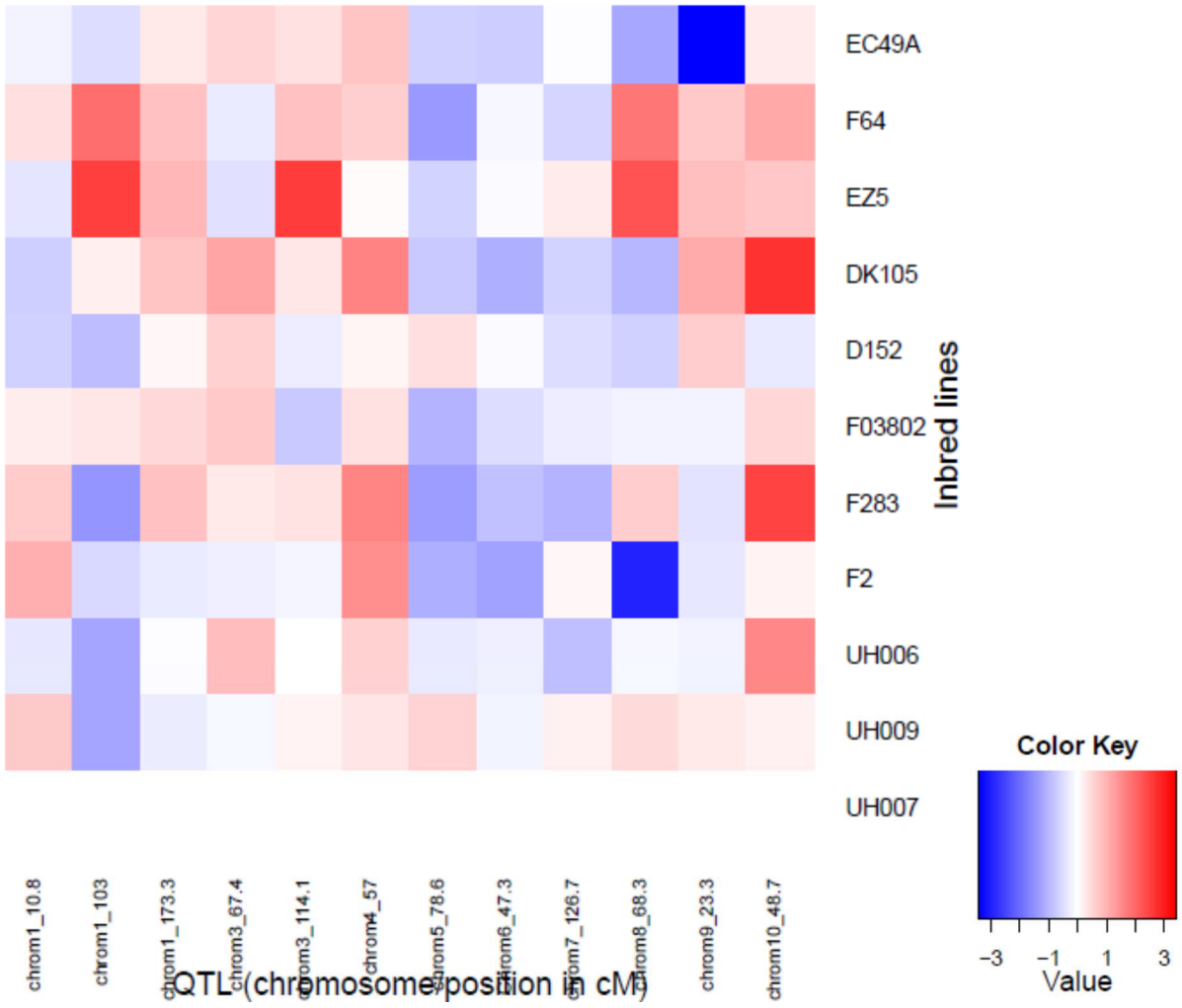

Figure S13 Allelic effects for the different flint lines for the QTL detected for DtTAS with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero. Allelic effects estimated for EP44 were not shown because the population where it segregates was too small (17 individuals) to obtain a reliable estimation.

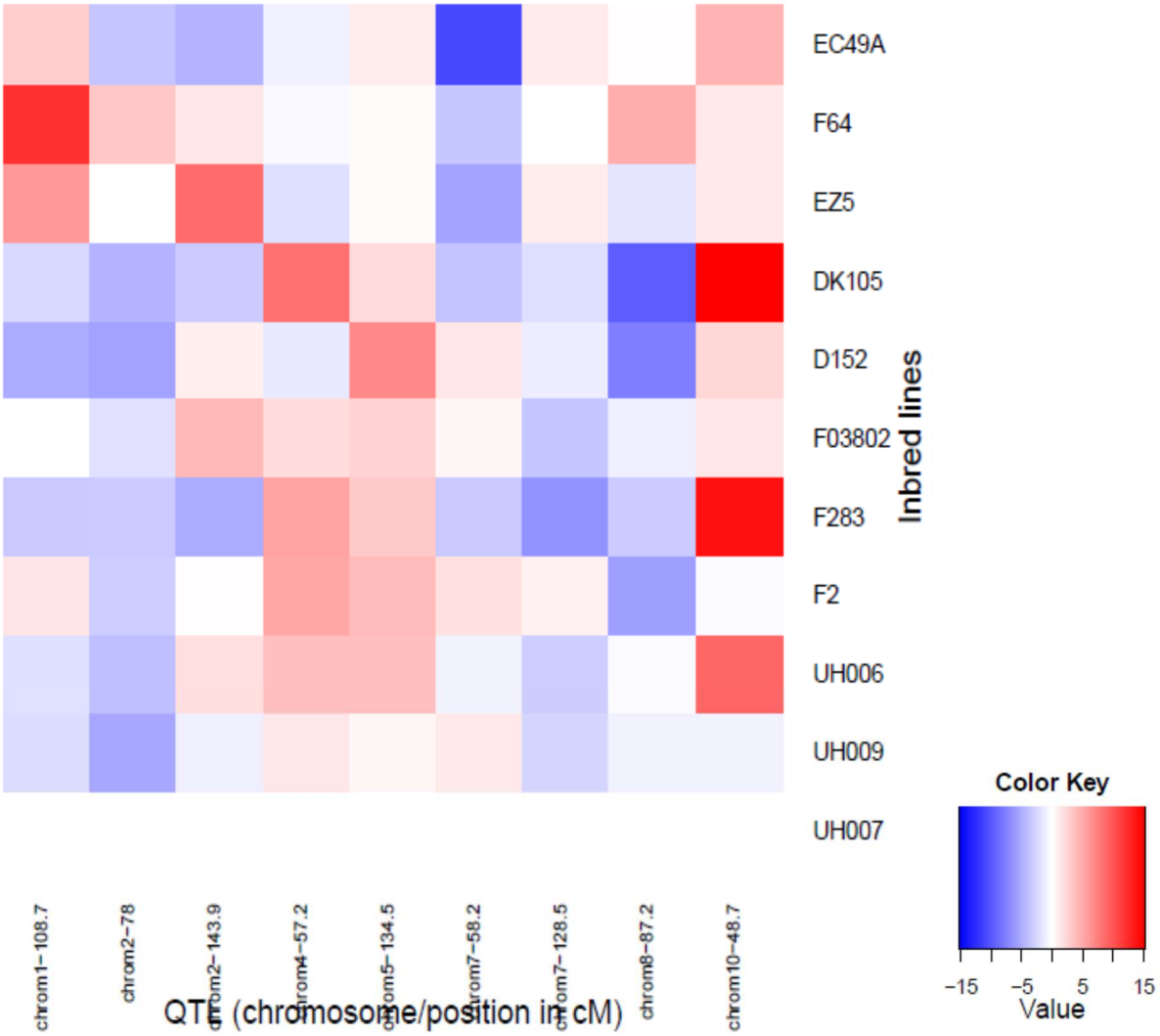

Figure S14 Allelic effects for the different flint lines for the QTL detected for PH with the connected model. Allelic effects are estimated in contrast to the central line allelic effect (F353), which was set to zero. Allelic effects estimated for EP44 were not shown because the population where it segregates was too small (17 individuals) to obtain a reliable estimation.

Table S1 Composition of the dent and flint designs with the number of DH lines in each family which were used for obtaining the consensus maps and the one which were phenotyped.

Dent design			**Flint design**		
Central line F353			Central line UH007		
Founder	Genotyped lines	Phenotyped lines	Founder	Genotyped lines	Phenotyped lines
B73	73	64	D152	112	72
D06	103	99	EC49A	53	29
D09	105	100	EP44	34[a]	17
EC169	77	66	EZ5	50	26
F252	105	96	F03802	129	129
F618	108	104	F2	77	54
Mo17	63	53	F283	134	133
UH250	99	94	F64	108	64
UH304	86	81	UH006	114	94
W117	100	84	UH009	117	98
			DK105	115	95
Total	**919**	**841**	**Total**	**1009**	**811**

[a] Population not used for mapping as too small

Table S2 Threshold values for the –log(p-value) for all the models and traits for the dent and flint groups and for the joint connected study.

	DMC	DMY	DtSILK	DtTAS	PH	**Mean**
Dent						
Connected	3.71	3.91	3.88	4.11	3.83	**3.89**
LDLA - 5cM	4.18	4.38	4.62	4.87	4.42	**4.49**
LDLA - 2cM	4.26	4.46	4.61	4.94	4.50	**4.55**
LDLA – 1-marker [a]	5.64	5.64	5.64	5.64	5.64	**5.64**
Flint						
Connected	3.43	3.65	4.15	4.37	3.83	**3.89**
LDLA - 5cM	4.07	4.20	4.84	5.13	4.53	**4.55**
LDLA - 2cM	4.30	4.36	5.04	5.23	4.60	**4.71**
LDLA – 1-marker [a]	5.61	5.61	5.61	5.61	5.61	**5.61**
Joint						
Connected	3.49	3.94	4.02	4.85	3.74	**4.00**

[a] The threshold value for the LDLA – 1-marker corresponded to a Bonferroni threshold for a genome-wide risk of 10 %. The other thresholds were calculated using 5,000 intra-family permutations of the phenotypes for a type I risk of 10 % across all families and the total genome.

Table S3 Results of the QTL detection in the flint design using the connected model. For each detected QTL, we showed its genetic position on the flint consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	Confidence interval	-log10(p)	R^2 (%)
DMC (%)	1	1	PZE_101172677	216581	148.3	142-149	5.4	3.2
	2	2	PZE_102191415	234096	177.2	176-181	6.6	3.6
	3	3	PZE_103033917	26445 - 138643	53.1	53-58	5.1	3
	4	4	PZE_104021549	18916 - 23142	50.3	49-55	6.7	3.7
	5	5	PZE_105085637	107137 - 138073	80.4	79-84	9.8	4.9
	6	5	PZE_105150268	201762	133	129-138	5.3	3.1
	7	8	PZE_108060399	107884 - 113068	64	50-65	7.4	4
	8	10	PZE_110049849	93025 - 115573	46	46-49	59.6	26.5
DMY (dt.ha-1)	1	1	PZE_101038989	25879 - 26917	52.5	52-54	7.4	3.8
	2	1	PZE_101147651	191513	118.9	117-120	14.4	6.4
	3	2	PZE_102152279	198891	129.6	127-141	8.5	4.2
	4	3	PZE_103137887	191279 - 196563	100.1	96-101	12.4	5.7
	5	4	PZE_104021283	22836	54.9	54-57	12.1	5.6
	6	4	PZE_104152609	237454 - 237702	153.3	127-154	4.5	2.7
	7	5	PZE_105078335	88752 - 139163	82.1	81-85	8.1	4.1
	8	6	PZE_106097864	151579	83.5	82-84	9.5	4.6
	9	7	PZE_107127708	170248	126.7	123-131	5.5	3.1
	10	8	PZE_108105216	159953	94.1	92-111	6	3.2
	11	10	PZE_110047350	88553 - 97551	44.6	44-45	34.2	14.2
DtSILK (d)	1	1	PZE_101005818	4452	10.3	9-12	6.9	2.2
	2	1	PZE_101143233	184616	113.9	101-115	26.3	6.7
	3	1	PZE_101181658	226175	153.1	148 -198	4.6	1.6
	4	2	PZE_102129781	178613 - 179981	102.1	101-162	5.4	1.8
	5	3	PZE_103121610	69965 - 179545	79.5	61-80	6.8	2.1
	6	4	PZE_104027603	23555 - 38027	55.1	52-57	31.8	8
	7	5	PZE_105033399	17477 - 18623	59.8	54-60	6.7	2.1
	8	5	PZE_105093579	135624 - 150088	84.2	80-86	7.3	2.2
	9	5	PZE_105163109	204822 - 205566	138.6	132-140	5.3	1.8
	10	6	PZE_106049962	99771	33.4	27-42	8	2.4
	11	7	PZE_107130719	171824	131.5	125-133	5.6	1.9

	12	8	PZE_108067425	119151	68.4	65-91	5.3	1.8
	13	8	PZE_108135936	175699 - 175734	139.7	117-140	6.1	2
	14	9	PZE_109009942	11080 - 11080	31.3	31-32	13	3.5
	15	10	PZE_110049100	91959 - 127347	49.7	49-50	94.6	27.5
DtTAS								
(d)	1	1	PZE_101005766	4609	10.8	9-13	6.8	2.5
	2	1	PZE_101108474	115220 - 161708	103	100-103	31.7	9
	3	1	PZE_101198020	246399 - 250557	173.3	164-175	6.2	2.3
	4	3	PZE_103098655	158895	67.4	67-68	9	3
	5	3	PZE_103153521	206703	114.1	110-115	8.5	2.9
	6	4	PZE_104025625	30431 - 37023	57	52-57	27.6	7.9
	7	5	PZE_105068746	71898 - 87721	78.6	78-79	21.6	6.3
	8	6	PZE_106061581	111966 - 112514	47.3	46-49	8.5	2.9
	9	7	PZE_107127708	170248	126.7	124-130	8.5	2.9
	10	8	PZE_108066752	118422 - 119082	68.3	68-69	15	4.6
	11	9	PZE_109007521	8233	23.3	23-38	7.6	2.6
	12	10	PZE_110048157	90243 - 122268	48.7	46-49	62.7	18.6
PH (cm)	1	1	PZE_101127891	162428 - 178788	108.7	108-109	12.7	5.4
	2	2	PZE_102074552	39031 - 55241	78	18-85	9	4.1
	3	2	PZE_102169535	213168	143.9	140-145	6.9	3.3
	4	4	PZE_104022475	23556 - 24765	57.2	55-60	12	5.1
	5	5	PZE_105151348	202416	134.5	132-136	6.4	3.2
	6	7	PZE_107061937	118305	58.2	53-62	4.9	2.6
	7	7	PZE_107128331	170536	128.5	121-129	7.2	3.4
	8	8	PZE_108098736	155052	87.2	87-92	17.8	7.1
	9	10	PZE_110048157	90243 - 122268	48.7	46-49	55.4	21.7

Table S4 Results of the QTL detection in the flint design using the LDLA – 5 cM model. For each detected QTL, we showed its genetic position on the flint consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	-log10(p)	R^2 (%)
DMC (%)	1	1	PZE_101147104	190602	119.4	7.3	2.9
	2	1	PZE_101250881	295590	225.9	5.6	2.2
	3	2	PZE_102025627	11947	39.2	7	1.8
	4	2	PZE_102046822	24366	63.7	7.3	3.4
	5	2	PZE_102183284	225854	157.9	8	3
	6	4	PZE_104021283	22836	54.9	11.2	4.8
	7	5	PZE_105085637	107137 - 138073	80.4	10.2	4.3
	8	5	PZE_105165365	208891 - 209048	148.7	5.2	1.6
	9	6	PZE_106076029	131411 - 134098	64.2	4.9	2.3
	10	8	PZE_108018911	18447	43.1	7.5	2.9
	11	10	PZE_110049849	93025 - 115573	46	70.2	29.2
DMY ($dt.ha^{-1}$)	1	1	PZE_101147651	191513	118.9	13.3	5.7
	2	1	PZE_101213494	263732	185.2	5.7	2.6
	3	3	PZE_103098382	158668 - 159808	66.8	7.9	4.1
	4	4	PZE_104020618	21905	55	13.5	6.3
	5	4	PZE_104123129	200190	129.8	5.3	2.9
	6	5	PZE_105068572	71700 - 72614	75.9	6.9	3.7
	7	6	PZE_106107736	156986	95.6	4.3	1.2
	8	7	PZE_107128866	170819	127.8	10	3.5
	9	8	PZE_108029326	27221 - 66473	50.1	6.6	3.4
	10	10	PZE_110045930	86778 - 109582	46.3	31.3	14.3
DtSILK (d)	1	1	PZE_101005818	4452	10.3	5.2	1.2
	2	1	PZE_101146834	190143	119.1	34.3	8.1
	3	1	PZE_101199192	248322	171.8	15	3.9
	4	2	PZE_102179704	222468	154.6	10.1	2.4
	5	3	PZE_103098382	158668 - 159808	66.8	8.2	2.4
	6	3	PZE_103121610	69965 - 179545	79.5	7.8	2
	7	4	PZE_104062511	44504 - 124929	62.6	31.7	8.1
	8	5	PZE_105078445	86146 - 140781	80.7	18.6	4.7
	9	5	PZE_105153835	204326 - 205504	137.6	10.9	2.4
	10	7	PZE_107133704	173181	139.3	7.6	1.6
	11	8	PZE_108066557	118189	68.7	7.4	1.9
	12	8	PZE_108133033	173617	133.8	6.6	1.1

	13	9	PZE_109009936	11079	31.6	11.6	2.8
	14	10	PZE_110048157	90243 - 122268	48.7	98.7	29.1
DtTAS (d)	1	1	PZE_101005765	4609	10	6.5	1.8
	1	1	PZE_101109004	116312 - 158005	105.3	7.3	2.5
	2	1	PZE_101147104	190602	119.4	6.9	2.2
	3	1	PZE_101213102	263154	185.3	9	2.7
	4	3	PZE_103098655	158895	67.4	8.9	2.7
	5	3	PZE_103158635	210426	116.9	7.7	2.2
	6	4	PZE_104044703	33362 - 96313	60.7	21.6	6.8
	7	5	PZE_105066936	69125 - 83278	77.8	21.9	6.5
	8	7	PZE_107136925	174718	144.3	5.7	1.3
	9	8	PZE_108019174	18351	42.8	6.8	2
	10	8	PZE_108073574	128549 - 128753	75.2	13.4	3.9
	11	9	PZE_109009220	10008 - 10009	28.4	9.5	2.4
	12	10	PZE_110048157	90243 - 122268	48.7	58.1	18.3
	13	1	PZE_101005765	4609	10	6.5	1.8
PH (cm)	1	1	PZE_101146427	189406	118.5	14.7	5.2
	2	2	PZE_102074552	39031 - 55241	78	9.3	3.8
	3	2	PZE_102173058	216192	146.8	7.4	2.9
	4	4	PZE_104045760	68246 - 68323	58.7	10.6	4.4
	5	4	PZE_104103602	179801 - 180054	102.8	6.6	2.4
	6	5	PZE_105150122	201632	128.7	8.7	2.7
	7	7	PZE_107128144	170420 - 170496	126.6	11.2	3.2
	8	8	PZE_108092331	149305 - 155644	87.8	20.5	7.8
	9	9	PZE_109008703	9311	25.5	4.8	1.3
	10	10	PZE_110049849	93025 - 115573	46	56.8	21.7

Table S5 Results of the QTL detection in the flint design using the LDLA – 2 cM model. For each detected QTL, we showed its genetic position on the flint consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	-log10(p)	R^2 (%)
DMC (%)	1	1	PZE_101146598	189773	118	6.2	2.5
	2	2	PZE_102185359	229130 - 229288	165	5	1.7
	3	4	PZE_104018885	18916 - 23142	50.3	6.9	3.4
	4	5	PZE_105085637	107137 - 138073	80.4	9.1	4.6
	5	5	PZE_105163718	208374	149.2	4.5	1.9
	6	8	PZE_108063241	113068 - 113206	64.1	6.5	3.2
	7	10	PZE_110049849	93025 - 115573	46	46.5	20.8
	8	10	PZE_110089009	139036	69.9	5	1.4
DMY ($dt.ha^{-1}$)	1	1	PZE_101146427	189406	118.5	14.9	5.3
	2	2	PZE_102172077	215135	144.9	6.9	2.5
	3	3	PZE_103097999	157939	65.4	6	1.9
	4	3	PZE_103142979	198520 - 198581	101.6	4.7	2.5
	5	4	PZE_104023433	26403 - 26403	52.7	15.8	5.8
	6	4	PZE_104122410	199546 - 199546	126.5	5.4	2.6
	7	5	PZE_105092759	133339 - 159961	84.4	7.4	3.7
	8	6	PZE_106050624	100745 - 103709	35.4	4.6	2.2
	9	6	PZE_106103665	155178	90.5	7.8	3.4
	10	7	PZE_107128846	170819	128.9	9	2.9
	11	8	PZE_108027746	26074 - 29164	49.5	6.7	3.4
	12	10	PZE_110047350	88553 - 97551	44.6	31.7	13.4
DtSILK (d)	1	1	PZE_101005818	4452	10.3	5.6	1.4
	2	1	PZE_101147104	190602	119.4	33.6	9.2
	3	1	PZE_101199859	248854 - 249092	173.9	14.2	4.1
	4	2	PZE_102181292	222435 - 223721	156.5	6.3	1.6
	5	3	PZE_103118006	176570	78.5	10.9	3
	6	3	PZE_103167997	216529	126.6	9.2	1.9
	7	4	PZE_104044892	42641 - 134020	62.5	29.6	8.8
	8	5	PZE_105039522	24542	63.6	18	5.4
	9	8	PZE_108133033	173617	133.8	9.2	2.1
	10	9	PZE_109010021	11134	30.1	11.9	2.9
	11	10	PZE_110060375	114622 - 114653	48.6	89.7	29.6
DtTAS(d)	1	1	PZE_101005770	4610	10.5	5.3	1.5
	2	1	PZE_101109004	116312 - 158005	105.3	16	4.7

	3	1	PZE_101147248	190703	121.1	6.4	1.7
	4	1	PZE_101213479	263702 - 265655	186.8	8.9	2.3
	5	3	PZE_103109418	170117 - 171781	75.9	9.3	2.1
	6	3	PZE_103157683	209726	116.3	8	2.1
	7	4	PZE_104044703	33362 - 96313	60.7	24	7.2
	8	5	PZE_105063310	62822 - 82069	76	24	7.1
	9	6	PZE_106064975	117082 - 122646	56.3	6.9	2.2
	10	7	PZE_107128144	170420 - 170496	126.6	6.1	1.4
	11	8	PZE_108019174	18351	42.8	6.8	1.9
	12	8	PZE_108073574	128549 - 128753	75.2	16.7	4.5
	13	9	PZE_109009220	10008 - 10009	28.4	11.2	2.5
	14	10	PZE_110049001	89438 - 108230	47.2	61.8	18.8
PH (cm)	1	1	PZE_101144184	187342 - 187381	118.6	14.2	4.6
	2	2	PZE_102076936	51554 - 59013	83	9.1	3.7
	3	2	PZE_102175167	217650	147.7	8	2.7
	4	4	PZE_104028514	34558 - 80248	60.2	11.3	4.8
	5	4	PZE_104104676	180887	105.6	6.2	1.5
	6	5	PZE_105144284	198198	130.3	7.5	1.9
	7	7	PZE_107057864	111123 - 112763	55.4	6	2.2
	8	7	PZE_107128144	170420 - 170496	126.6	9.7	2.8
	9	8	PZE_108092331	149305 - 155644	87.8	18.3	6.6
	10	9	PZE_109008133	8741	25.8	5.2	1.5
	11	10	PZE_110049849	93025 - 115573	46	56	21.3

Table S6 Results of the QTL detection in the flint design using the LDLA – 1-marker model. For each detected QTL, we showed its genetic position on the flint consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	-log10(p)	R^2 (%)
DMC (%)	1	2	PZE_102185353	229130	164.2	5.7	1.7
	2	4	PZE_104033064	40693	60	6.7	2.1
	3	5	PZE_105079359	90584	80.6	11.5	3.9
	4	5	PZE_105143697	197846	126.4	5.7	1.7
	5	8	PZE_108063319	113212	63.6	7.2	2.3
	6	10	PZE_110050010	94199	45.9	44.3	18.0
	7	10	PZE_110086343	137505	68.9	6.3	1.9
DMY ($dt.ha^{-1}$)	1	1	PZE_101128881	164375	105.8	6.3	1.8
	2	1	PZE_101144216	187381	118.6	15.6	5.2
	3	3	PZE_103097999	157939	65.4	10.3	3.3
	4	4	PZE_104017088	17150	48.7	11.4	3.7
	5	4	PZE_104021665	23190	51.4	5.8	1.7
	6	4	PZE_104122007	198999	124	6.0	1.7
	7	5	PZE_105094114	137392	81.5	9.7	3.1
	8	6	PZE_106104239	155466	90.7	7.8	2.4
	9	7	PZE_107128846	170819	128.9	8.2	2.5
	10	8	PZE_108028156	29898	51.1	9.1	2.9
	11	10	PZE_110050010	94199	45.9	33.0	12.2
DtSILK (d)	1	1	PZE_101004387	3883	8.5	8.2	1.5
	2	1	PZE_101088198	79735	90.5	10.0	1.9
	3	1	PZE_101106156	109635	102.4	17.1	3.6
	4	1	PZE_101151084	194731	125.8	9.4	1.8
	5	1	PZE_101200614	249700	173.1	11.3	2.2
	6	3	PZE_103098779	158974	61.9	11.3	2.2
	7	4	PZE_104021514	23073	51	9.0	1.7
	8	4	PZE_104079162	153502	69	10.8	2.1
	9	4	PZE_104152590	237693	155.7	6.4	1.1
	10	5	PZE_105069912	74335	76.8	21.1	4.6
	11	5	PZE_105143119	197706	127.6	5.7	1.0
	12	7	PZE_107128331	170536	128.5	6.5	1.2
	13	8	PZE_108070056	122950	71	12.0	2.4
	14	9	PZE_109009591	10597	30.3	7.1	1.3
	15	10	PZE_110016138	16504	38.6	11.2	2.2

	16	10	PZE_110050010	94199	45.9	45.7	11.1
DtTAS(d)	1	1	PZE_101004387	3883	8.5	6.1	1.2
	2	1	PZE_101115961	138907	103.5	14.5	3.3
	3	1	PZE_101144216	187381	118.6	23.2	5.6
	4	1	PZE_101160171	202307	128	6.7	1.3
	5	1	PZE_101200614	249700	173.1	8.5	1.8
	6	3	PZE_103007349	4064	11.7	8.6	1.8
	7	3	PZE_103098779	158974	61.9	14.0	3.2
	8	4	PZE_104021514	23073	51	9.7	2.1
	9	4	PZE_104079162	153502	69	10.0	2.1
	10	5	PZE_105069912	74335	76.8	24.7	6.0
	11	6	PZE_106066817	119166	57.4	5.5	1.0
	12	8	PZE_108067255	118970	68.8	5.8	1.1
	13	8	PZE_108074213	129415	75.5	10.3	2.2
	14	9	PZE_109111133	151251	109.1	6.5	1.3
	15	10	PZE_110018448	22128	38.9	8.9	1.9
	16	10	PZE_110050010	94199	45.9	57.5	16.1
PH (cm)	1	1	PZE_101145493	188172	115.6	8.3	2.4
	2	2	PZE_102074558	55249	79.5	9.8	3.0
	3	4	PZE_104042538	60023	63.1	11.6	3.6
	4	5	PZE_105091638	129996	83.4	7.8	2.3
	5	5	PZE_105134752	195420	117.6	9.9	3.0
	6	6	PZE_106097959	151785	84.1	6.1	1.7
	7	7	PZE_107127637	170111	126.3	7.4	2.1
	8	8	PZE_108105216	159953	94.1	16.6	5.3
	9	10	PZE_110050010	94199	45.9	53.9	20.6

Table S7 Results of the QTL detection in the dent design using the connected model. For each detected QTL, we showed its genetic position on the dent consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	Confidence interval	-log10(p)	R^2 (%)
DMC (%)	1	1	PZE_101031077	19101	35.2	32-39	11.7	4.8
	2	2	PZE_102011868	5425	17.1	16-18	13.4	5.3
	3	2	PZE_102149235	195177 - 197936	94.4	88-100	5.4	2.6
	4	3	PZE_103091082	150173 - 165855	63.5	58-64	16.8	6.5
	5	4	PZE_104079076	153406	59	57-61	10	4.2
	6	5	PZE_105026024	13303 - 13313	42.8	42-45	10.6	4.4
	7	6	PZE_106002839	3588 - 3869	2.9	1-6	10.4	4.3
	8	6	PZE_106098045	151822	75.1	72-79	9.6	4.1
	9	8	PZE_108058161	103705 - 103897	51.3	51-58	18	6.9
	10	9	PZE_109009836	10943	29.3	27-30	6	2.8
	11	9	PZE_109096235	141951	74.4	70-76	6.8	3.1
	12	10	PZE_110048796	91481 - 107902	46.4	27-53	5.7	2.7
DMY (dt.ha-1)	1	1	PZE_101071870	54342	59.2	57-62	5.7	3.8
	2	1	PZE_101215677	266310 - 266369	144.6	123-160	5.5	3.7
	3	3	PZE_103108908	169730 - 172477	70.8	66-76	5.5	3.7
	4	3	PZE_103160673	211719 - 212707	116.2	115-129	6.2	4
	5	3	PZE_103185177	229665	147.3	146-148	6.3	4.1
	6	6	PZE_106038467	86549	16.8	9-20	14.9	8
	7	7	PZE_107066645	123598 - 126465	58.2	57-61	11.7	6.6
	8	8	PZE_108057442	102536 - 108663	52.9	52-53	14.2	7.7
DtSILK (d)	1	1	PZE_101033622	21685	38.5	32-39	11.6	4.7
	2	1	PZE_101081841	69289 - 70518	66.2	65-67	4.6	2.3
	3	1	PZE_101194503	241368 - 244469	129.1	128-133	6.6	3
	4	2	PZE_102148927	195747 - 196529	93.9	93-96	9.1	3.8
	5	3	PZE_103110415	170772 - 174828	72	65-72	15.9	6.1
	6	3	PZE_103147207	201536 - 202769	103.4	101-110	5.5	2.6
	7	6	PZE_106095147	150309	72.4	70-74	13	5.1
	8	7	PZE_107072681	129265	63.2	43-67	8.4	3.6
	9	8	PZE_108057885	103311	50.4	50-54	26.4	9.6
	10	9	PZE_109020361	18684 - 20598	42.9	38-45	5	2.5
	11	10	PZE_110057591	110540 - 120784	49.8	49-52	7.2	3.2

DtTAS (d)	1	1	PZE_101033622	21685	38.5	36-39	10.8	5.4
	2	2	PZE_102157405	204235	99.8	99-100	6	3.4
	3	3	PZE_103101981	162179 - 167076	63.7	61-65	23.2	10.4
	4	5	PZE_105144068	198031	91.6	91-96	7.3	3.9
	5	7	PZE_107076807	132075	67.1	48-68	13.7	6.5
	6	8	PZE_108058411	104281 - 104625	53.7	50-55	19.4	8.8
	7	9	PZE_109092637	139196 - 140154	71.3	56-74	7.1	3.9
PH (cm)	1	1	PZE_101018818	10905	23.7	21-24	7.5	3
	2	1	PZE_101133561	172881 - 172940	81.5	80-84	13.2	4.7
	3	1	PZE_101196829	245032 - 245219	131	127-133	17.9	6.1
	4	3	PZE_103110278	170548	71.8	69-73	8.1	3.2
	5	4	PZE_104073340	138154 - 144727	55.2	53-58	6.9	2.8
	6	5	PZE_105065019	66038 - 79496	59.4	58-60	7.1	2.9
	7	6	PZE_106040994	89408 - 91643	19.9	19-20	27.7	9.1
	8	7	PZE_107005418	3665 - 3667	12.3	2-16	5.1	2.2
	9	7	PZE_107080996	135892	71.4	71-72	17.6	6
	10	8	PZE_108056028	100939 - 102711	49.7	49-52	23.3	7.7
	11	8	PZE_108078317	130737 - 134065	68.8	68-69	14.9	5.2
	12	9	PZE_109025803	25986	49.1	48-50	4	1.9
	13	9	PZE_109086708	134570 - 135460	67.8	67-70	7.8	3.1
	14	10	PZE_110008028	6072	22.6	22-26	12.2	4.4

Table S8 Results of the QTL detection in the dent design using the LDLA – 5 cM model. For each detected QTL, we showed its genetic position on the dent consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	-log10(p)	R^2 (%)
DMC (%)	1	1	PZE_101028121	16789 - 17963	31.6	11	3.9
	2	1	PZE_101150204	193868 - 194764	92	4.9	2
	3	1	PZE_101202934	249700 - 251159	134.9	7.9	3.1
	4	2	PZE_102006385	3379	9.6	6.4	1.5
	5	2	PZE_102150016	196649	94	8.9	2.9
	6	3	PZE_103038564	33572 - 56014	46	17.6	6.4
	7	3	PZE_103151042	204999	105	5.4	2.3
	8	4	PZE_104081311	155805	59.6	10.2	4
	9	5	PZE_105047074	35783 - 36699	52.3	12.4	4.9
	10	6	PZE_106007445	18846 - 21466	9.9	11.7	4.3
	11	6	PZE_106096901	150891	71.7	12.7	4.6
	12	7	PZE_107040665	154074	66.8	3.6	1.7
	13	8	PZE_108057885	103311	50.4	19.1	6.1
	14	9	PZE_109089324	137410	68.6	9	3.5
	15	10	PZE_110012467	10879	31.8	7.7	2.2
DMY ($dt.ha^{-1}$)	1	3	PZE_103116584	175989	78.9	8	3.9
	2	3	PZE_103162977	213416	117.8	7.8	4.3
	3	6	PZE_106038467	86549	16.8	13.8	8.1
	4	7	PZE_107066645	123598 - 126465	58.2	8.8	5.6
	5	8	PZE_108057442	102536 - 108663	52.9	10.3	5.8
DtSILK (d)	1	1	PZE_101033622	21685	38.5	22.6	7.2
	2	1	PZE_101205734	251079 - 254464	136.4	6.7	2.4
	3	2	PZE_102152020	198672	94.5	15	4.5
	4	3	PZE_103086165	142732 - 157202	61.7	10	3.6
	5	3	PZE_103122617	180515	78.6	16.3	4.7
	6	5	PZE_105049624	41635 - 58706	56.3	4.8	2.2
	7	6	PZE_106094705	149930	71	16.5	5.4
	8	7	PZE_107045046	25471 - 104886	43.9	7.4	3
	9	7	PZE_107107125	158951 - 158952	82.1	5.8	2
	10	8	PZE_108058411	104281 - 104625	53.7	27.7	8.9
	11	9	PZE_109098496	143352	77.3	5.5	2
	12	10	PZE_110057591	110540 - 120784	49.8	8	3.1
DtTAS(d)	1	1	PZE_101032015	19641 - 21075	35.7	13.2	4.8

	2	2	PZE_102159907	206081 - 207151	102.9	7.1	2.8
	3	3	PZE_103098157	158352	60.8	21	7.6
	4	3	PZE_103143600	199245 - 201331	102.2	8	3.1
	5	5	PZE_105143985	197957 - 200116	95.6	5.1	2
	6	6	PZE_106033981	79499 - 86347	16	6.1	2.7
	7	6	PZE_106092387	148530 - 150461	72	6	2.5
	8	7	PZE_107045046	25471 - 104886	43.9	7.6	3.4
	9	7	PZE_107099124	152685 - 155704	79.9	6.1	2.7
	10	8	PZE_108062375	111291	54.3	23.6	8.3
	11	9	PZE_109094832	141175	73.4	11.1	3.9
PH (cm)	1	1	PZE_101018868	10962	21.4	5.5	2.1
	2	1	PZE_101133356	172811	81.4	12.9	4.4
	3	1	PZE_101196829	245032 - 245219	131	17.9	6.1
	4	3	PZE_103111112	171438 - 175550	75.5	10.8	3.9
	5	4	PZE_104073340	138154 - 144727	55.2	4.9	2
	6	4	PZE_104136077	202589 - 227111	109.7	7.4	2.6
	7	5	PZE_105068432	70082 - 86033	60.4	6.1	2.7
	8	6	PZE_106040975	89403 - 89404	19.3	28.4	9.3
	9	7	PZE_107076796	132076	65.4	19	6.3
	10	8	PZE_108056028	100939 - 102711	49.7	25.7	8.2
	11	8	PZE_108079422	133563 - 138524	71.4	11.7	4.1
	12	9	PZE_109085253	133933	68.1	21.3	7.2
	13	10	PZE_110014332	11179 - 13553	33.8	11	3.3

Table S9 Results of the QTL detection in the dent design using the LDLA – 2 cM model. For each detected QTL, we showed its genetic position on the dent consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	-log10(p)	R^2 (%)
DMC (%)	1	1	PZE_101036345	23712	38	12.4	4
	2	1	PZE_101154088	194939 - 197272	93.9	6.2	2.3
	3	1	PZE_101203104	250888	131.9	7.1	2.8
	4	2	PZE_102002360	1724	4.7	8.5	2
	5	2	PZE_102017964	8279	23.4	7.7	2.2
	6	2	PZE_102152020	198672	94.5	8.5	3
	7	3	PZE_103093079	154090 - 160936	63.9	16.7	5.4
	8	3	PZE_103148259	202185	104.1	5.4	2
	9	4	PZE_104076988	151510 - 151684	56.9	10.3	3.5
	10	5	PZE_105047074	35783 - 36699	52.3	13.3	4.9
	11	6	PZE_106020123	14400 - 24611	10	14.3	4.4
	12	6	PZE_106097959	151785	73.1	14.1	4.3
	13	7	PZE_107045895	24563 - 103626	43.6	5.8	2.4
	14	8	PZE_108061901	110744 - 115294	57.2	19.7	6.4
	15	9	PZE_109091148	138616 - 138617	69.6	12.7	4.2
	16	10	PZE_110012769	11241	33	8.2	2.2
DMY ($dt.ha^{-1}$)	1	1	PZE_101183895	228556	119.6	4.8	2.3
	2	3	PZE_103113115	172857 - 178134	78.3	8.4	4
	3	3	PZE_103159262	210755 - 210760	114.6	7.8	3.7
	4	6	PZE_106032535	75517 - 86627	15.5	13.3	6.9
	5	7	PZE_107069530	126351	58.1	9.3	5
	6	8	PZE_108057745	103023 - 103457	53	10.5	5.8
DtSILK (d)	1	1	PZE_101035008	19696 - 22646	37.8	22	6.9
	2	1	PZE_101205734	251079 - 254464	136.4	7.4	2.2
	3	2	PZE_102151348	197954	94.7	12.9	4.1
	4	3	PZE_103086165	142732 - 157202	61.7	10.2	3.7
	5	3	PZE_103122617	180515	78.6	16.3	4.7
	6	5	PZE_105049624	41635 - 58706	56.3	4.9	2.3
	7	6	PZE_106095370	150525 - 150588	72.6	17.5	5.1
	8	7	PZE_107045046	25471 - 104886	43.9	8.7	3.3
	9	7	PZE_107106025	158126	81.1	6.1	2
	10	8	PZE_108062521	111781	54.6	28.2	8.9
	11	9	PZE_109098632	143808	78.5	6.1	2.1

	12	10	PZE_110057591	110540 - 120784	49.8	7.8	3
DtTAS(d)	1	1	PZE_101032015	19641 - 21075	35.7	12.9	5
	2	2	PZE_102161022	207043	102.3	6.5	2.6
	3	3	PZE_103098157	158352	60.8	21.2	8.3
	4	3	PZE_103143600	199245 - 201331	102.2	7.8	3.2
	5	5	PZE_105143697	197846 - 200369	95.5	5	2.2
	6	7	PZE_107045046	25471 - 104886	43.9	6.1	2.9
	7	7	PZE_107099124	152685 - 155704	79.9	6	2.9
	8	8	PZE_108058411	104281 - 104625	53.7	20	7.6
	9	9	PZE_109090152	137787 - 138020	70.1	9.2	3.7
PH (cm)	1	1	PZE_101132703	171230 - 178401	82.3	10.6	4.1
	2	1	PZE_101196829	245032 - 245219	131	15.9	6
	3	3	PZE_103119393	178152 - 178564	79.1	8.9	2.8
	4	4	PZE_104073794	145614	55.7	6.4	2.5
	5	4	PZE_104138654	204861 - 226068	109.4	6.2	2.5
	6	5	PZE_105070660	74660 - 145496	61.9	5.2	2.5
	7	6	PZE_106044620	93734	20.4	20.4	6.4
	8	7	PZE_107077092	132190	66.5	19.9	6.5
	9	8	PZE_108038271	26346 - 65101	43.6	22.4	8.1
	10	8	PZE_108081297	133441 - 143002	71.9	8.6	3.3
	11	9	PZE_109085253	133933	68.1	14.6	5.4
	12	10	PZE_110013838	12922	35.9	10.4	4.2

Table S10 Results of the QTL detection in the dent design using the LDLA – 1-marker model. For each detected QTL, we showed its genetic position on the dent consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	-log10(p)	R^2 (%)
DMC (%)	1	1	PZE_101043094	29355	43.3	8.5	2.6
	2	2	PZE_102015152	6459	20.9	15.3	5.1
	3	2	PZE_102150016	196649	94	6.9	2.0
	4	3	PZE_103038375	33073	45.8	14.2	4.7
	5	4	PZE_104077580	151818	57	8.5	2.6
	6	5	PZE_105043990	31444	54.4	9.7	3.0
	7	6	PZE_106020569	16293	9.3	6.4	1.9
	8	6	PZE_106092387	148530	72	6.3	1.8
	9	8	PZE_108054499	97030	50.2	9.8	3.1
	10	8	PZE_108104357	159498	81.9	5.6	1.6
	11	9	PZE_109098884	143900	78.9	9.8	3.1
	12	10	PZE_110073412	130077	53.5	8.5	2.6
DMY ($dt.ha^{-1}$)	1	3	PZE_103115334	175353	78.1	9.0	3.6
	2	3	PZE_103162977	213416	117.8	9.0	3.7
	3	6	PZE_106031833	74460	14.6	14.0	6.0
	4	7	PZE_107026145	29198	41.6	7.7	3.0
	5	8	PZE_108054494	97029	50.5	10.4	4.3
DtSILK (d)	1	1	PZE_101032230	19696	37.8	8.8	2.4
	2	1	PZE_101076734	60701	61	11.4	3.3
	3	2	PZE_102150016	196649	94	8.9	2.5
	4	3	PZE_103104448	165863	71.3	15.7	4.7
	5	3	PZE_103132614	188530	91.1	8.1	2.2
	6	5	PZE_105054634	51432	55.4	6.4	1.7
	7	6	PZE_106092387	148530	72	13.8	4.1
	8	7	PZE_107023943	25471	43.9	13.0	3.8
	9	8	PZE_108026961	27634	45.5	5.7	1.5
	10	8	PZE_108063387	113292	57.3	10.1	2.9
	11	10	PZE_110062675	117753	50.1	6.9	1.8
DtTAS(d)	1	1	PZE_101035341	23055	39.2	13.3	4.6
	2	3	PZE_103104448	165863	71.3	26.8	10.1
	3	6	PZE_106092387	148530	72	5.8	1.8
	4	7	PZE_107061937	118305	50.8	6.8	2.1
	5	7	PZE_107100713	155251	79.6	6.7	2.1

	6	8	PZE_108057325	102454	51.5	22.5	8.3
	7	9	PZE_109071914	116752	59.8	7.3	2.3
PH (cm)	1	1	PZE_101132469	170852	80.9	6.5	1.9
	2	1	PZE_101191970	238427	127.1	8.4	2.6
	3	2	PZE_102068532	46438	61.8	6.8	2.0
	4	3	PZE_103104448	165863	71.3	6.0	1.7
	5	4	PZE_104082879	156995	61.4	6.0	1.7
	6	6	PZE_106040890	89137	19.4	19.4	6.7
	7	7	PZE_107084200	139526	76.4	15.9	5.4
	8	8	PZE_108014288	14088	32	15.9	5.4
	9	9	PZE_109080822	128851	65.9	18.5	6.4
	10	10	PZE_110009551	7563	29.4	6.3	1.8

Table S11 Results of the QTL detection in the joint analysis using the connected model. For each detected QTL, we showed its genetic position on the dent-flint consensus map, its confidence interval, its level of significance and the partial percentage of variance explained. We also showed the name of one of the markers located at the detected position and their range of physical position(s) on the B73 v2 genome (Gore *et al.* 2009).

Trait	Nb	Chr	Marker	Physical position (kb)	Genetic position (cM)	-log10(p)	R^2 (%)
DMC (%)	1	1	PZE_101032230	19696 - 19975	41.3	8.5	2.3
	2	1	PZE_101103995	104611 - 113689	93.3	5.2	1.7
	3	1	PZE_101202934	251103 - 251159	160.4	4.5	1.6
	4	1	PZE_101247063	292581	200.4	4.3	1.5
	5	2	PZE_102012595	5556	17.2	21.4	4.5
	6	2	PZE_102178263	220854	131.3	6.9	2
	7	3	PZE_103033638	26310 - 30050	45.4	9.3	2.5
	8	3	PZE_103100449	160755	65.5	7.5	2.1
	9	4	PZE_104032843	40344 - 65470	54.6	13.8	3.3
	10	4	PZE_104143137	231732	130.7	4.4	1.6
	11	5	PZE_105025123	12581	42.6	17.6	3.8
	12	6	PZE_106005094	6514	6.3	10.1	2.6
	13	6	PZE_106082658	139918 - 142454	66	11.8	2.9
	14	7	PZE_107012564	9201	33.9	4.7	1.6
	15	8	PZE_108063387	112547 - 113298	63.4	21.7	4.5
	16	9	PZE_109010670	11079 - 11504	30.8	3.9	1.5
	17	9	PZE_109096248	141983	82.5	5.4	1.7
	18	10	PZE_110047687	89209 - 111680	47.7	68.5	12.5
DMY (dt.ha-1)	1	1	PZE_101145302	188026 - 188087	109.4	12.2	3.6
	2	1	PZE_101215394	266047	170.9	10.4	3.2
	3	2	PZE_102013856	5997 - 6049	18.5	5.2	2.1
	4	2	PZE_102066516	44332	67.4	7.2	2.6
	5	3	PZE_103010658	5853	21.6	7.6	2.6
	6	3	PZE_103098655	158895 - 161562	65.7	9.5	3.1
	7	3	PZE_103162977	213416	120.4	5.3	2.1
	8	4	PZE_104025845	28986 - 32061	51.5	13.5	3.9
	9	5	PZE_105103128	155811 - 160460	78	8.8	2.9
	10	6	PZE_106037747	81440 - 86559	18.3	9.9	3.1
	11	6	PZE_106050075	99944	34	11.4	3.4
	12	6	PZE_106106971	156749	88.4	9.2	3
	13	7	PZE_107025551	28013 - 100690	44.8	8.3	2.8
	14	7	PZE_107127637	170111 - 170248	119.3	5.9	2.3
	15	8	PZE_108060398	107884 - 111781	62.1	12.1	3.6

	16	10	PZE_110043381	82670 - 84599	43.2	38.6	8.9
DtSILK (d)	1	1	PZE_101005770	4452 - 4610	9.4	6.4	1.8
	2	1	PZE_101034085	21984 - 21992	42.7	10	2.3
	3	1	PZE_101105390	102985 - 118116	93.6	19.2	3.7
	4	1	PZE_101195591	244158 - 244596	155.9	11.2	2.5
	5	2	PZE_102161485	206123 - 207224	123	13.3	2.8
	6	3	PZE_103098655	158895 - 161562	65.7	12.7	2.7
	7	3	PZE_103128597	185274 - 187610	93.2	9.1	2.2
	8	4	PZE_104025181	29345 - 30933	51.8	22.8	4.2
	9	5	PZE_105050638	42662 - 51518	61.2	16.8	3.3
	10	6	PZE_106097991	151792	77.3	11.7	2.6
	11	7	PZE_107072354	128648 - 128709	68.9	14.8	3.1
	12	8	PZE_108061059	107884 - 109378	60.7	27.9	5
	13	9	PZE_109010476	11398	30.3	11.4	2.5
	14	9	PZE_109094832	141175	82.6	5.1	1.6
	15	10	PZE_110047800	89438 - 106051	47.4	93	15.2
DtTAS(d)	1	1	PZE_101033489	21569 - 22464	43	12.1	2.5
	2	1	PZE_101140981	182104 - 184245	105	41.4	6.6
	3	1	PZE_101216412	267537 - 267568	171.5	12.3	2.5
	4	3	PZE_103098655	158895 - 161562	65.7	30.9	5.1
	5	3	PZE_103152007	205694	109.8	11.3	2.4
	6	4	PZE_104022348	23525 - 25988	49.6	25.7	4.4
	7	5	PZE_105059330	58137 - 72409	66	29.1	4.9
	8	5	PZE_105138874	193728	108.2	7.8	1.9
	9	6	PZE_106090469	147428	71.3	6.8	1.7
	10	7	PZE_107040665	66316 – 171898	75.4	13.6	2.7
	11	7	PZE_107130789	171926	126.3	6.2	1.6
	12	8	PZE_108018453	18973	42.2	11.8	2.4
	13	8	PZE_108070788	123843	69.2	20.3	3.6
	14	9	PZE_109020361	20598 - 20829	47.7	10.7	2.3
	15	9	PZE_109089874	137784	78	10.3	2.2
	16	9	PZE_109119196	153947	120.8	6.8	1.7
	17	10	PZE_110050293	94969 - 106961	47.5	77	12
PH (cm)	1	1	PZE_101021455	12363	29.3	5.1	1.4
	2	1	PZE_101106839	111278 - 150672	93.9	20.8	3.7
	3	1	PZE_101184213	229073	145.9	13.7	2.7
	4	2	PZE_102011812	5379	17.1	5.9	1.6
	5	2	PZE_102076989	59015 - 62213	74.2	20.2	3.6
	6	2	PZE_102169349	212884	128.1	9.6	2.1
	7	3	PZE_103017768	10455	33.7	8.1	1.9

8	3	PZE_103132826	188571 - 188925	94.8	11.8	2.5
9	3	PZE_103175533	221582 - 221583	135.2	7.1	1.8
10	4	PZE_104022152	23948 - 24979	49.7	13.8	2.7
11	4	PZE_104132688	215436 - 227111	120.7	9	2
12	5	PZE_105084182	101590 - 150275	73.1	9.6	2.1
13	5	PZE_105152260	203315	120.3	7.4	1.8
14	6	PZE_106049618	98629	30.4	21.2	3.8
15	7	PZE_107072030	128141 - 128146	66.4	18	3.3
16	7	PZE_107126258	168905	115.7	10.4	2.3
17	8	PZE_108009237	9875	25.6	7.3	1.8
18	8	PZE_108056704	101776 - 102656	57	8.1	1.9
19	8	PZE_108096469	152593 - 153140	85.5	17.3	3.2
20	9	PZE_109077113	124694 - 130885	70.9	19.7	3.6
21	10	PZE_110047799	89438 - 97551	46.6	77.4	12.2

Appendix II: Supporting information of Chapter 2

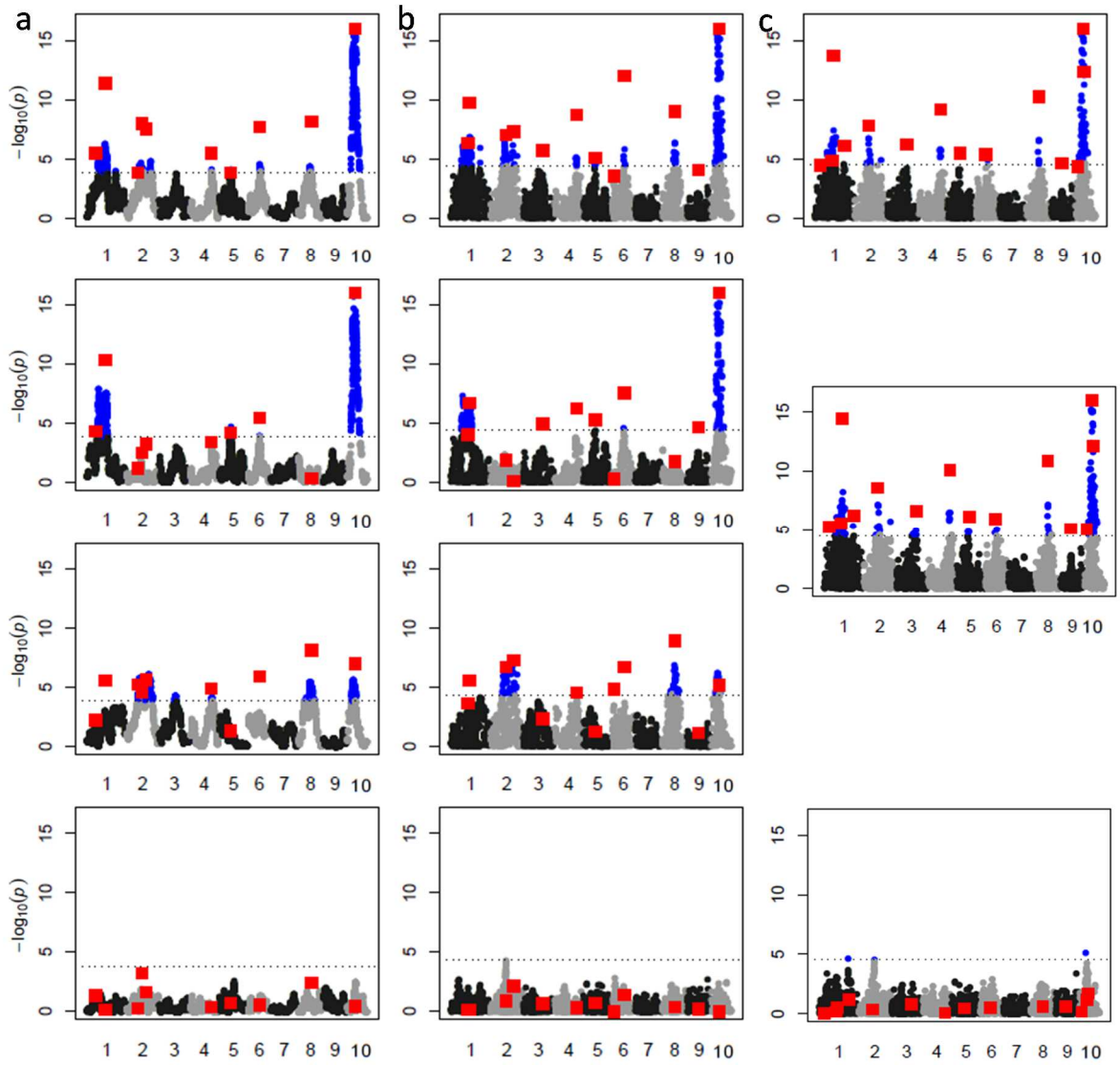

Figure S1 -log(p-value) curves of QTL detection for Dry Matter Content (DMC) with (a) the "Founder alleles" model, (b) the "SNP within-group" model, (c) the "Hybrid genotype" model. The chromosome number is indicated on the abscissa. For the "Founder alleles" (a) and the "SNP within-group" (b) models, the graphic at the top corresponds to the global effects at the markers. The other graphics correspond to the different components of the decomposed effects: from the top to the bottom: the flint GCA, the dent GCA, the SCA. For the "Hybrid genotype" model, the graphic at the top corresponds to the global effect at the markers, the middle part to the additive effect of the markers and the bottom part to the dominance effect of the markers. The grey and black dots correspond to the significance levels of tests below the threshold at the first step of the forward procedure, the blue dots correspond to those that were above the threshold. The red squares correspond to the –log(p-value) of the QTLs that were included in the final multi-locus model. An upper limit for the –log(pval) was fixed to 16.

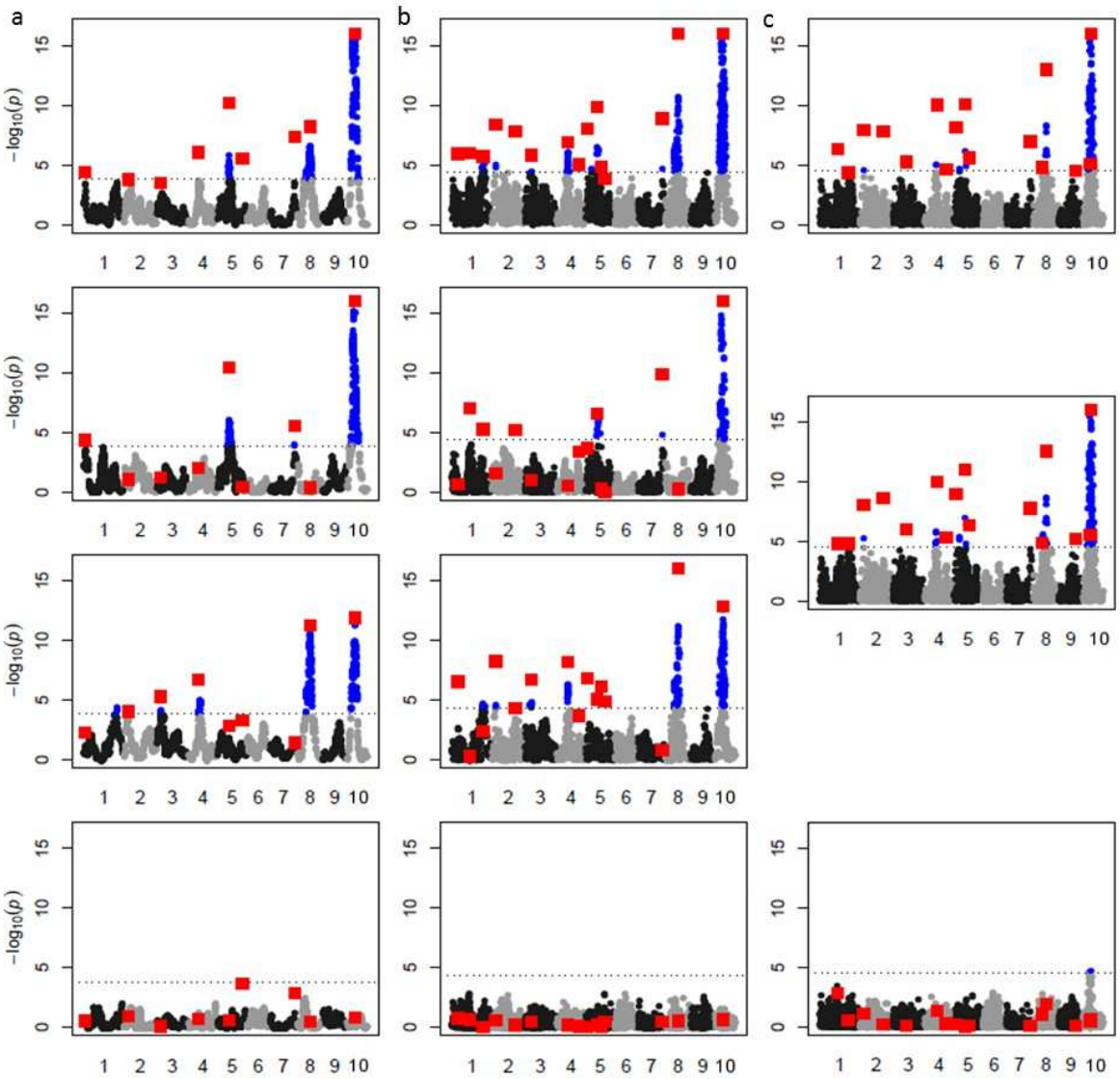

Figure S2 -log(p-value) curves of QTL detection for Female Flowering (DtSILK) with (a) the "Founder alleles" model, (b) the "SNP within-group" model, (c) the "Hybrid genotype" model. The chromosome number is indicated on the abscissa. For the "Founder alleles" (a) and the "SNP within-group" (b) models, the graphic at the top corresponds to the global effects at the markers. The other graphics correspond to the different components of the decomposed effects: from the top to the bottom: the flint GCA, the dent GCA, the SCA. For the "Hybrid genotype" model, the graphic at the top corresponds to the global effect at the markers, the middle part to the additive effect of the markers and the bottom part to the dominance effect of the markers. The grey and black dots correspond to the significance levels of tests below the threshold at the first step of the forward procedure, the blue dots correspond to those that were above the threshold. The red squares correspond to the –log(p-value) of the QTLs that were included in the final multi-locus model. An upper limit for the –log(pval) was fixed to 16.

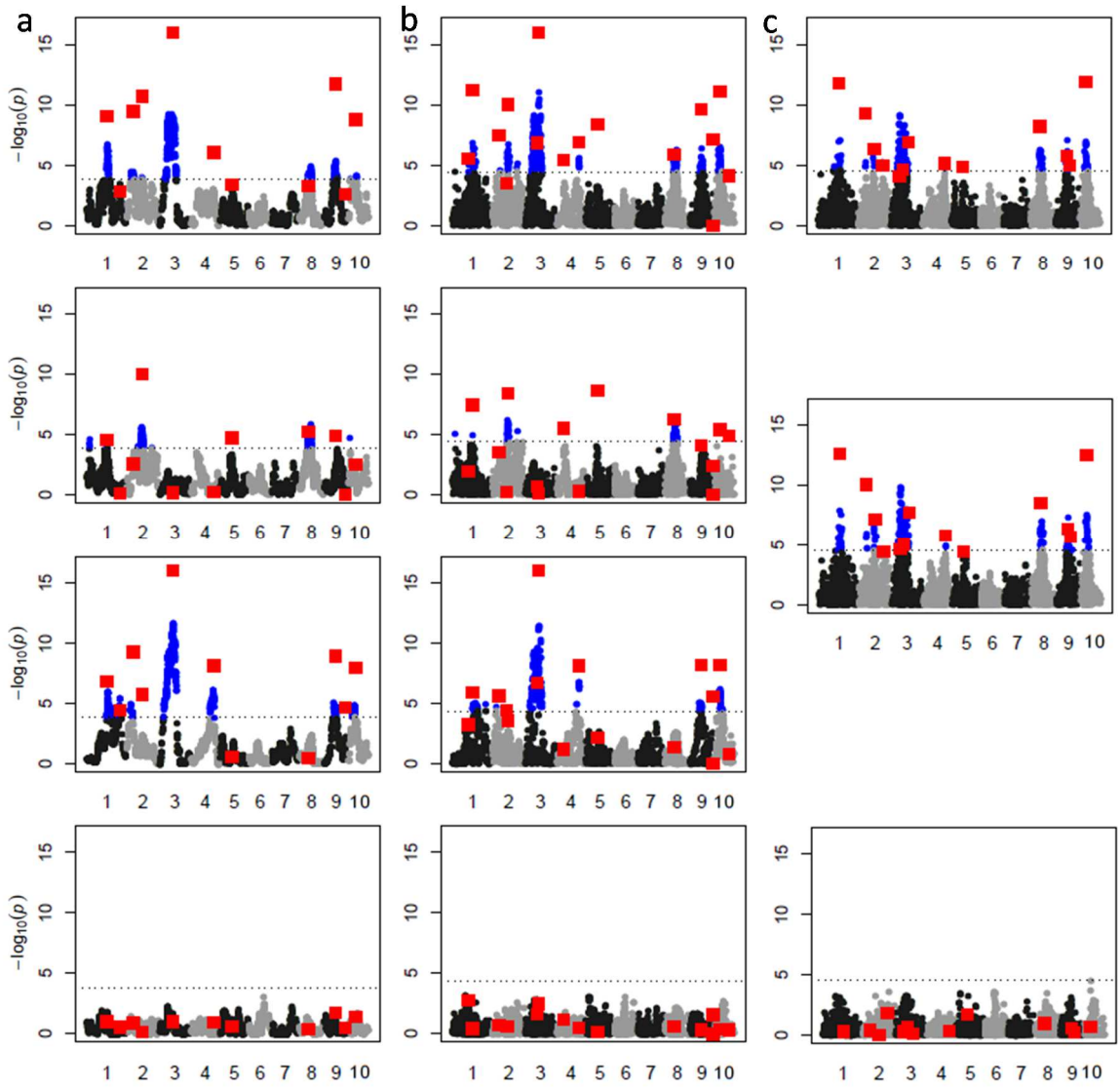

Figure S3 -log(p-value) curves of QTL detection for Plant height (PH) with (a) the "Founder alleles" model, (b) the "SNP within-group" model, (c) the "Hybrid genotype" model. The chromosome number is indicated on the abscissa. For the "Founder alleles" (a) and the "SNP within-group" (b) models, the graphic at the top corresponds to the global effects at the markers. The other graphics correspond to the different components of the decomposed effects: from the top to the bottom: the flint GCA, the dent GCA, the SCA. For the "Hybrid genotype" model, the graphic at the top corresponds to the global effect at the markers, the middle part to the additive effect of the markers and the bottom part to the dominance effect of the markers. The grey and black dots correspond to the significance levels of tests below the threshold at the first step of the forward procedure, the blue dots correspond to those that were above the threshold. The red squares correspond to the –log(p-value) of the QTLs that were included in the final multi-locus model. An upper limit for the –log(pval) was fixed to 16.

Table S1 Results of the QTLs detection for the "Hybrid Genotype" model for the four studied traits: Dry Matter Content (DMC), Dry Matter Yield (DMY), Silking Date (DtSILK), Plant Height (PH). For each QTL, the chromosome (Chr), the marker (Mk), the genetic position in cM (Gen pos), the physical position in kbp (Phys pos), the –log(p-value) of the additive part of its effect (-log(Ad)), the –log(p-value) of the dominance part of its effect (-log(Dom)), the –log(p-value) of its global effect (-log(Global)), the explained R^2 in % (R^2) are indicated.

Trait	Nb	Chr	Mk	Gen pos	Phys pos	-log(Ad)	-log(Dom)	-log(Global)	R^2
DMC									
	1	1	PZE-101023852	28.7	14032	5,2	0,04	4,46	0,91
	2	1	PZE-101107138	97.9	112147	5,6	0,17	4,83	0,98
	3	1	PZE-101141198	107.1	182293	14,5	0,49	13,68	2,76
	4	1	PZE-101210621	173.7	260145	6,2	1,17	6,10	1,23
	5	2	PZE-102080558	82.3	64362	8,6	0,36	7,83	1,60
	6	3	PZE-103142982	113.4	198521	6,1	0,45	5,47	1,24
	7	4	PZE-104129789	130.5	210477	10,1	0,08	9,17	1,87
	8	5	PZE-105100982	73.2	151631	6,6	0,77	6,19	1,10
	9	6	PZE-106068323	51.9	121253	6,0	0,49	5,38	1,09
	10	8	PZE-108075290	71	130926	10,9	0,60	10,22	2,09
	11	9	PZE-109038235	58.7	56424	5,1	0,59	4,62	0,93
	12	10	PZE-110007567	10.9	5875	5,1	0,17	4,34	0,87
	13	10	PZE-110049918	44.3	94001	37,7	1,18	37,2	7,47
	14	10	PZE-110060381	48.8	114656	12,1	1,71	12,32	2,48
DMY									
	1	2	PZE-102096468	86	111010	7,9	1,28	7,860	2,21
	2	2	PZE-102160945	135.8	207038	5,0	0,04	4,251	1,20
	3	3	PZE-103108225	76.5	169072	8,4	1,90	8,876	2,48
	4	3	PZE-103125956	95.8	183577	5,3	0,09	4,522	1,28
	5	4	PZE-104078143	74.5	152444	7,4	0,16	6,553	1,90
	6	4	PZE-104129789	130.5	210477	5,6	0,36	4,988	1,44
	7	5	PZE-105077552	65.2	87115	4,8	1,94	5,430	1,49
	8	6	PZE-106055176	39.1	105598	5,1	0,19	4,404	1,25
	9	8	PZE-108088583	77.5	145907	11,9	1,70	12,163	3,46
	10	9	PZE-109052698	60.5	91043	4,9	0,32	4,252	1,18
	11	10	PZE-110020953	33.7	27697	10,6	0,02	9,647	2,75
DtSILK									
	1	1	PZE-101138117	105.5	179183	4,8	2,90	6,306	1,53
	2	1	PZE-101199598	164.9	248737	4,8	0,60	4,335	1,05
	3	2	PZE-102018300	22.3	8782	8,0	1,16	7,881	1,91
	4	2	PZE-102160945	135.8	207038	8,6	0,27	7,799	1,92
	5	3	PZE-103109418	77.8	170117	6,0	0,13	5,237	1,27
	6	4	PZE-104078143	74.5	152444	10,0	1,39	10,009	2,45
	7	4	PZE-104117192	125.1	193348	5,3	0,30	4,654	1,14
	8	5	PZE-105012348	12.9	5384	9,0	0,24	8,134	1,98
	9	5	PZE-105077552	65.2	87115	11,0	0,00	10,042	2,44
	10	5	PZE-105117653	89.7	174358	6,4	0,14	5,583	1,36
	11	7	PZE-107132427	141.6	172725	7,8	0,17	6,930	1,67
	12	8	PZE-108028681	49.8	26352	4,9	1,10	4,784	1,18

	13	8	PZE-108077879	71.8	133563	12,5	1,94	12,954	3,10
	14	9	PZE-109098682	95.5	143862	5,2	0,12	4,494	1,09
	15	10	PZE-110046358	42.1	87170	5,6	0,53	5,040	1,24
	16	10	PZE-110049918	44.3	94001	50,6	0,73	49,7	12,01
PH									
	1	1	PZE-101152239	117.8	195684	12,6	0,30	11,75	2,67
	2	2	PZE-102037297	43.3	17988	10,1	0,41	9,30	2,11
	3	2	PZE-102119036	93.8	160729	7,2	0,03	6,33	1,41
	4	2	PZE-102168063	142.1	211948	4,4	1,83	4,94	1,08
	5	3	PZE-103051361	49.5	56901	4,7	0,29	4,03	0,92
	6	3	PZE-103102119	69.8	162433	5,0	0,71	4,63	1,02
	7	3	PZE-103128864	98.8	185839	7,7	0,16	6,90	1,54
	8	4	PZE-104129789	130.5	210477	5,8	0,40	5,12	1,16
	9	5	PZE-105075570	64.1	83398	4,4	1,70	4,87	1,04
	10	8	PZE-108036758	52.6	57215	8,5	1,00	8,20	1,86
	11	9	PZE-109061773	64	103353	6,3	0,59	5,79	1,31
	12	9	PZE-109082918	77.1	131575	5,6	0,27	4,93	1,13
	13	10	PZE-110020737	33.7	27318	12,5	0,71	11,87	2,68

Table S2 Results of the QTLs detection for the "SNP within-group" model for the four studied traits: Dry Matter Content (DMC), Dry Matter Yield (DMY), Silking Date (DtSILK), Plant Height (PH). For each QTL, the chromosome (Chr), the marker (Mk), the genetic position in cM (Gen pos), the physical position in kbp (Phys pos), the –log(p-value) of the flint GCA part of its effect (-log(GCAf)), the –log(p-value) of the dent GCA part of its effect (-log(GCAd)), the –log(p-value) of the SCA part of its effect (-log(SCA)), the –log(p-value) of its global effect (-log(Global)), the explained R^2 in % (R^2) are indicated.

Trait	Nb	Chr	Mk	Gen pos	Phys pos	-log (GCAf)	-log (GCAd)	-log (SCA)	-log (Global)	R^2
DMC										
	1	1	PZE-101107138	97,9	112147	4,01	3,63	0,17	6,31	1,42
	2	1	PZE-101141198	107,1	182293	6,72	5,61	0,15	9,73	2,15
	3	2	PZE-102116089	92,1	153797	1,86	6,73	0,91	7,04	1,58
	4	2	PZE-102159268	134,4	205898	0,11	7,26	2,14	7,29	1,64
	5	3	PZE-103142982	113,4	198521	4,95	2,35	0,67	5,70	1,30
	6	4	PZE-104129789	130,5	210477	6,22	4,59	0,35	8,68	1,94
	7	5	PZE-105096988	70,6	142855	5,26	1,26	0,72	5,10	1,17
	8	6	PZE-106006210	5,2	7781	0,27	4,84	0,01	3,56	0,84
	9	6	PZE-106090096	70	147212	7,53	6,71	1,40	12,01	2,62
	10	8	PZE-108074836	70,8	130409	1,76	8,92	0,36	8,99	1,98
	11	9	PZE-109057266	62,2	98436	4,64	1,15	0,23	4,07	0,95
	12	10	PZE-110049918	44,3	94001	38,78	5,18	0,02	41,13	8,64
DMY										
	1	2	PZE-102120732	95	164884	8,96	3,59	0,10	9,81	2,95
	2	3	PZE-103008521	9,4	4704	0,82	4,71	0,49	3,83	1,29
	3	3	PZE-103124449	93,4	181769	0,71	12,18	0,17	10,66	3,33
	4	4	PZE-104144717	147	233539	2,73	3,59	0,50	4,64	1,50
	5	5	PZE-105077135	65	86222	5,23	3,42	0,14	6,77	2,10
	6	6	PZE-106048775	29,9	97805	0,24	5,03	0,40	3,92	1,33
	7	8	PZE-108073565	70,2	128542	3,69	12,84	0,81	14,40	4,42
	8	9	PZE-109052698	60,5	91043	4,55	0,60	0,32	3,61	1,16
	9	10	PZE-110045871	41,9	86698	5,27	7,23	0,02	10,20	3,16
DtSILK										
	1	1	PZE-101029748	35.2	17832	0,65	6,54	0,80	5,93	1,52
	2	1	PZE-101138117	105.5	179183	7,02	0,30	0,70	5,99	1,56
	3	1	PZE-101217474	180.1	268814	5,26	2,41	0,07	5,69	1,50
	4	2	PZE-102018300	22.3	8782	1,56	8,25	0,63	8,38	2,11
	5	2	PZE-102160945	135.8	207038	5,20	4,35	0,24	7,80	2,00
	6	3	PZE-103026244	37.7	19256	1,02	6,71	0,50	5,82	1,50
	7	4	PZE-104058608	65.7	112760	0,60	8,22	0,22	6,87	1,76
	8	4	PZE-104126472	129.5	204417	3,38	3,72	0,13	5,02	1,33
	9	5	PZE-105006205	7.5	3143	3,68	6,81	0,06	8,02	2,01
	10	5	PZE-105077552	65.2	87115	6,56	5,09	0,21	9,82	2,47
	11	5	PZE-105117617	89.7	174351	0,34	6,13	0,13	4,85	1,27
	12	5	PZE-105138874	112.5	193728	0,00	4,94	0,48	3,82	1,01
	13	7	PZE-107132427	141.6	172725	9,88	0,83	0,53	8,85	2,23
	14	8	PZE-108077879	71.8	133563	0,29	19,40	0,57	17,84	4,30
	15	10	PZE-110049918	44.3	94001	47,23	12,82	0,69	57,53	13,72

PH										
	1	1	PZE-101091535	90,9	83644	1,94	3,27	2,81	5,50	1,15
	2	1	PZE-101150835	116,8	194674	7,42	5,91	0,45	11,24	2,39
	3	2	PZE-102032234	37,1	15025	3,48	5,64	0,71	7,48	1,62
	4	2	PZE_102084168	84	71896	0,24	4,44	0,65	3,50	0,79
	5	2	PZE-102115483	91,8	152503	8,42	3,61	0,61	10,02	2,15
	6	3	PZE-103100427	68,3	160672	0,67	6,75	1,67	6,81	1,44
	7	3	PZE-103109418	77,8	170117	0,12	27,23	2,48	27,10	5,31
	8	4	PZE-104019121	39	19433	5,50	1,21	1,20	5,43	1,22
	9	4	PZE-104126472	129,5	204417	0,28	8,15	0,53	6,89	1,45
	10	5	PZE-105094920	69,9	139235	8,63	2,20	0,17	8,35	1,83
	11	8	PZE-108038507	52,7	62847	6,24	1,41	0,60	5,89	1,33
	12	9	PZE-109061773	64	103353	4,08	8,17	0,33	9,59	2,06
	13	9	PZE-109121844	131,7	155733	2,41	5,59	1,57	7,10	1,54
	14	10	PZE_110035195	35	66734	5,40	8,20	0,37	11,07	2,36
	15	10	PZE-110094832	85,2	142069	4,90	0,83	0,34	4,09	0,94

Table S3 Results of the QTLs detection for the "Founder Alleles" model for the four studied traits: Dry Matter Content (DMC), Dry Matter Yield (DMY), Silking Date (DtSILK), Plant Height (PH). For each QTL, the chromosome (Chr), the marker (Mk), the genetic position in cM (Gen pos), the physical position in kbp (Phys pos), the –log(p-value) of the flint GCA part of its effect (-log(GCAf)), the –log(p-value) of the dent GCA part of its effect (-log(GCAd)), the –log(p-value) of the SCA part of its effect (-log(SCA)), the –log(p-value) of its global effect (-log(Global)), the explained R^2 in % (R^2) are indicated.

Trait	Nb	Chr	Mk	Gen pos	Phys pos	-log (GCAf)	-log (GCAd)	-log (SCA)	-log (Global)	R^2
DMC										
	1	1	PZE_101043600	50.6	29682	4,29	2,23	1,38	5,48	2,05
	2	1	PZE_101139638	106.1	180401	10,32	5,55	0,19	11,37	3,35
	3	2	PZE_102057464	67.1	36146	1,21	5,28	0,29	3,84	1,67
	4	2	PZE_102115483	91.8	152503	2,48	4,54	3,26	7,95	2,60
	5	2	PZE_102139190	113.6	187622	3,23	5,65	1,65	7,53	2,50
	6	4	PZE_104114190	122	189586	3,42	4,90	0,41	5,49	2,06
	7	5	PZE_105087042	67.5	111541	4,20	1,32	0,75	3,86	1,67
	8	6	PZE_106084109	64.5	141285	5,47	5,90	0,56	7,70	2,55
	9	8	PZE_108090858	79.2	148086	0,36	8,12	2,46	8,17	2,62
	10	10	PZE_110050269	44.5	94965	33,4	6,98	0,46	34,5	7,91
DMY										
	1	2	PZE_102109928	89.3	141598	4,34	0,74	1,73	4,56	2,25
	2	2	PZE_102136109	111.2	185694	5,36	0,22	2,37	5,52	2,55
	3	3	PZE_103005573	6.2	3433	0,01	5,02	1,97	4,57	2,27
	4	3	PZE_103108225	76.5	169072	2,36	1,80	2,74	4,93	2,33
	5	3	PZE_103125892	95.8	183555	1,20	5,65	0,08	3,31	1,92
	6	4	PZE_104138099	136.3	225269	1,83	4,47	1,95	5,86	2,66
	7	5	PZE_105020816	23.5	10009	4,59	0,62	0,62	3,47	1,90
	8	5	PZE_105078900	65.8	89718	5,93	2,61	0,41	5,69	2,57
	9	6	PZE_106044414	25	93501	0,06	5,80	0,35	3,26	1,90
	10	8	PZE_108062087	62.4	110990	3,25	12,31	0,96	12,10	4,39
	11	9	PZE_109061773	64	103353	6,12	0,20	0,96	4,47	2,24
	12	10	PZE_110051444	44.9	96836	2,86	4,53	0,42	4,93	2,39
DtSILK										
	1	1	PZE-101004387	7.7	3883	4,37	2,28	0,56	4,37	2,40
	2	2	PZE-102021400	25.3	9973	1,09	4,05	0,96	3,77	2,22
	3	3	PZE-103016459	20.4	9311	1,26	5,30	0,10	3,48	2,13
	4	4	PZE-104034811	58.3	43970	2,02	6,75	0,74	6,02	2,98
	5	5	PZE-105075954	64.3	83980	10,43	2,84	0,61	10,16	4,16
	6	5	PZE-105166980	144.8	210286	0,47	3,32	3,68	5,51	2,71
	7	7	PZE-107133704	143.1	173181	5,53	1,46	2,87	7,32	3,23
	8	8	PZE-108077879	71.8	133563	0,39	11,26	0,54	8,21	3,63
	9	10	PZE-110050273	44.5	94968	32,5	11,90	0,83	39,5	11,68
PH										
	1	1	PZE_101150835	116.8	194674	4,54	6,87	1,01	9,01	2,88
	2	1	PZE_101233132	192.4	281321	0,13	4,42	0,57	2,79	1,42
	3	2	PZE_102031753	36.6	14800	2,56	9,27	0,90	9,42	3,03
	4	2	PZE_102116344	92.2	154273	9,96	5,75	0,20	10,67	3,32
	5	3	PZE_103108648	77.1	169577	0,19	20,8	1,04	17,1	4,42

6	4	PZE_104134644	132.4	218941	0,24	8,16	0,94	6,05	2,17
7	5	PZE_105097751	70.9	144099	4,71	0,61	0,62	3,34	1,57
8	8	PZE_108038053	52.7	61790	5,22	0,50	0,41	3,29	1,59
9	9	PZE_109061773	64	103353	4,85	8,94	1,74	11,71	3,49
10	9	PZE_109115897	122	152795	0,01	4,68	0,50	2,59	1,34
11	10	PZE_110045521	41.5	85840	2,52	7,94	1,41	8,74	2,80

Appendix III: Supporting information of Chapter 3

Table S1 Variance components of the phenotypic data, estimated directly on field plot data, for the four variance decomposition models presented in Giraud *et al.* (submitted), for the different traits: dry matter content DMC (in % of dry matter), dry matter yield DMY (in tons. per ha), female flowering time DtSILK (in days after 1st of January), plant height PH (in cm). For each variance, precision corresponding to 1.96 x the standard error of the estimated variance is indicated. We indicated the min and max values of the residual variances in the different trials.

Model	**Hybrid (1)**	**Population structure + Hybrid (2)**	**Flint GCA + Dent GCA + SCA (3)**	**Population structure + Flint GCA + Dent GCA + SCA (4)**
DMC				
Hybrid variance	3.68 ± 0.37	2.46 ± 0.27		
Flint GCA variance			2.31 ± 0.51	0.98 ± 0.44
Dent GCA variance			0.92 ± 0.53	0.96 ± 0.44
SCA variance			0.43 ± 0.54	0.53 ± 0.49
Residual variances: min and max	1.44 ± 0.18 : 6.15 ± 0.60	1.52 ± 0.19 : 6.05 ± 0.59	1.45 ± 0.18 : 6.13 ± 0.59	1.52 ± 0.18 : 6.04 ± 0.58
DMY				
Hybrid variance	1.24 ± 0.14	0.94 ± 0.11		
Flint GCA variance			0.28 ± 0.20	0.30 ± 0.19
Dent GCA variance			0.74 ± 0.19	0.44 ± 0.17
SCA variance			0.22 ± 0.22	0.20 ± 0.20
Residual variances: min and max	1.01 ± 0.12 : 3.53 ± 0.35	1.02 ± 0.12 : 3.56 ± 0.34	1.00 ± 0.11 : 3.53 ± 0.35	1.02 ± 0.12 : 3.56 ± 0.35
DtSILK				
Hybrid variance	2.38 ± 0.24	2.06 ± 0.21		
Flint GCA variance			0.74 ± 0.38	0.51 ± 0.37
Dent GCA variance			1.26 ± 0.35	1.09 ± 0.33
SCA variance			0.39 ± 0.37	0.47 ± 0.37
Residual variances: min and max	0.96 ± 0.11 : 6.12 ± 0.57	0.96 ± 0.11 : 6.11 ± 0.57	0.96 ± 0.11 : 6.13 ± 0.57	0.96 ± 0.11 : 6.12 ± 0.57
PH				
Hybrid variance	154.7 ± 15.7	97.6 ± 10.7		
Flint GCA variance			42.9 ± 22.9	38.4 ± 17.4
Dent GCA variance			92.8 ± 21.6	45.1 ± 16.0
SCA variance			19.3 ± 23.2	13.4 ± 17.3
Residual variances: min and max	49.7 ± 6.5 : 244.2 ± 23.7	46.4 ± 6.0 : 243.6 ± 23.5	50.0 ± 6.5 : 243.6 ± 23.6	46.7 ± 6.0 : 243.2 ± 23.5

Appendix IV: Résumé substantiel en français

Darwin en 1876 conclut après de multiples observations au bénéfice dans de nombreux cas de la pollinisation croisée par rapport à l'autofécondation. Cette observation a été par la suite théorisée parallèlement par Shull (1908) et East (1908) qui en se basant sur des observations sur le maïs (*Zea mays* L.) ont défini le concept d'hétérosis (Shull 1914). L'hétérosis se manifeste tant au niveau inter que intra-spécifique. On se concentrera dans cette thèse sur l'hétérosis intra-spécifique. L'hétérosis du sélectionneur est le fait que l'individu issu du croisement de deux parents d'une certaine nature (lignées, populations, clones individus hétérozygotes) est supérieur au meilleur de ses parents. L'hétérosis du généticien correspond à l'avantage d'un hybride par rapport à la moyenne de ses parents. L'hétérosis est l'opposé conceptuel de la dépression de consanguinité, qui correspond à une réduction graduelle de la vigueur suite à la reproduction sur plusieurs générations d'individus apparentés. Son importance diffère selon le système de reproduction (supérieure chez les espèces allogames) et le caractère étudié. Les bases génétiques de l'hétérosis sont peu connues mais plusieurs hypothèses non-exclusives existent pour l'expliquer : la dominance, la super-dominance, la pseudo-superdominance, la superdominance marginale.

Chez le maïs, la théorisation de l'hétérosis par East et Shull a été accompagnée aux USA par le développement d'hybrides. Les hybrides F1 permettaient de produire à grande échelle les meilleures combinaisons de gamètes dérivés de variétés populations. En 1922-1924 les meilleurs hybrides, comme par exemple Cooper Cross, présentaient un avantage répétable d'environ 10% par rapport aux meilleures variétés populations (Charcosset 2002). Des groupes hétérotiques (structurant la diversité génétique) ont été développés à partir des années 50 et les variétés actuelles correspondent généralement à des hybrides issus du croisement entre les lignées appartenant à des groupes hétérotiques complémentaires.

La valeur d'un hybride se décompose traditionnellement en deux parties (Sprague and Tatum 1942). La première est la somme des Aptitudes Générales à la Combinaison (AGC) des deux lignées parentales, chacune définies comme la valeur moyenne des hybrides obtenus par croisement de cette lignée avec les lignées du groupe complémentaire. La deuxième partie est l'Aptitude Spécifique à la Combinaison (ASC) de la paire de lignées parents de l'hybride. C'est la différence entre la valeur de l'hybride et sa valeur prédite sur la base des AGCs. Dans un programme de sélection, à cause de considérations pratiques, toutes les combinaisons de lignées inter-groupes hétérotiques ne peuvent être évaluées phénotypiquement. Par conséquent, la sélection est effectuée en deux étapes. Dans un premier temps, les potentielles lignées candidates de chaque groupe sont sélectionnées sur la base de la valeur de leur descendance en croisement avec une ou quelques lignées représentatives du groupe complémentaire, appelées

testeurs. Dans un second temps, un nombre limité de combinaisons entre les lignées sélectionnées des deux groupes sont évaluées pour identifier les meilleurs hybrides. Dans ce schéma, la majeure partie de la sélection est effectuée durant la première étape. A cause du faible nombre de testeurs considérés, la sélection des lignées est basée sur une combinaison de leur AGC et de l'ASC avec le(s) testeur(s). La compréhension de la part des AGCs et de l'ASC dans la variation des caractères d'intérêt est par conséquent importante pour évaluer à quel point le choix du (des) testeur(s) peut affecter l'estimation du potentiel de futures lignées.

L'étude des caractères quantitatifs *i.e.* des caractères présentant une variation continue, ce qui est le cas de la majorité des caractères d'intérêts agronomiques, cherche à comprendre les bases génétiques de ces caractères complexes. Les outils de génotypage développés depuis les années 80 permettent de rechercher des associations entre les variations moléculaires au niveau de l'ADN et celle des phénotypes. On peut distinguer deux intérêts principaux : la compréhension du déterminisme génétique des caractères (détection de QTL) et la sélection basée sur les QTLs détectés (sélection assistée par marqueurs) ou sur une valeur prédite de l'individu impliquant l'ensemble des marqueurs (sélection génomique). Chez les plantes, avec le développement des marqueurs moléculaires (RFLP, AFLP puis SSR notamment), la détection de QTL s'est développée essentiellement au sein de populations biparentales (Paterson *et al.* 1988). L'essor des SNPs et le développement du marquage moléculaire haut-débit ont entraîné une forte diminution des coûts de génotypage ce qui a permis d'augmenter le nombre d'individus génotypés et le nombre de points de génotypage. Ceci a induit des modifications dans la manière de penser la détection de QTL et les méthodes de sélection. En effet, cela a permis de développer la détection de QTL dans des populations à base génétique large issues de plus de deux parents et de prendre en compte en plus des informations de linkage, l'existence d'haplotypes communs entre parents, définis par la présence de marqueurs proches en déséquilibre de liaison. Des méthodes de type LDLA : « linkage disequilibrium and linkage analysis » ont été développées. Cette révolution technique a aussi permis le développement d'approches de sélection prenant globalement en compte l'ensemble des marqueurs du génome, ce qui a été défini sous le terme de sélection génomique (Meuwissen *et al.* 2001).

A notre connaissance, aucune détection de QTL n'a été effectuée dans un dispositif expérimental entre lignées issues directement de populations en ségrégation disponibles aux premiers stades de la sélection dans deux groupes hétérotiques complémentaires. Avec le développement des techniques d'haploïdes doublés, les sélectionneurs peuvent directement générer à chaque cycle de sélection des populations en ségrégation composées de lignées pures.

Au lieu d'utiliser un faible nombre de testeurs du groupe complémentaire pour évaluer les candidats à la sélection, il peut être plus pertinent d'évaluer directement des hybrides entre lignées non sélectionnées des deux groupes. Détecter des QTLs d'AGC et d'ASC permettrait d'identifier les meilleures combinaisons hybrides possibles et d'optimiser l'amélioration de futures lignées parentales au sein de chaque groupe.

Dans ce contexte, l'objectif de cette thèse est de comprendre les bases génétiques de la valeur hybride chez le maïs pour la production de biomasse. Pour cela, la première partie du travail a consisté à mettre en œuvre une approche de détection de QTL dans deux dispositifs multiparentaux, correspondant chacun à un groupe hétérotique (les cornés versus les dentés), évalué pour sa valeur en croisement avec un testeur du groupe complémentaire. Dans un deuxième temps nous avons réalisé une détection de QTL dans un factoriel obtenus en croisant deux dispositifs multiparentaux correspondant chacun à un groupe hétérotique. Ce dispositif nous a permis d'évaluer l'importance relative de l'ASC par rapport à l'AGC et de rechercher les locus impliqués dans chacune de ses composantes. Enfin nous avons estimé quel était son potentiel pour la mise en œuvre de la sélection génomique.

L'analyse de type « Linkage Disequilibrium Linkage analysis » de dispositifs multiparentaux révèle différents QTLs multi-alléliques pour la performance hybride dans les groupes hétérotiques cornés et dentés de maïs

Ces travaux ont été publiés dans la revue Genetics de Décembre 2014 (Genetics 198 : 1717-1734). Deux dispositifs de type « Nested Association Mapping » de maïs, adaptés aux conditions européennes, ont été dérivés à partir des groupes hétérotiques complémentaires dentés et cornés, utilisés pour la production d'hybrides en Europe du Nord. Ces dispositifs ont été créés dans le cadre du projet européen « CornFed ». Dix familles biparentales dentées (N=841) et 11 familles biparentales cornées (N=811) ont été génotypées avec 56110 marqueurs SNP (Ganal *et al.* 2011) et évaluées sur testeur, en utilisant la lignée centrale du dispositif réciproque. Cinq caractères ont été phénotypés : le contenu en matière sèche, le rendement ensilage, la date de floraison mâle, la date de floraison femelle et la hauteur de plante. Des cartes génétiques consensus dentée, cornée et cornée-dentée ont été construites, correspondant respectivement à 21878 marqueurs (6808 positions génétiques), 20406 marqueurs (7272 positions génétiques) et 25472 marqueurs (8124 positions génétiques). Pour chaque dispositif, différents modèles de détection de QTL ont été appliqués correspondant à différents codages au niveau des allèles. Les allèles ont été définis soit par rapport aux lignées parentales, soit

comme des allèles haplotypiques basés sur une identité par descendance (IBD) entre lignées parentales, soit comme les allèles observés au niveau des SNPs. Le regroupement des allèles parentaux a été effectué pour chaque dispositif à l'aide du package R « clusthaplo » (Leroux *et al.* 2014) en choisissant la taille de la fenêtre en se basant sur la décroissance du déséquilibre de liaison. Différents modèles de détection de QTL multilocus ont été testés. Le premier correspond à un modèle connecté multifamille conventionnel. La connexion entre familles est assurée par la lignée centrale et l'hypothèse est faite que chaque parent porte un allèle différent aux QTLs. Les deuxième et troisième modèles testés sont des modèles connectés LDLA multifamilles, utilisant le regroupement effectué à l'aide de « clushaplo » (deux tailles de fenêtre différentes). Ces trois modèles ont été mis en œuvre à l'aide du logiciel MCQTL_LD (Jourjon *et al.* 2005) en utilisant une méthode de détection QTL nommée « iterative composite interval QTL mapping » (Charcosset *et al.* 2000) et un risque de type I de 10% au niveau du génome en considérant toutes les familles ensemble. Le dernier modèle testé est un modèle LDLA simple marqueur, considérant que deux lignées parentales présentant le même allèle au marqueur sont IBD pour ce marqueur. Ce modèle a été mis en œuvre dans R en utilisant un script R dérivé de l'approche basée sur un modèle mixte multilocus présentée dans Segura *et al.* (2012) en considérant également un risque de type I de 10% au niveau du génome. Ensuite, les deux dispositifs ont été analysés conjointement en utilisant le modèle connecté conventionnel. Les effets des QTLs détectés avec l'analyse conjointe ont été testés dans chacun des deux dispositifs pris séparément.

Sur la base de la décroissance du déséquilibre de liaison deux tailles de fenêtre glissante ont été choisies pour effectuer le regroupement haplotypique : 2 et 5cM. Dans chaque dispositif, la fenêtre glissante de 5 cM a conduit à un nombre d'allèles ancestraux plus important. Ce nombre varie le long du génome. Le regroupement est plus important dans les régions télomériques que centromériques, où le nombre d'allèles ancestraux est souvent proche du nombre de lignées parentales.

Entre cinq et 16 QTLs ont été détectés selon le modèle, le caractère et le groupe génétique considéré. Dans le dispositif corné, un QTL majeur (R^2=27%) présentant des effets pléiotropiques, a été détecté sur le chromosome 10. Les autres QTLs présentent des effets plus faibles (R^2<10%). En moyenne les modèles de type LDLA ont détecté plus de QTLs mais expliquent un plus faible pourcentage de la variance. Le modèle connecté pour quasiment tous les caractères explique un pourcentage de variance plus important, ce qui est en accord avec le fait que la plupart des QTLs présentent des séries alléliques avec des valeurs relativement continues des effets. En comparant les positions des QTLs détectés dans chacun des dispositifs

par le modèle connecté conventionnel, seuls 15% des QTLs ont été trouvés comme communs aux deux dispositifs.

Table 1 Nombre de QTLs détectés (Nb) et pourcentage ajusté de variance expliqué par les QTLs détectés (R^2) pour les cinq caractères dans les deux dispositifs séparés pour chaque modèle et dans le dispositif conjoint pour le modèle connecté. Le nombre total de QTLs détectés sur l'ensemble des caractères et le pourcentage moyen de variance expliqués (colonne « Total ») sont indiqués.

	DMC		DMY		DtSILK		DtTAS		PH		**Total**	
	Nb	R^2 (%)	Nb	R^2 (%)	Nb	R^2 (%)	Nb	R^2 (%)	Nb	R^2 (%)	**Nb**	**R^2 (%)**
Dent												
Connected	12	51.4	8	32.7	11	52.3	7	41.2	14	57.1	**52**	**46.9**
LDLA - 5cM	15	51.1	5	22.5	12	53.7	11	49.2	13	54.1	**56**	**46.1**
LDLA - 2cM	16	53.6	6	23.4	12	53.2	9	45.1	12	49.5	**55**	**45.0**
LDLA – 1-marker	12	37.4	5	18.6	11	43.2	7	33.3	10	36.4	**45**	**33.8**
Flint												
Connected	8	46.0	11	48.6	15	69.3	12	65.3	9	52.3	**55**	**56.3**
LDLA - 5cM	11	49.2	10	41.9	14	67.5	13	61.1	10	51.7	**58**	**54.3**
LDLA - 2cM	8	42.1	12	45.3	11	62.0	14	62.2	11	51.9	**56**	**52.7**
LDLA – 1-marker	7	36.1	11	39.0	16	61.7	16	58.0	9	41.9	**59**	**47.3**
Joint												
Connected	18	54.6	16	45.5	15	59.7	17	61.4	21	61.2	**87**	**56.5**

L'analyse conjointe des deux dispositifs a permis de détecter 87 QTLs soit entre 15 et 21 QTLs selon le caractère. Parmi ces derniers, entre 27% (floraison femelle) et 41% (floraison mâle) étaient significatifs dans les deux groupes hétérotiques. Pour chaque caractère, un nombre supérieur ou égal de QTLs a été détecté avec l'analyse conjointe que dans les analyses par groupe hétérotique. Toutes les lignées présentent des QTLs avec des effets positifs et négatifs sur le rendement. La présence d'allèles favorables dans les deux groupes ouvre des perspectives pour l'amélioration de la production de biomasse du maïs ensilage par sélection assistée par marqueurs. Contrairement aux autres études sur maïs réalisées dans des dispositifs NAM évalué *per se*, les QTL trouvés dans notre étude correspondent à des caractères mesurés sur hybrides (lignées croisées à un testeur) qui reflètent directement la variation génétique utile en sélection dans les deux groupes hétérotiques.

Une efficacité variable des différents modèles de détection de QTL selon les caractères et la

région a été observée ce qui montre la complémentarité des différents codages alléliques pour déchiffrer les séries alléliques. Contrairement à des études précédentes (Bardol *et al.* 2013), un clair avantage des modèles LDLA utilisant l'information haplotypique n'a pas été observé ce qui est cohérent avec le faible apparentement des lignées parentales de nos dispositifs. Le faible nombre de QTLs communs entre les deux groupes confirme l'ancienne divergence des groupes hétérotiques cornés et dentés : plus de 500 ans (Tenaillon and Charcosset 2011). Peu ou pas de QTLs communs ont été détectés pour le rendement ensilage, soit via l'analyse conjointe soit par comparaison directe des QTLs détectés dans les analyses par groupe hétérotique. Ceci peut être dû au fait que le rendement ensilage, contrairement à la floraison (et indirectement au contenu en matière sèche) a été soumis à une sélection directionnelle et non stabilisatrice. Cette sélection directionnelle a pu avoir tendance à fixer des allèles différents entre les deux groupes hétérotiques.

La détection de QTL sur des hybrides de maïs (*Zea mays* L.) dérivés de deux dispositifs multiparentaux a permis la détection simultanée de QTLs d'aptitudes générales et spécifique à la combinaison pour les performances ensilage

L'objectif de ce travail était d'évaluer si, grâce au génotypage dense maintenant disponible à moindre coût, il peut être plus pertinent d'évaluer directement des hybrides entre lignées candidates de deux groupes génétiques plutôt que des hybrides avec un faible nombre de testeurs du groupe complémentaire. Par rapport à une évaluation sur testeur, un tel dispositif permet de décomposer les performances hybrides en AGC et ASC (composantes qui sont confondues dans le cas d'une évaluation sur testeur) et de réduire d'un facteur deux le nombre d'hybrides à phénotyper. Détecter des QTLs d'AGC et d'ASC permettrait d'identifier les meilleures combinaisons hybrides possibles et d'optimiser l'amélioration de futures lignées parentales au sein de chaque groupe. L'objectif de cette partie est d'estimer l'importance relative d'ASC et de l'AGC dans un tel dispositif et d'évaluer son intérêt pour la détection de QTL d'AGC et d'ASC. Ce travail s'inscrit dans le cadre du projet SAM-MCR financés par 7 entreprises de sélection privées (Caussade, Euralis, Limagrain, Maïsadour, Pioneer, RAGT et Syngenta) membres de l'association Promaïs. Pour évaluer cette stratégie basée sur un factoriel, un dispositif multiparental corné et un denté ont été crées (Figure 1). Pour chaque groupe hétérotique, six populations biparentales ont été développées à partir de quatre lignées fondatrices et les lignées de ces populations ont été croisées avec celles de l'autre groupe selon un factoriel incomplet. 1044 hybrides cornés-dentés ont été obtenus par croisement de 863

lignées dentées et 879 lignées cornées. Les huit lignées fondatrices (quatre cornées et quatre dentées) ont été génotypées avec une puce Illumina 50 K SNPs (Ganal et al. 2011) et les lignées parentales des hybrides avec une puce Affymetrix® de 18480 SNPs. Seuls les marqueurs PANZEA communs aux deux puces et polymorphes chez les fondateurs ont été considérés. Après filtres sur le taux de données manquantes, l'hétérozygotie et la fréquence de l'allèle minoritaire, 9643 marqueurs ont été gardés. Après filtres sur les lignées parentales, 1758 lignées (875 dentées et 883 cornées) ont été considérées pour la suite des analyses. Une carte génétique consensus cornée-dentée a été construite et est constituée de 9548 marqueurs polymorphes dans au moins un des deux groupes hétérotiques.

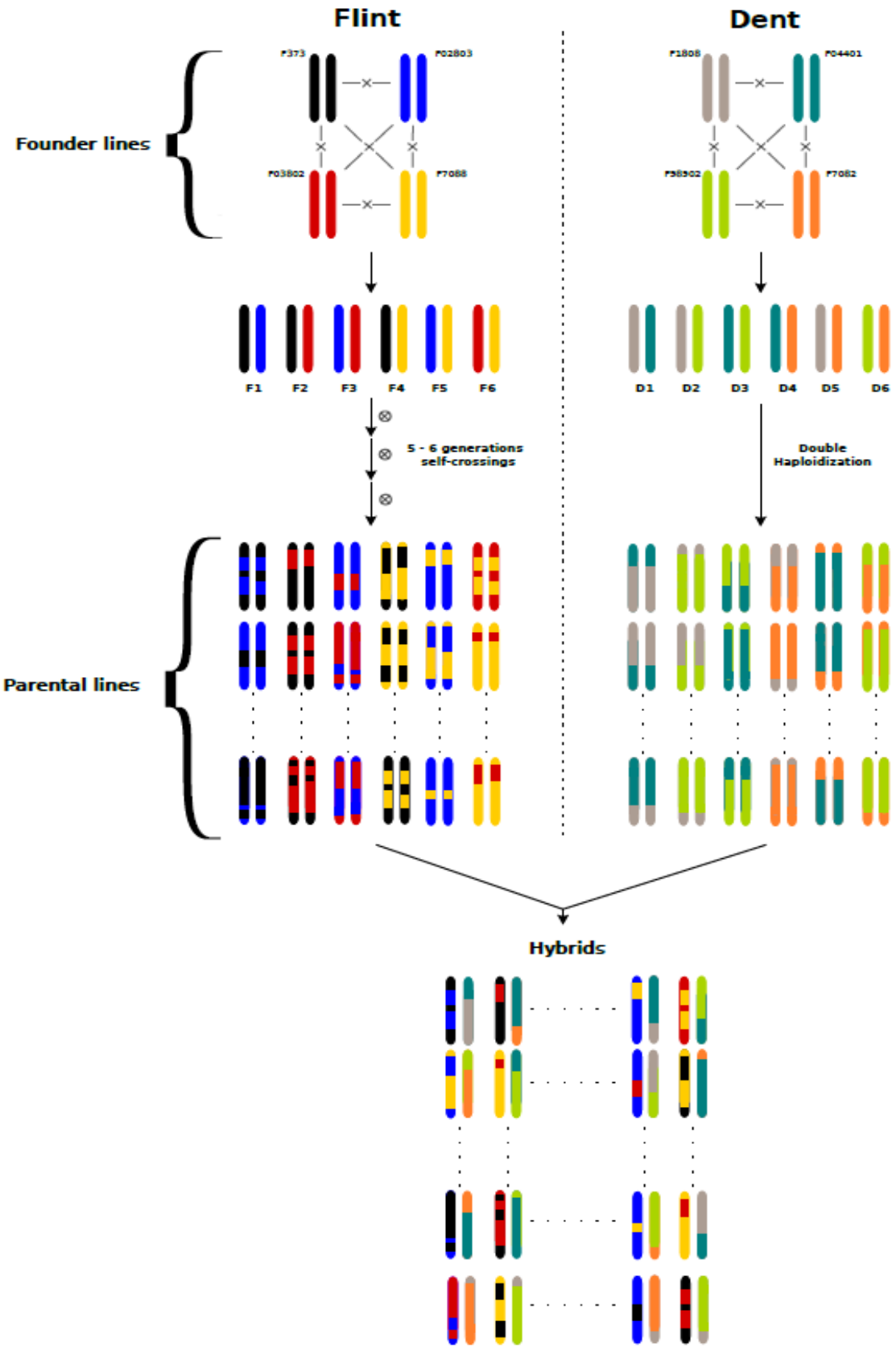

Figure 1 : Représentation schématique du dispositif expérimental.

Les hybrides ont été phénotypés pour le rendement ensilage, la matière sèche (% poids frais), la floraison femelle (jours après semis), la hauteur de plante (cm) dans 8 environnements (4 en 2013 et 4 en 2014) dans le nord de la France et en Allemagne. Les essais ont été mis en place selon un dispositif de type « augmented p-rep » et comprenaient 1088 parcelles. Dans un environnement donné, les hybrides entre les quatre lignées fondatrices et environ 17% des hybrides expérimentaux ont été évalués deux fois.

Différents modèles ont été mis en œuvre pour évaluer la part de l'AGC et l'ASC dans la variance génétique. Ils ont été mis en œuvre sur les données parcellaires concomitamment à la correction pour les hétérogénéités spatiales. Dans le premier modèle, il n'y a pas de décomposition de la valeur hybride. Dans le second, il y a prise en compte de la structure en populations (6 populations cornées, 6 populations dentées, 36 interactions). Dans le troisième, il y a décomposition de la valeur hybride en AGC cornée, AGC dentée et ASC. Dans le dernier, il y a prise en compte de la structure en population et décomposition en AGC cornée intra-population, AGC dentée intra-population et ASC intra-population. A partir de ces décompositions, des paramètres synthétiques sont pu être calculés. Les héritabilités au sens large sont fortes : entre 0.814 (rendement) et 0.892 (matière sèche), tout comme les héritabilités intra-population : entre 0.767 (rendement) et 0.876 (floraison femelle). La part de la variance génétique intrapopulation est élevée : entre 63.1% (hauteur de plante) et 86.7% (floraison femelle). L'ASC explique entre 13.8 et 22.6% de la variance hybride intra-population et entre 11.7% et 17.4% de la variance hybride selon le caractère,

Trois modèles mixtes de détection de QTL ont été développés dans ASReml-R (Butler *et al.* 2007 ; R Core Team 2013), considérant chacun un codage allélique différent, et ont été mis en œuvre sur les moyennes ajustées des hybrides. Le premier ou modèle « Allèles fondateurs » considère que chacune des huit lignées fondatrices porte un allèle différent au QTL (linkage analysis). A chaque marqueur, les probabilités des quatre lignées fondatrices dentées (respectivement cornées) ont été inférées à l'aide du logiciel PlantImpute (Hickey *et al.* 2015) pour chacun des 9548 marqueurs cartographiés. Le second modèle ou « SNP intra-groupe » considère directement l'allèle au SNP mais des effets des QTLs différents dans chaque groupe hétérotique. Le dernier modèle ou « Génotype de l'hybride » considère également directement l'allèle au SNP mais suppose de plus que les effets des QTLs sont identiques dans chaque groupe hétérotique. Les données de génotypage manquantes pour les deux derniers modèles ont été imputées à l'aide Beagle v3.0 (Browning and Browning 2009) et la détection a porté uniquement sur les 4758 marqueurs cartographiés et polymorphes dans les deux groupes. Pour les trois modèles, on a considéré un seuil de significativité de 5% au niveau de l'ensemble du

génome, basé sur le nombre de marqueurs efficaces (Gao *et al.* 2008). Une procédure multi-marqueurs a été implémentée composée de deux étapes : une sélection « forward » des marqueurs suivie par une sélection « backward ». Au total entre 42 et 54 QTLs ont été détectés selon le modèle. Les modèles « SNP intra-groupe » et « Génotype de l'hybride » ont détecté plus de QTLs au total et pour tous les caractères sauf le rendement. En considérant arbitrairement les QTLs détectés à moins de 10cM l'un de l'autre comme identiques, seuls 16 QTLs ont été détectés par les trois modèles. Certains ont été détectés avec seulement deux modèles, d'autres avec un seul. La plupart des QTLs détectés avec les modèles « Allèles fondateurs » et « SNP intra-groupe » sont significatifs au sein d'un seul des groupes hétérotiques. Quel que soit le modèle, la plupart des QTLs présentent de petits effets (expliquant moins de 5% de la variation à l'exception du QTL pour la hauteur de plante détecté sur le chromosome 3 par le modèle « SNP intra-groupe » et du QTL détectés à 44.5 cM sur le chromosome 10 par les 3 modèles. Ce QTL explique environ 8% de la variance pour la matière sèche et 13% pour la floraison femelle. Il correspond au gène de précocité *ZmmCCT*. 12 QTLs présentent des effets de dominance/ASC significatifs avec un risque individuel de 5% et sont localisés tout le long du génome. Les QTLs détectés avec le modèle « Allèles fondateurs » pour le rendement, le contenu en matière sèche et la hauteur de plante sont ceux expliquant le plus de variance phénotypique. Pour la floraison femelle, ce sont ceux détectés avec le modèle « SNP intra-groupe ». Considérant le meilleur modèle pour chaque caractère, les QTLs détectés expliquent entre 30.7 % (hauteur de plante, avec ASC) et 37.6 % (floraison femelle, avec ASC) de la variance phénotypique.

Ce dispositif a donc permis de détecter des QTLs conjointement dans les deux groupes hétérotiques, sans utiliser de testeur. Chaque hybride apporte donc de l'information sur les deux groupes, ce qui à moyens de phénotypage constants permet d'évaluer deux fois plus de lignées. Des QTLs ont été détectés pour tous les caractères avec tous les modèles et, comme dans l'étude réalisée sur testeur, nous avons mis en évidence des QTLs spécifiques de l'AGC cornée et de l'AGC dentée. De l'ASC a été observée pour tous les caractères mais peu de QTLs ayant un effet sur l'ASC ont été détectés. On peut faire l'hypothèse que l'ASC est due à de nombreux QTLs à faibles effets, à de l'épistasie etc. En évitant l'utilisation de testeurs, on peut supposer que notre dispositif conduit à de meilleures estimations des effets d'AGC des QTLs.

La sélection génomique peut-elle modifier profondément les premières étapes d'un programme de sélection pour la valeur hybride ? Cas des performances ensilage de maïs (*Zea mays* L.) d'hybrides obtenus dans un croisement de deux dispositifs multiparentaux.

Le dispositif étudié est le même que celui présenté dans la partie précédente. L'objectif de cette partie était de chercher à prédire dans un tel dispositif les composantes d'AGC et d'ASC de la valeur hybride chez le maïs. En effet, avec la technique d'haplo-diploïdisation il est maintenant possible d'obtenir directement de larges population de lignées en ségrégation ce qui offre de nouvelles possibilités pour revisiter les schémas de sélection pour les hybrides. Au lieu d'utiliser des testeurs, on peut directement explorer une large gamme de combinaisons de lignées parentales en créant directement les hybrides entre lignées non encore sélectionnées. Ces données peuvent être ensuite utilisées pour calibrer des équations de prédiction de la valeur hybride et identifier les meilleures combinaisons. Le même modèle peut être utilisé pour prédire les AGC de toutes les lignées candidates et sélectionner les plus prometteuses pour un nouveau cycle de sélection.

Les prédictions génomiques ont été effectuées en utilisant des modèles de type GBLUP (Genomic Unbiased Linear Prediction) (Massman *et al.* 2013 ; Technow *et al.* 2014). Le modèle complet considère la structure en populations en effets fixes ainsi que les AGC et l'ASC. Pour la covariance entre les effets d'AGC, nous avons estimé l'apparentement entre lignées de chaque groupe à partir des informations de marquage selon la méthode 1 de Van Raden (2008). La covariance entre les effets d'ASC de deux hybrides a été calculée à partir du produit des apparentements intra-groupe des lignées parentales des deux hybrides (en faisant l'hypothèse d'une absence d'apparentement inter-groupe). Des variantes du modèle présentant ou non la structure en populations et/ou la composante d'ASC ont été testées ainsi que des modèles considérant des matrices d'apparentement basées sur les allèles fondateurs et non sur les SNP observés. La qualité de prédiction des différents modèles a été évaluée par validation croisée en utilisant quatre cinquième du jeu de données pour la calibration et un cinquième pour la validation et en répétant 100 fois l'échantillonnage. La qualité de prédiction a été estimée via la corrélation valeurs observées / valeurs prédites sur le jeu de validation. En considérant pour chaque caractère le meilleur modèle, la corrélation entre données observées et données prédites varie entre 0.652 pour le rendement et 0.771 pour la hauteur de plante. Le modèle permettant la meilleure qualité de prédiction est celui ne prenant pas en compte la structure en populations et considérant les matrices d'apparentement basées sur les SNPs observés. La prise en compte de l'ASC dans le modèle a un très faible impact. Les prédictions génomiques sont meilleures

que celles obtenues sur la base des QTLs dans la partie précédente.
L'effet de la taille du jeu de calibration a également été étudié. Pour tous les caractères la qualité des prédictions augmente avec la taille du jeu de calibration jusqu'à atteindre un plateau quand le jeu de calibration contient 12 individus par population dentée-cornée (432 hybrides). L'impact de la composition du jeu de calibration a également été étudié : impact de la présence de demi-frères, impact de la présence d'hybrides issus des mêmes lignées fondatrices. La qualité de prédiction des AGCs dentées et cornées a été étudiée. La précision observée sur leurs prédictions est élevée et varie entre 0.739 (floraison femelle) et 0.827 (matière sèche) pour les AGCs dentées et entre 0.830 (rendement) et 0.938 (floraison femelle) pour les AGCs cornées. Les prédictions génomiques appliqués dans notre dispositif ont donc permis de prédire précisément les composantes d'AGCs des lignées parents des hybrides malgré le faible nombre de contribution de chaque lignée au factoriel. Ceci ouvre d'intéressantes perspectives pour revisiter les schémas de sélection hybride.

Discussion générale

La compréhension de l'architecture génétique de la valeur hybride chez le maïs est importante tant d'un point de vue théorique que appliqué. Dans cette thèse deux dispositifs ont été étudiés mettant en jeu des hybrides : le premier des hybrides sur testeur, le second des hybrides issus d'un croisement factoriel entre deux dispositifs connectés multiparentaux. Des QTLs ont été détectés dans les deux dispositifs et des modèles de prédiction génomique ont été mis en œuvre dans le second.
Un aspect important non pris en compte dans cette thèse tant pour la détection de QTL que la sélection génomique est la prise en compte des interactions GxE. De nombreuses applications et perspectives existent, principalement concernant le deuxième dispositif, pour lequel des caractères liés à la qualité de l'ensilage ont également été phénotypés. Des hybrides issus de ce dispositif font actuellement l'objet d'évaluation pour être inscrits au Catalogue Français des Variétés. Une amélioration des lignées parentales par sélection assistée par marqueurs peut également être envisagée. D'un point de vue économique, notre dispositif présente des avantages par rapport aux dispositifs sur testeur quant à l'évaluation des lignées parentales, puisqu'il permet pour un nombre donné de lignées de chaque groupe de diminuer par un facteur deux le nombre d'hybrides évalués agronomiquement. Cependant il engendrerait un changement des pratiques des sélectionneurs. Des recherches complémentaires sont nécessaires pour évaluer l'intérêt de notre dispositif par rapport à un dispositif sur testeur et pour évaluer

l'impact de modifications du dispositif expérimental (nombre de lignées par famille, nombre d'hybrides évalués par lignées…) sur la précision de prédiction des AGCs et des ASCs.

Références bibliographiques

Bardol, N., M. Ventelon, B. Mangin, S. Jasson, V. Loywick *et al.*, 2013 Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. Theor. Appl. Genet. 126: 2717–2736.

Browning, B. L., and S. R. Browning, 2009 A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84: 210–223.

Butler, D. B. R. Cullis, A. R. Gilmour, and B. J. Gogel, 2007 ASReml-R reference manual. The State of Queensland, Department of Primary Industries and Fisheries.

Charcosset, A., B. Mangin, L. Moreau, L. Combes, M.-F. Jourjon *et al.*, 2000 Heterosis in maize investigated using connected RIL populations, pp. 89-98 in *Quantitative Genetics and Breeding Methods: The Way Ahead*, edited by A. Gallais, C. Dillman, and I. Goldringer. INRA, Paris, France.

Charcosset, A., 2002 Le fait hybride, conditions de l'innovation et choix stratégiques. "Colloque l'amélioration des plantes, continuités et ruptures", Montpellier, France, October 2002.

Darwin, C., 1876 The effects of cross and self fertilization in the vegetable kingdom. John Murray, London.

East, E. M., 1908 Inbreeding in corn, pp.419-428 in *Reports of the Connecticut Agricultural Experimental Station for years 1907-1908*. Connecticut Agricultural Experiment Station, New Haven.

Ganal, M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler *et al.*, 2011 A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS ONE 6: e28334.

Gao, X., J. Starmer, and E. R. Martin, 2008 A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet. Epidemiol. 32: 361-369.

Giraud, H., C. Lehermeier, E. Bauer, M. Falque, V. Segura *et al.*, 2014 Linkage disequilibrium with linkage analysis of multiline crosses reveals different multiallelic QTL for hybrid performance in the flint and dent heterotic groups of maize. Genetics 198: 1717-1734.

Hickey, J. M., G. Gorjanc, R. K. Varshney, and C. Nettelblad, 2015 Imputation of single nucleotide polymorphism genotypes in biparental, backcross, and topcross populations with a hidden Markov model. Crop Sci. 55: 1934-1946.

Jourjon, M.-F., S. Jasson, J. Marcel, B. Ngom, and B. Mangin, 2005 MCQTL: multi-allelic QTL mapping in multi-cross design. Bioinformatics 21: 128–130.

Leroux, D., A. Rahmani, S. Jasson, M. Ventelon, F. Louis *et al.*, 2014 Clusthaplo: a plug-in for MCQTL to enhance QTL detection using ancestral alleles in multi-cross design. Theor. Appl. Genet. 127: 921-933.

Massman, J. M., A. Gordillo, R. E. Lorenzana, and R. Bernardo, 2013 Genomewide predictions from maize single-cross data. Theor. Appl. Genet. 126: 13-22.

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

Paterson, A. H., E. S. Lander, J. D. Hewitt, S. Peterson, S. E. Lincoln *et al.*, 1988 Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335: 721-726.

R Core Team, 2013 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Segura, V., B. J. Vilhjálmsson, A. Platt, A. Korte, U. Seren *et al.*, 2012 An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations.

Nat. Genet. 44:825-830.

Shull, G. H., 1908 The composition of a field of maize. Am. Breed. Assoc. Rep. 5: 51-59.

Shull, G. H., 1914 Duplicate genes for capsule-form in *Bursa bursa pastoris*. Zeitschrift ind. Abst. U. Verebsgl. 12: 97-149.

Sprague, G. F., and L. A. Tatum, 1942 General *vs.* specific combining ability in single crosses of corn. J. Am. Soc. Agron. 34: 923-932.

Technow, F., T. A. Schrag, W. Schipprack, E. Bauer, H. Simianer *et al.*, 2014 Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. Genetics 197: 1343-1355.

Tenaillon, M. I., and A. Charcosset, 2011 A European perspective on maize history. C. R. Biol. 334: 221–228.

Van Raden, P. M., 2008 Efficient methods to compute genomic predictions. J. Dairy Sci. 91: 4414-4423.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Agriculture, alimentation, biologie, environnement, santé (ABIES)

Titre : Analyse génétique de la valeur hybride chez le maïs fourrage dans des dispositifs multiparentaux : détection de QTL et sélection génomique

Mots clés : maïs, valeur hybride, détection de QTL, sélection génomique, ensilage

Résumé : Le maïs (*Zea mays* L.) est une culture majeure présentant un hétérosis important pour les caractères liés à la biomasse. Pour exploiter ce phénomène, les programmes de sélection ont été organisés en groupes hétérotiques et la majorité des surfaces cultivées correspond à des hybrides F1. La valeur hybride peut être décomposée en Aptitudes Générales à la Combinaison (AGC) des lignées parentales et en Aptitude Spécifique à la Combinaison (ASC).
L'objectif de cette thèse était d'apporter de nouveaux éclairages sur la valeur hybride, concernant tant la compréhension de ses déterminismes génétiques sous-jacents que sa prédiction. Deux dispositifs multiparentaux connectés ont été analysés, mettant en jeu deux des principaux groupes hétérotiques utilisés pour la production de maïs fourrage dans le nord de l'Europe : les cornés et les dentés. Le premier dispositif était constitué de deux populations de type Nested Association Mapping (NAM). La population dentée comprenait dix familles biparentales et la cornée onze. Ces populations ont été évaluées pour leur valeur en croisement avec un testeur du groupe complémentaire. Le deuxième dispositif était constitué d'un factoriel entre deux populations multiparentales de lignées : une cornée et une denté, dérivées de deux demi-diallèles entre quatre lignées fondatrices. Les lignées ont été croisées entre elles selon un plan factoriel incomplet afin d'obtenir des hybrides inter-groupes. Pour les deux dispositifs, le phénotypage des hybrides a porté sur le rendement ensilage, le contenu en matière sèche, la date de floraison femelle et la hauteur de plante. Les hybrides des dispositifs NAM ont également été phénotypés pour la date de floraison mâle.
Une détection des locus impliqués dans la variation des caractères quantitatifs (QTL) a été mise en œuvre en utilisant des modèles prenant en compte différents codages alléliques : allèles fondateurs haplotypes ou allèle observé au marqueur. Ces codages ont permis de mettre en œuvre des modèles de types LA (Linkage Analysis) ou LDLA (Linkage Disequilibrium - Linkage Analysis). Une complémentarité des modèles a été observée. Certains modèles ont permis la détection de QTLs multi-alléliques. Des QTLs différents ont été détectés dans les deux groupes hétérotiques, confirmant leur divergence ancienne. Pour le contenu en matière sèche et le rendement, nous n'avons pas détecté de QTL à effet majeur dans aucun des dispositifs à l'exception d'un QTL corné de floraison à effet pléiotropique dans le groupe corné. Bien que l'ASC représentait 20% de la variance génétique intra-population totale pour la plupart des caractères (sauf la hauteur de plante) le dispositif factoriel n'a permis la détection que d'un faible nombre de QTLs ayant un effet sur la dominance/ l'ASC.
Différents modèles de sélection génomique de type GBLUP ont été mis œuvre dans le dispositif factoriel. Nous avons obtenu de bonnes qualités de prédictions, bien que la majorité des lignées ne fût parentes que d'un seul hybride. Nous avons montré qu'il était important de calibrer les prédictions sur des hybrides issus de lignées apparentées aux hybrides à prédire. Nos modèles ont montré leurs limites quant à la capacité de prédiction de l'ASC mais ont permis d'obtenir de bonnes qualités de prédiction pour les AGC des lignées parentales. Notre étude ouvre de nouvelles perspectives pour reconsidérer les schémas de sélection des hybrides de maïs. Les évaluations des lignées candidates sur testeur pourraient être remplacées par une évaluation d'hybrides obtenus selon un plan de croisement factoriel incomplet entre groupes hétérotiques. Des évaluations complémentaires sont nécessaires pour comparer plus directement notre stratégie avec celle basée sur l'utilisation de testeurs. Cependant des premiers résultats encourageants ont été obtenus.

Title : Genetic analysis of hybrid value for silage maize in multiparental designs: QTL detection and genomic selection

Keywords : maize, hybrid value, QTL detection, genomic selection, silage

Abstract : Maize (*Zea mays* L.) is a major crop presenting strong heterosis for traits linked to biomass. To exploit this phenomenon, F1 hybrids are cultivated and breeding programs are organized in heterotic groups. Hybrid value can be decomposed in General Combining Abilities (GCA) of the parental lines and Specific Combining Ability (SCA).
The goal of this thesis was to bring new insights to the understanding and the prediction of hybrid value and its components for silage maize. To do so, two multiparental connected designs, involving hybrids between the dent and flint heterotic groups, main heterotic groups used for silage breeding in Northern Europe, were analyzed. The first design consisted in two Nested Association Mapping (NAM) populations involving test-cross hybrids. The dent population consisted of ten biparental dent families and the flint one of 11 biparental families. The second design consisted of two multiparental populations of inbred lines, one dent and one flint, derived from two half diallels between four founder lines. Inbred lines from both groups were crossed according to an incomplete factorial to produce experimental hybrids. Hybrids were phenotyped for dry matter yield, dry matter content, female flowering date, plant height, and only for the NAM designs male flowering date.
Quantitative Trait Loci (QTL) detection using Linkage Analysis (LA) and Linkage Disequilibrium – Linkage Analysis (LDLA) methods was performed in both designs. QTL detection models were found to be complementary. Some of the QTL detection models allowed us to detect multiallelic series at the QTLs, showing the importance of not directly considering the observed biallelic genotype. Different QTLs were detected in the two heterotic groups, highlighting their long-term divergence. No QTL showed major effect for DMC and DMY except one pleiotropic QTL for flowering time in the flint group. The reciprocal multiparental design allowed the detection of QTLs for dominance/SCA. However, very few QTLs had significant dominance/SCA effects even if SCA represented around 20% of the within-genetic variance, except for PH for which it was lower.
Genomic selection was conducted in the second design using different GBLUP models. We obtained good predictive abilities even though predicted hybrids were obtained through a highly incomplete factorial mating where the majority of the inbred lines were parents of only one hybrid. Variation within populations sharing three founders with populations used for calibration could be predicted efficiently. Our models showed a limited ability for predicting SCA but predicted well the GCA of the parental lines. This opens new prospects for reconsidering maize breeding. Indeed test-cross evaluation, usually used in the breeding companies, could be replaced by single-cross evaluation, according to an incomplete factorial design, between the two heterotic groups to improve. Further evaluations are needed for comparing the efficiency of this strategy with the one of tester-based designs, but first obtained results are encouraging.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France