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

Gene mining in doubled haploid lines from European maize landraces with association mapping

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

von
M. Sc. der Agrarwissenschaften
Alexander Carl Georg Strigens
aus Wiesbaden

Stuttgart–Hohenheim
2014

Die vorliegende Doktorarbeit wurde am 2. Februar 2014 von der Fakultät Agrarwissenschaften der Universität Hohenheim als „Dissertation zur Erlangung des Grades eines Doktors der Agrarwissenschaften (Dr. sc. agr.)“ angenommen.

Tag der mündlichen Prüfung: 7. Juli 2014

1. Prodekan:	Prof. Dr. M. Rodehutscord
Berichterstatter, 1. Prüfer:	Prof. Dr. A. E. Melchinger
Mitberichterstatter, 2. Prüfer:	Prof. Dr. J. C. Reif
3. Prüfer	Prof Dr. Zebitz

CONTENTS

[1] Strigens, A., C. Grieder, B.I. Haussmann, and A.E. Melchinger. 2012. Genetic variation among inbred lines and testcrosses of maize for early growth parameters and their relationship to final dry matter yield. Crop Science 52: 1084–1092.

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

[3] Strigens, A., W. Schipprack, J.C. Reif, and A.E. Melchinger. 2013b. Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding. PloS one 8: e57234.

ABBREVIATIONS

Locations

EWE	Eckartsweier
HOH	Hohenheim high N
HOL	Hohenheim low N
KLH	Kleinhohenheim
OLI	Oberer Lindenhof

Populations

BU	Bugard
GB	Gelber Badischer
SC	Schindelmeiser
EU-F	European elite Flint
EU-D	European elite Dent
NA-D	North-American Dent

Other

BLUE	Best linear unbiased estimation
BLUP	Best linear unbiased prediction
DH	Doubled haploid
GDD	Growing day degrees
GWA	Genome wide association
LD	Linkage disequilibrium
LR	Landrace
MAF	Minor allele frequency
MAS	Marker assisted selection
N_e	Effective population size
PCoA	Principal coordinate analysis
QTL	Quantitative trait loci
SNP	Single nucleotide polymorphism

Traits

ASIN	Anthesis-silking interval
BAST	Barren stalks
CHLO	Leaf chlorosis score
EADI	Ear diameter
EAHT	Ear insertion height
EALE	Ear length
EASH	Ear shank score
EDMC	Ear dry matter content
$EFMA_X$	Early fresh mass at X-leaf stage
EMER	Emergence score
$EPHT_X$	Early plant height at X-leaf stage
$EVIG_X$	Early vigor at X-leaf stage
FFLO	Female flowering
GERM	Germination
GRYD	Kernel yield
HUCO	Husk coverage score
HUFL	Husk flag leaves score
IFUS	Ear rot incidence
KERO	Kernels by row
KOIL	Kernel oil content
LODG	Lodging
$MAPL_{8;OLI}$	Fresh mass per plant at 8-leaf stage in OLI
MFLO	Male flowering
PLHT	Plant height at maturity
REGR	Relative growth rate (mean over locations)
$REGR_Y$	Relative growth rate at location Y
ROWS	Kernel rows
SFUS	Ear rot severity
SMUT	Common smut
$TFMA_{8;OLI}$	Total fresh mass per plot at 8-leaf stage in OLI
THKW	Thousand kernel weight

Chapter 1

General Introduction

Importance of maize cultivation and its challenges

Maize is one of the three most important crops cultivated for human nutrition together with rice and wheat. In 2011, maize production covered a total of 170 million hectares worldwide, producing 883 million metric tons of grain, while 704 million metric tons of wheat were produced on 220 million hectares (FAOSTAT, 2013). Germany, and generally north-western Europe, where long considered as areas with only marginal potential for maize growing due to the cold sensitivity of maize. However, the dramatic increase in maize production in Germany over the last decades (DMK 2012), shows how breeding and new cultivation practices can lead to the adaptation of a crop to new areas. This expansion of maize cultivation to northern latitudes was achieved by the development of varieties of maize able to cope with the cool temperatures and high humidity of those climates (Frei, 2000) and by the extensive use of maize for silage production (DMK, 2012). To further improve the productivity and yield stability of the species, continuous efforts have to be undertaken to increase their tolerance to abiotic (*e.g.*, heat, drought, chilling) and biotic (*e.g.*, insects, fungi) stresses. A further challenge of maize production will certainly also be the maintenance of high productivity with reduced fertilizer input, because prizes for nitrogen and phosphor fertilizers are increasing and a continuation of this trend can be anticipated (World Bank, 2013).

Maize landraces as genetic resource

Since its domestication, maize has been shaped by farmers selecting preferred plants for the next growing season. Over the centuries, this resulted in a broad diversity of open-pollinated maize populations adapted to the farmer's preferences and needs. Through the ongoing natural selection, these so called landraces became at the same time well adapted to the local climatic and edaphic conditions. Since the introduction of Tropical Flint maize into southern Europe by Colombus in 1492 and of Northern Flint into north-western Europe by further discoverer of the 16th century (Rebourg et al., 2003), open-pollinated varieties were also cultivated and selected by farmers across the European continent. Over the centuries, the hybridization of landraces from the southern and northern Flint introductions in the Pyrenean region resulted in a completely new genetic pool: the European Flint (Tenaillon and Charcosset, 2011). In parallel, the originally rather cold sensitive maize got adapted to the cool and wet climate of Europe, allowing its cultivation even north of the Alps. This resulted in a broad diversity of European Flint landraces with a unique genetic composition and specific adaptation.

Because landraces were developed before chemical pesticides and mineral fertilizers were available and widely used, it is expected that the landraces harbor numerous genes or alleles for abiotic stress tolerance and pest resistance (Lafitte, 1997; Hoisington et al., 1999; Malvar et al., 2004, 2007; Warburton et al., 2008; Peter et al., 2009a; b). However, with the advent of hybrid breeding (Shull, 1908), hybrid varieties exploiting heterosis more optimally gradually replaced landraces in the U.S.A. in the 1930s' (Crow, 1998). The superior yield, uniformity and stability of hybrids were key factors for their success in the developing mechanized agriculture of that time (Barrière et al., 2006). Since the 1950s', the well adapted European landraces were also replaced by hybrid varieties exploiting the strong heterosis observed between the U.S. Corn Belt Dent and European Flint heterotic groups (Gouesnard et al., 2005;

Reif et al., 2005; Tenaillon and Charcosset, 2011). The development of inbred lines from several European Flint landraces significantly contributed to this success, but the genetic diversity captured in these first-cycle inbred lines was just a fraction of the available diversity (Messmer et al., 1992; Reif et al., 2005)

Fortunately, the value of landraces as genetic resources was recognized before their extinction. They were collected at their growing locations and are being conserved *ex situ* in gene banks. Thus, alleles for abiotic stress tolerance and pest resistance needed to further improve maize productivity and yield stability might still be found in the large collections of landraces accessions (~50,000) stored in gene banks around the world (Hoisington et al., 1999). The European landraces might especially be of great interest to improve the European elite material, due to their specific adaptation to the cool and wet climate prevailing in Europe (Reif et al., 2005; Peter et al., 2009a; b; Tenaillon and Charcosset, 2011).

Evaluation and characterization of the European landraces stored in the gene banks was performed to classify the collected material and to identify interesting properties that might be introduced in the elite material (for a review see Gouesnard et al., 2005). Landraces with superior cold tolerance (Revilla et al., 1998, 2006; Rodríguez et al., 2007, 2010; Peter et al., 2009a; b; Schneider et al., 2011), pest resistance (Malvar et al., 2004, 2007) and digestibility (Barrière et al., 2010) could be identified. Genetic analyses of this material further showed the huge genetic diversity present in these landraces in comparison with elite breeding material (Gauthier et al., 2002; Reif et al., 2005; Eschholz et al., 2008).

The limitations of landraces for breeding

Even though landraces appear to be very valuable genetic resources for broadening the genetic base of elite material as well as for the mining of new properties, their use in breeding remained so far limited (Hoisington et al., 1999). This can be attributed to the heterogeneous

nature of these open-pollinated populations combined with the presence of unfavorable traits and detrimental alleles, the so called genetic load, in this unselected material. The first hampers a precise evaluation of the landraces, because completely new and unique heterozygous individuals are produced at each generation and cannot be reproduced for evaluation in different environments. It further complicates the removal of the second by mass selection, because recessive alleles remain hidden at heterozygous loci. Inbreeding, as done for the development of the parents of the first hybrids, enables to remove these recessive alleles from the landraces (Crnokrak and Barrett, 2002). However, this is a very tedious work, because of the strong inbreeding depression and because lethal recessive alleles might still be uncovered in very advanced selfing generations, ruining the efforts of the breeders (Schnell, 1959). Additionnaly, unwanted properties tightly associated with the desirable ones might reduce the breeding value of the developed inbred lines, because negative properties will unintentionally be introduced into the breeding germplasm by linkage drag.

Use of the DH technique to unlock the diversity of landraces

To get a more efficient and rapid access to the genetic diversity harbored in landraces, Reif et al. (2005) proposed the use of the doubled haploid (DH) technique to produce DH lines out of the landraces. This method takes advantage of the aptitude of specific inbred lines, so called inducers, to produce haploid embryos when used as pollinators (Coe, 1959; Eder and Chalyk, 2002; Röber et al., 2005; Prigge and Melchinger, 2012). A still unknown mechanism (either chromosome elimination or parthogenesis) leads to the development of haploid embryos. These haploid plants are generally male sterile (Coe, 1959; Coe and Sarkar, 1964; Kleiber et al., 2012) and an artificial chromosome doubling is necessary to obtain male fertile DH lines. The alkaloid Colchicine is commonly used for chromosome doubling. It blocks the building of microtubuli and, thus, the separation of the sister chromatids during the anaphase of mitosis, resulting in undivided cells with a doubled amount of DNA (Deimling et al., 1997).

As a consequence, DH plants are perfectly homozygous samples of the maternal gametes. Besides all the advantages of obtaining fixed inbred lines within one step instead of repeated selfings for 7 generations (Geiger and Gordillo, 2010), it was postulated that the genetic load present in the induced material might be purged by the DH technique (Reif et al., 2005; Prigge et al., 2012). Parts of the lethal recessive alleles are expected to be expressed and lead to mortality at the haploid stage (Charlesworth and Charlesworth, 1992).

Producing DH lines from landraces would, thus, overcome the drawbacks limiting the use of landraces as genetic resources. Ideally it should allow (i) fixing of the complete genetic diversity present in the landraces, (ii) *ad libitum* multiplication of the genetic material without any genetic drift, (iii) precise evaluation of the phenotypic diversity present in landraces in replicated multilocation trials, and (iv) reducing the genetic load present in landraces.

Identifying new alleles by genome wide association mapping

A broad set of DH lines derived from various landraces is, therefore, a formidable mine of genetic diversity. Because no artificial selection was performed on this material, large phenotypic and genotypic variances can be expected. New advantageous properties might be identified in this material. Further, the possibility to perform replicated trials allows estimating variance components and trait heritability, and, thus, quantifying the selection gains that can be expected from the introgression of the identified traits into the elite germplasm.

Genotyping of such libraries of DH lines derived from landraces with the recently developed high throughput and high density single nucleotide polymorphism (SNP) marker platforms yielding thousands of marker points (Ganal et al., 2011) would give a very deep insight in the molecular diversity of the landraces. It would allow very precise estimation of genetic diversity, kinship and population structure (Eding and Meuwissen, 2001). It might further

allow determining the effect of the DH method on gamete sampling and purging of lethal recessive alleles as well as estimating the effective population size of the landraces.

Because low linkage disequilibrium (LD) was observed in European landraces (Reif et al., 2005; Tenaillon and Charcosset, 2011), a similarly low LD can be expected in DH lines derived from landraces. Combined with a large phenotypic and genetic diversity, as well as the availability of dense marker coverage, this makes such libraries a perfect tool for high resolution genome wide association (GWA) mapping approaches (Yu et al., 2006; Stich et al., 2008). Association mapping exploits the historical linkage between genetic markers and causative genes in diverse populations, allowing the precise identification of quantitative trait loci (QTL) and underlying candidate genes. This allows targeted introgression of desired traits from the landraces into elite breeding material, without introducing unwanted properties by linkage drag. Further, it gives insights in the genetic architecture underlying trait expression, allowing deeper understanding of physiological and metabolic pathways (Riedelsheimer et al., 2012).

Objectives of this study

The goal of this research was to use the advantages of the DH technique to unlock the diversity of European Flint landraces and mine for new genes and alleles by GWA mapping in the DH lines derived from landraces. A strong focus was put on early growth and cold tolerance, because adaptation to the cool and wet climate of Europe is one of the most important features and contribution to elite material of the European Flint landraces. In particular, the objectives were to

(1) develop a robust method to quantify early growth with a non-destructive remote sensing platform developed at the University of Hohenheim (Montes et al., 2011),

(2) evaluate the importance of *per se* early growth performance of inbred lines with regard to their early growth and yield performance in testcrosses,

(3) determine the potential of GWA mapping to identify genes and alleles underlying early growth and cold tolerance related traits under controlled and field conditions,

(4) evaluate the phenotypic and genotypic diversity recovered in 132 DH lines derived from the European Flint landraces *Bugard*, *Gelber Badischer* and *Schindelmeiser* for morphological and agronomic traits in comparison with a set of elite flint inbred lines,

(5) estimate the effect of the DH method on the recovered genetic diversity and of an eventual purging of lethal recessive alleles from the landraces by comparing the original landraces with synthetic landraces obtained from the recombination of the respective DH lines.

(6) perform gene mining by GWA mapping in a panel of DH lines derived from landraces together with elite Flint and elite Dent inbred lines to identify new genes or alleles underlying morphological and agronomical properties,

(7) discuss the potential of DH lines derived from landraces to perform gene mining and improve the genetic diversity and performance of current elite European Flint breeding germplasm.

REFERENCES

Barrière, Y., D. Alber, O. Dolstra, C. Lapierre, M. Motto, A. Ordás, J. Van Waes, L. Vlasminkel, C. Welcker, and J.P. Monod. 2006. Past and prospects of forage maize breeding in Europe. II. History, germplasm evolution and correlative agronomic changes. Maydica 51: 435–449.

Barrière, Y., A. Charcosset, D. Denoue, D. Madur, C. Bauland, and J. Laborde. 2010. Genetic variation for lignin content and cell wall digestibility in early maize lines derived from ancient landraces. Maydica 55: 65–74.

Charlesworth, D., and B. Charlesworth. 1992. The effects of selection in the gametophyte stage on mutational load. Evolution 46: 703–720.

Coe, E.H. 1959. A line of maize with high haploid frequency. The American Naturalist 93: 381–382.

Coe, E.H., and K.R. Sarkar. 1964. The detection of haploids in maize. Journal of Heredity 55: 231–233.

Crnokrak, P., and S.C.H. Barrett. 2002. Perspective: purging the genetic load: a review of the experimental evidence. Evolution 56: 2347–2358.

Crow, J.F. 1998. 90 Years Ago : The Beginning of Hybrid Maize. Genetics 148: 923–928

Deimling, S., F.K. Röber, and H.H. Geiger. 1997. Methodik und Genetik der Haploiden-Induktion bei Mais. Vortr. Pflanzenzüchtung 38: 203–224.

Deutsches Maiskomitee e.V. (DMK). 2012. DMK-Geschäftsbericht 2011/2012. DMK, Bonn.

Eder, J., and S. Chalyk. 2002. In vivo haploid induction in maize. Theoretical and Applied Genetics 104: 703–708.

Eding, H., and T.H.E. Meuwissen. 2001. Marker based estimates of between and within population kinships for the conservation of genetic diversity. Journal of Animal Breeding and Genetics 118: 141–159.

Eschholz, T.W., R. Peter, P. Stamp, and A. Hund. 2008. Genetic diversity of Swiss maize (*Zea mays* L. ssp. *mays*) assessed with individuals and bulks on agarose gels. Genetic Resources and Crop Evolution 55: 971–983.

FAOSTAT, 2013. Available at http://faostat3.fao.org/home/index.html

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

Ganal, M.W., G. Durstewitz, A. Polley, A. Bérard, E.S. Buckler, A. Charcosset, J.D. Clarke, E.-M. Graner, M. Hansen, J. Joets, M.-C. Le Paslier, M.D. McMullen, P. Montalent, M. Rose, C.-C. Schön, Q. Sun, H. Walter, O.C. Martin, and M. Falque. 2011. A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PloS one 6: e28334.

Gauthier, P., B. Gouesnard, J. Dallard, R. Redaelli, C. Rebourg, A. Charcosset, and A. Boyat. 2002. RFLP diversity and relationships among traditional European maize populations. Theoretical and Applied Genetics 105: 91–99.

Geiger, H.H., and G.A. Gordillo. 2010. Doubled haploids in hybrid maize breeding. 2010. Maydica 54: 485–499.

Gouesnard, B., J. Dallard, P. Bertin, A. Boyat, and A. Charcosset. 2005. European maize landraces: genetic diversity, core collection definition and methodology of use. Maydica 50: 115–234.

Hoisington, D., M. Khairallah, T. Reeves, J.-M. Ribaut, B. Skovmand, S. Taba, and M.L. Warburton. 1999. Plant genetic resources: what can they contribute toward increased crop productivity? Proceedings of the National Academy of Sciences 96: 5937–5943.

Kleiber, D., V. Prigge, A.E. Melchinger, F. Burkard, F. San Vicente, G. Palomino, and G.A. Gordillo. 2012. Haploid Fertility in Temperate and Tropical Maize Germplasm. Crop Science 52: 623-630.

Lafitte, H. 1997. Adaptive strategies identified among tropical maize landraces for nitrogen-limited environments. Field Crops Research 49: 187–204.

Malvar, R.A., A. Butrón, A. Álvarez, B. Ordás, P. Soengas, P. Revilla, and A. Ordás. 2004. Evaluation of the European Union maize landrace core collection for resistance to Sesamia nonagrioides (*Lepidoptera*: *Noctuidae*) and Ostrinia nubilalis (*Lepidoptera*: *Crambidae*). Journal of Economic Entomology 97: 628–634.

Malvar, R.A., A. Butrón, A. Álvarez, G. Padilla, M. Cartea, P. Revilla, and A. Ordás. 2007. Yield performance of the European Union Maize Landrace Core Collection under multiple corn borer infestations. Crop Protection 26: 775–781.

Messmer, M., A.E. Melchinger, J. Boppenmaier, R.G. Herrmann, and E. Brunklaus-Jung. 1992. RFLP analyses of early-maturing European maize germ plasm I . Genetic diversity among flint and dent inbreds. Theoretical and Applied Genetics 83: 1003–1012.

Montes, J.M., F. Technow, B.S. Dhillon, F. Mauch, and A.E. Melchinger. 2011. High-throughput non-destructive biomass determination during early plant development in maize under field conditions. Field Crops Research 121: 268–273.

Peter, R., T.W. Eschholz, P. Stamp, and M. Liedgens. 2009a. Early growth of flint maize landraces under cool conditions. Crop Science 49: 169–178.

Peter, R., T.W. Eschholz, P. Stamp, and M. Liedgens. 2009b. Swiss Flint maize landraces—A rich pool of variability for early vigour in cool environments. Field Crops Research 110: 157–166.

Prigge, V., R. Babu, B. Das, M.H. Rodriguez, G.N. Atlin, and A.E. Melchinger. 2012. Doubled haploids in tropical maize: II. Quantitative genetic parameters for testcross performance. Euphytica 185: 453–463.

Prigge, V., and A.E. Melchinger. 2012. Production of haploids and doubled haploids in maize. *In* Loyola-Vargas, V., Ochoa-Alejo, N. (eds.), Plant cell culture protocols. 3rd ed. Humana Press - Springer Verlag, Totowa, New Jersey.

Rebourg, C., M. Chastanet, B. Gouesnard, C. Welcker, P. Dubreuil, and A. Charcosset. 2003. Maize introduction into Europe: the history reviewed in the light of molecular data. Theoretical and Applied Genetics 106: 895–903.

Reif, J.C., S. Hamrit, M. Heckenberger, W. Schipprack, H. Peter Maurer, M. Bohn, and A.E. Melchinger. 2005. Genetic structure and diversity of European flint maize populations determined with SSR analyses of individuals and bulks. Theoretical and Applied Genetics 111: 906–913.

Revilla, P., A. Boyat, A. Álvarez, B. Gouesnard, B. Ordás, V.M. Rodríguez, A. Ordás, and R.A. Malvar. 2006. Contribution of autochthonous maize populations for adaptation to European conditions. Euphytica 152: 275–282.

Revilla, P., R.A. Malvar, M. Cartea, and A. Ordás. 1998. Identifying open-pollinated populations of field corn as sources of cold tolerance for improving sweet corn. Euphytica 101: 239–247.

Riedelsheimer, C., J. Lisec, A. Czedik-Eysenberg, R. Sulpice, A. Flis, C. Grieder, T. Altmann, M. Stitt, L. Willmitzer, and A.E. Melchinger. 2012. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. Proceedings of the National Academy of Sciences.

Röber, F.K., G.A. Gordillo, and H.H. Geiger. 2005. In vivo haploid induction in maize-performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 50: 275–283.

Rodríguez, V.M., R.A. Malvar, A. Butrón, A. Ordás, and P. Revilla. 2007. Maize Populations as Sources of Favorable Alleles to Improve Cold-Tolerant Hybrids. Crop Science 47: 1779.

Rodríguez, V.M., M.C. Romay, A. Ordás, and P. Revilla. 2010. Evaluation of European maize (*Zea mays* L.) germplasm under cold conditions. Genetic Resources and Crop Evolution 57: 329–335.

Schneider, D.N., N.M. Freitag, M. Liedgens, B. Feil, and P. Stamp. 2011. Early growth of field-grown swiss flint maize landraces. Maydica 56: 1702.

Schnell, F.W. 1959. Mais. p. 140—141. *In* Rudorf, W. (ed.), Dreißig Jahre Züchtungsforschung. Fischer Verlag, Stuttgart.

Shull, G. 1908. The composition of a field of maize, Am. Breeders Assoc. Rep. 4. : 296–301.

Stich, B., J. Möhring, H.-P. Piepho, M. Heckenberger, E.S. Buckler, and A.E. Melchinger. 2008. Comparison of mixed-model approaches for association mapping. Genetics 178: 1745–54.

Tenaillon, M.I., and A. Charcosset. 2011. A European perspective on maize history. Comptes Rendus Biologies 334: 221–228.

Warburton, M.L., J.C. Reif, M. Frisch, M. Bohn, C. Bedoya, X.C. Xia, J. Crossa, J. Franco, D. Hoisington, K. Pixley, S. Taba, and A.E. Melchinger. 2008. Genetic Diversity in CIMMYT Nontemperate Maize Germplasm: Landraces, Open Pollinated Varieties, and Inbred Lines. Crop Science 48: 617.

World Bank. 2013. Available at http://databank.worldbank.org/data/home.aspx

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

Chapter 2

Genetic variation among inbred lines and testcrosses of maize for early growth parameters and their relationship to final dry matter yield

Alexander Strigens, Christoph Grieder, Bettina I.G. Haussmann, and Albrecht E. Melchinger

Alexander Strigens, Christoph Grieder, Bettina I.G. Haussmann, and Albrecht E. Melchinger, Institute of Plant Breeding, Seed Science and Population Genetics (350), University of Hohenheim, D-70593 Stuttgart, Germany

Crop Science 52 : 1084-1092
DOI: 10.2135/cropsci2011.08.0426
Article published: May 2012

ABSTRACT

Reduced early growth of maize has a negative impact on subsequent biomass accumulation, and, therefore, on final whole plant dry matter yield (DMY). Quantitative-genetic studies on biomass growth rates and their relation to final DMY in large germplasm sets were so far hampered by a lack of suitable phenotyping tools. In this study, we took advantage of a recently developed non-destructive phenotyping platform to (i) determine early biomass and growth rates in a set of 285 dent inbred lines and their testcrosses with two flint testers grown at three locations in 2008 and 2009, based on non-destructive measurements of biomass between the four- and eight-leaf stage, (ii) estimate variance components and heritability for these traits, (iii) investigate the association of early growth with final DMY and other agronomic traits, and (iv) calculate correlations between line per se performance (LP) and general combining ability (GCA) for these traits. We observed significant genetic variance and high heritabilities for early growth traits, though genotype-by-environment variances were larger than for agronomic traits. Early growth traits showed weak (GCA) to moderate (LP) correlations with final DMY. Correlations between LP and GCA were only moderate for early growth traits, most probably due to masking effects of the testers. Since correlations among early growth traits were tight, visual scoring of early vigor seems sufficient for selection of promising testcrosses.

Chapter 3

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

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

Alexander Strigens, Christoph Grieder, Christian Riedelsheimer, Tobias A. Schrag, and Albrecht E. Melchinger, Institute of Plant Breeding, Seed Science and Population Genetics (350), University of Hohenheim, D-70593 Stuttgart, Germany

Niclas M. Freitag, Xavier Gilbert, and Rainer Messmer, Group of Crop Science, Swiss Federal Institute of Technology Zurich, CH-8092 Zurich, Switzerland

Plant, Cell & Environment
DOI: 10.1111/pce.12096
Article published: May 2013

The original publication is available at:
http://onlinelibrary.wiley.com/doi/10.1111/pce.12096/abstract

ABSTRACT

Chilling sensitivity of maize is a major limitation for its cultivation in cooler areas, because reduced growth in early stages impairs on later biomass accumulation. Efficient breeding for chilling tolerance is hampered by both the complex physiological response of maize to chilling temperatures and the difficulty to accurately measure chilling tolerance in the field under fluctuating climatic conditions. For this research we used genome-wide association (GWA) mapping to identify genes underlying chilling tolerance under both controlled and field conditions in a broad germplasm collection of 375 maize inbred lines genotyped with 56,110 SNPs. We identified nineteen highly significant association signals explaining between 5.7 and 52.5 % of the phenotypic variance observed for early growth and chlorophyll fluorescence parameters. The effect of several quantitative trait loci (QTL) identified for early growth was varying with temperature and incident radiation. Candidate genes involved in ethylene signaling, brassinolide, and lignin biosynthesis were found in their vicinity. Candidate genes involved into signaling or gene expression regulation may explain the complex response of photosynthetic performance and early growth to climatic conditions, and support pleiotropism as a major cause of co-locations of QTL for these highly polygenic traits.

Chapter 4

Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding

Alexander Strigens, Wolfgang Schipprack, Jochen C. Reif and Albrecht E. Melchinger

Alexander Strigens, Wolfgang Schipprack, and Albrecht E. Melchinger, Institute of Plant Breeding, Seed Science and Population Genetics (350), University of Hohenheim, D-70593 Stuttgart, Germany

Jochen C. Reif, Former: State Institute of Agricultural Engineering and Bioenergy (720), Universität Hohenheim, D-70593 Stuttgart, Germany. Present: Leibniz Institute of Plant Genetics and Crop Plant Research, D-06466 Gatersleben, Germany.

PLoS ONE 8(2): e57234.
DOI: 10.1371/journal.pone.0057234
Article published: February 2013

ABSTRACT

Landraces are valuable genetic resources for broadening the genetic base of elite germplasm in maize. Their extensive use was so far hampered by their genetic heterogeneity and heavy genetic load. In this study we assessed the use of the in-vivo doubled haploid (DH) technique to overcome these limitations. We evaluated 132 DH lines derived from three European landraces and 106 elite flint lines at five locations in Germany in 2010 for several agronomic traits. The landraces were further compared with synthetic populations produced by intermating DH lines derived from the respective landrace. Our objectives were to (i) compare the unselected DH lines with elite flint lines, (ii) determine their usefulness for broadening the elite germplasm, and (iii) discuss the potential of the DH technique for conserving landraces and purging them from their genetic load. The lower mean performance of the DH lines was largely compensated by the huge genetic variances: the 40% best DH lines had a grain yield comparable to that of elite flint lines. Selected DH lines with superior early growth may thus be crossed to elite germplasm without tremendous losses on yield level. Enhanced fitness of the synthetic populations, with no reduction of their phenotypic variance suggests mild purging of the genetic load by the use of the in-vivo DH technique. Altogether, our results suggest that DH lines derived from landraces are representative for their genetic diversity. This opens new opportunities for preserving, characterizing and using the genetic diversity stored in gene banks.

Chapter 5

Gene mining in doubled haploids derived from European maize landraces

Maize (*Zea mays* L.) landraces are open pollinated populations of maize that were grown and selected by farmers all over the world for centuries. Through natural and artificial selection, they became adapted to very various environmental and climatic conditions. Compared to modern elite breeding material, their phenotypic and genotypic diversity is tremendous (Gouesnard et al., 1997; Vigouroux et al., 2008) and largely untapped (Hoisington et al., 1999). Through their specific adaptation to local and often marginal environments, it is further expected that they carry useful alleles, not yet tapped in commercial breeding. Introgression of genetic material from landraces in the current elite germplasm would, thus, allow a broadening of its genetic base and ensure ongoing selection gains (Reif et al., 2005). Yet, the heterogeneity and heterozygosity of open pollinated populations makes it particularly difficult to identify and select the traits of interest for breeding, because each plant has a unique and non-reproducible genotype.

Producing doubled haploid (DH) lines out of landraces is a very effective way to capture and fix the diversity present in such open pollinated population (Strigens et al., 2013b). First, it allows an infinite reproduction of the alleles captured and, second, it enables precise phenotying of the genetic material in different locations and over years. Additionally, recessive detrimental alleles not expressed in heterozygous plants might be removed during this process (Prigge et al., 2012). However, the genetic burden present in landraces is not completely removed during the DH line production. Especially the grain yield level of such

DH lines derived from landraces remains behind that of modern elite material (Strigens et al., 2013b). Higher stress and competition tolerance of modern inbred lines compared to first cycle inbred lines developed from landraces by selfing in the 1950's is seen as one of the drivers of the yield increase of hybrids over the past decades (Duvick, 2005; Troyer and Wellin, 2009) and may explain the yield difference observed between the elite material and the DH derived from landraces.

Therefore, introgression of DH lines derived from landraces into elite material to broaden its genetic base or to introduce specific traits (e.g., good early vigor, pest resistances, kernel quality) is always linked with the risk of introducing undesirable properties (e.g., barren stalks, lodging, poor kernel set). A precise identification of the genes involved in the expression of positive and negative traits would allow a targeted introgression of the alleles of interest or exclusion of recombinants carrying the negative alleles. The rapid decay of linkage disequilibrium (LD) observed in landraces (Reif et al., 2005; Tenaillon and Charcosset, 2011) and DH lines derived from them (Strigens et al., 2013b), together with the development of high-density and high-throughput genotyping platforms (Ganal et al., 2011), should allow high resolution genome wide association (GWA) mapping of quantitative trait loci (QTL) in such material.

Population structure and relatedness among genotypes has a strong impact on the number of QTL identified in GWA approaches and especially on the detection of false positives. Different GWA models correcting for population structure and/or kinship were developed to minimize the number of false positives (Yu et al., 2006; Stich et al., 2006). However, the optimal correction factor is largely depending on the trait and population under study (Mezmouk et al., 2011; Riedelsheimer et al., 2012b). Yet, in contrast to elite breeding populations, the very even distribution of genetic distances among DH lines derived from a same landrace suggests an almost complete absence of population structures within this

material (Strigens et al., 2013b). Therefore, it would be interesting to determine whether and how corrections for population structure and kinship in the GWA models affect QTL detection in a mapping population composed of DH lines derived from landraces. Additionally, assessing the effect of including elite material from the same or different heterotic pool into the mapping population on QTL detection, would give some insights in the tradeoff between mapping population size and structure. Further, with increasing population sizes, minor allele frequencies thresholds might have a strong impact on the discovery of rare alleles, which we are actually looking for in landraces.

Therefore, the aim of our study was to (i) identify QTL for several agronomic and morphological traits by GWA mapping in a panel composed of DH lines derived from European flint landraces, as well as of elite dent and flint inbred lines, (ii) evaluate the impact of mapping population composition, population structure and minor allele frequency threshold on QTL detection, and (iii) discuss the use of DH lines developed from landraces for gene mining and improvement of elite material.

MATERIALS & METHODS

Plant material & genotyping

A set of 132 DH lines was produced by KWS SAAT AG (Einbeck, Germany) from the European maize landraces *Bugard* (DH-BU, n = 36), *Gelber Badischer* (DH-GB, n = 31), and *Schindelmeiser* (DH-SC, n = 65) by a proprietary *in-vivo* haploid induction technique similar to the one described by Röber et al. (2005). This set of DH derived from landraces (LR-DH) was evaluated together with a panel of 256 inbred lines developed at the University of Hohenheim (Stuttgart, Germany). This elite breeding material can be assigned to the "European dent" (EU-D, 128 inbred lines) and "European flint" (EU-F, 128 inbred lines) heterotic groups according to pedigree information (Annex 1). Phenotypic and genotypic

diversity of the EU-F material and the LR-DH were described by Strigens et al. (2013b), whereas the EU-D lines were described together with the EU-F inbred lines in a further experiment Strigens et al. (2013a).

Genomic DNA from the 388 inbred lines was extracted from pooled leaf tissue samples of five seedlings per genotype using the CTAB method (CIMMYT, 2005). Each line was genotyped with 56,110 single nucleotide polymorphisms (SNP) using the MaizeSNP50 BeadChip (Illumina Inc., San Diego, USA). Quality control of the SNP marker data was performed according to Strigens et al. (2013a) with minor modifications. Inbred lines showing more than 2% heterozygous loci were excluded. Lines and SNP markers with call rates below 0.95 were excluded from further analysis. Four sets of genotypes and SNPs were defined by combinations of germplasm groups and MAF thresholds (Table 1): Set 1 composed of LR-DH lines only and with a MAF of 0.05; Set 2 composed of LR-DH and EU-F lines, with a MAF of 0.05; Set 3 composed of LR-DH, EU-F and EU-D, with a MAF of 0.05; and Set 4 composed of LR-DH, EU-F and EU-D, with a MAF of 0.025. For each set, only SNPs with an allele frequency over the MAF threshold within the respective set of genotypes were retained for further analysis.

Linkage-disequilibrium (LD) was calculated within each population (DH-BU, DH-GB, DH-SC, EU-F, EU-D) as well as in the whole mapping panel (Set 3) as squared allele frequency correlation (r^2) between pairs of loci for each chromosome (Hill and Robertson, 1966). Obtained values were binned according to the distance between markers in steps of 50 kbp and averaged over chromosomes. To determine the extent of LD, a threshold of $r^2 = 0.1$ was set, below which LD was considered non-significant (Zhu et al., 2008). This distance was considered as confidence interval for the detected QTLs and significant SNP × trait association falling within this distance were considered as a single QTL.

Table 1 Set, mapping population, minor allele frequency (MAF), number of principal components (Q), population size (N), number of polymorphic single nucleotide polymorphisms (SNP), significance level (β), number of significant SNP×trait associations, and number of detected quantitative trait loci (QTL) for thirteen scenarios used to perform genome wide association scans.

Scenario	Set	Population	MAF [%]	Q	N	No. of SNPs	β	No. of SNP×Trait associations	No. of QTL
1	Set1	LR-DH	5	0	141	29279	1.71x10-6	9	5
2	Set1	LR-DH	5	3	141	29279	1.71x10-6	6	2
3	Set1	LR-DH	5	5	141	29279	1.71x10-6	6	2
4	Set1	LR-DH	5	10	141	29279	1.71x10-6	1	1
5	Set2	LR-DH/EU-F	5	0	238	30711	1.63x10-6	25	16
6	Set2	LR-DH/EU-F	5	3	238	30711	1.63x10-6	22	16
7	Set2	LR-DH/EU-F	5	5	238	30711	1.63x10-6	10	9
8	Set2	LR-DH/EU-F	5	10	238	30711	1.63x10-6	4	4
9	Set3	LR-DH/EU-F/EU-D	5	0	364	34137	1.46x10-6	39	25
10	Set3	LR-DH/EU-F/EU-D	5	3	364	34137	1.46x10-6	23	12
11	Set3	LR-DH/EU-F/EU-D	5	5	364	34137	1.46x10-6	20	10
12	Set3	LR-DH/EU-F/EU-D	5	10	364	34137	1.46x10-6	13	8
13	Set4	LR-DH/EU-F/EU-D	2.5	5	364	36328	1.38x10-6	27	17

Phenotyping & statistical analysis

The 388 DH and inbred lines were evaluated in field trials conducted in 2010 in five environments in South Germany (EWE=Eckartsweier, HOH=Hohenheim high N, HOL=Hohenheim low N, KLH=Kleinhohenheim, and OLI=Oberer Lindenhof) contrasting in mean air temperature, elevation, nitrogen supply, and cultivation practice. Detailed information about the experimental design and description of emergence rate (GERM), leaf chlorosis (CHLO), relative growth rates (REGR), female flowering (FFLO), anthesis-silking interval (ASIN), plant height (PLHT), ear insertion height (EAHT), ear shank (EASH) and husk flag leaves (HUFL) score, ear length (EALE) and diameter (EADI), number of kernel rows per ear (ROWS), number of kernels per row (KERO), thousand kernel weight (THKW), ear dry matter content (EDMC), and grain yield (GRYD) were given by Strigens et al. (2013b).

In addition we evaluated the following traits on a plot basis for all lines: Plant emergence score a few days after emergence (EMER), early vigor scores at the four-leaf ($EVIG_4$) and eight-leaf ($EVIG_8$) stage were given on a 1 (good) to 9 (poor) scale. Total standing biomass ($TFMA_{8;OLI}$) in kg was evaluated at the eight-leaf stage at OLI by destructive harvest. Biomass per plant($MAPL_{8;OLI}$) in g was obtained by dividing the harvested fresh biomass by the number of standing plants. Male flowering (MFLO) was recorded as sum of growing day degrees (GDD, base temperature= 10°C) from sowing until 50% of the plants were shedding pollen and leaf greenness at flowering (SPAD) was measured with a SPAD-502 Chlorophyll Meter (Konica Minolta Sensing Inc., Sakai. Osaka, Japan) on the mid part of the top ear leaf as indicator of the nutritional state of the plants. Occurrence of lodging (LODG), common smut (*Ustilago maydis,* SMUT) and barren stalks (BAST) was recorded on mature plants and converted to percentage of affected plants before the ears of five plants were harvested by hand. A husk coverage score (HUCO) was given on a 1 (good) to 9 (poor) scale. Before

shelling, incidence (IFUS) and severity (SFUS) of ear rot (*Fusarium* ssp.) infestation in % was evaluated on the main ears of five harvested plants. Oil content of the grains (KOIL) in % was measured on the material harvested in EWE with nuclear magnetic resonance on four samples of five kernels for each plot.

Early plant height in cm at the four-leaf ($EPHT_4$), six-leaf ($EPHT_6$), and eight-leaf ($EPHT_8$) stage, as well as early fresh standing biomass in g m^{-2} at the four-leaf ($EFMA_4$), six-leaf ($EFMA_6$), and eight-leaf ($EFMA_8$) stage were measured with a non-destructive phenotyping platform combining spectral reflectance and light curtain (Montes et al., 2011). Mean REGR over environments were computed from the standing biomass as described in Strigens et al. (2012). Additionally, we considered the REGR measured at single locations ($REGR_{EWE}$, $REGR_{HOH}$, $REGR_{HOL}$, $REGR_{KLH}$, $REGR_{OLI}$) as individual traits (Strigens et al., 2013a).

For estimation of adjusted means of the genotypes, best linear unbiased estimates (BLUEs) were computed with the following model:

$$y_{iklmn} = \mu + g_i + e_k + ge_{ik} + t_{kl} + r_{klm} + b_{klmn} + \varepsilon_{iklmn}\,, \qquad (1)$$

where y_{iklmn} is the observed plot value, μ the overall mean, g_i the effect of genotype i, e_k is the effect of environment k, ge_{ik} the interaction between genotype i and environment k, t_{kl} the effect of trial l within environment k, r_{klm} the effect of replication m within trial l, b_{klmn} the effect of incomplete block n within replication m, and ε_{iklmn} the residual. All effects in Eq. (1) except μ and g_i were considered as random. Heterogeneity of residual variance among environments was taken into account and the pooled residual variance was calculated as the average of the individual estimates. For ear dry matter content, the sum of GDD from female flowering to harvest was additionally taken as covariate to adjust for different harvest dates. To estimate BLUEs of traits evaluated at single locations only, the terms e_k and ge_{ik} were dropped from Eq. (1).

For estimation of variance components across the whole panel, all effects in Eq. (1) except μ were considered as random. Estimates of the genotypic variance (σ^2_g), the variance of genotype × environment interactions ($\sigma^2_{g\times e}$), and error variance (σ^2_ε) were computed by restricted maximum likelihood. Heritabilities (h^2) were calculated according to Hallauer et al. (2010). To compare genotypic variances among the different populations (EU-D, EU-F, LR-DH), an additional term p_i was added in Eq. (1) to account for population effects and the variance components σ^2_g and $\sigma^2_{g\times e}$ were estimated within population by using a diagonal variance-covariance structure for both terms (Strigens et al., 2013a).

All calculations were performed within the R-environment (R Development Core Team, 2011). Mixed model analyses were performed using the package ASReml for the R-environment (Butler et al., 2007).

Genome wide association mapping

We adopted the two-step approach described by Stich et al. (2008) to perform GWA mapping and used BLUEs of genotypic means instead of best linear unbiased predictions (BLUPs) to avoid artifacts arising from a two-fold shrinkage of genotypic effects in two-step approaches (Piepho et al., 2012). The GWA analysis was conducted with models correcting for population structure (Q) and kinship (K) to avoid spurious trait × marker associations resulting from the confounding of population structure and phenotypic values. The K and Q matrices were computed separately for each SNP set. The K matrix was computed as proportion of shared SNP alleles (Eding and Meuwissen, 2001). The Q matrix was evaluated by principal coordinate analysis (PCoA) based on modified Rogers' distances between genotypes (Gower, 1966). To assess the effect of population structure correction on SNP detection, we used different levels of population structure correction by including no (Q_0), three (Q_3), five (Q_5) or ten (Q_{10}) first principal coordinates in the K + Q model.

Genome-wide association scans were performed for forty-one traits with the thirteen scenarios described in Table 1, using the maximum likelihood implementation in the function polygenic of GenABEL 1.6-5 (Aulchenko et al., 2007; Chen and Abecasis, 2007). The principal coordinates Q were considered as fixed effects and the K matrix as variance-covariance matrix for random genotype effects. *P* values were obtained with a one-degree of freedom score test implemented in the function mmscore of GenABEL. Genome-wide inflation (λ) was calculated as the regression coefficient of observed *P* values on expected *P* values with a zero intercept. The significance threshold ($\alpha = 0.05$) was Bonferroni-corrected according to the number of tested SNPs to obtain a significance level β for each SNP set (Table 1). Gene models associated with significant SNP were obtained from Ganal et al. (2011).

RESULTS

Means & variance components

Large phenotypic variation was observed for all traits (Table 2). Heritabilities over the whole panel were moderate to very high except for fresh mass at the 4-leaf stage and relative growth rates in HOH. The LR-DH showed a stronger early vigor and early biomass accumulation than the elite material in all stages but smaller plants and lower grain yields at maturity (Strigens et al., 2013b ; Annex 2). In contrast to the EU-F and EU-D material, the LR-DH had an increased occurrence of barren stalks, lodging, and common smut infestation, as well as longer husk flag leaves. Whereas the incidence of ear rot was similar across all sets of material, the severity of infestation was only slightly higher in LR-DH compared to EU-D materials. The LR-DH and EU-F material had a significantly ($P<0.001$) higher kernel oil content compared to the EU-D lines, while no difference could be found among the flint populations (Figure 1a). A few DH-SC lines with slightly shrunken brown kernels had a significantly higher ($P<0.05$) kernel oil content compared to the remaining DH-SC lines with yellow kernels (Figure 1b).

†Traits are: GERM, emergence rate; EMER, emergence score; CHLO, leaf chlorosis; $EVIG_4$, early vigor score at the four-leaf; $EVIG_8$, early vigor score eight-leaf; $EFMA_4$, standing biomass at the four-leaf stage (remote-sensing); $EFMA_6$, standing biomass at the six-leaf stage (remote-sensing); $EFMA_8$, standing biomass at the eight-leaf stage (remote-sensing); $EPHT_4$, plant height at the four-leaf stage (remote-sensing); $EPHT_6$ plant height at the six-leaf stage (remote-sensing); $EPHT_8$, plant height at the eight-leaf stage (remote-sensing); $TFMA_8$, total fresh mass at the eight-leaf stage (destructive); $MAPL_8$, biomass per plant (destructive); REGR, relative growth rates; FFLO, female flowering; MFLO, male flowering; ASIN, anthesis-silking interval; SPAD, leaf greenness; PLHT, plant height; EAHT, ear insertion height; EASH, ear shank score; HUCO, husk coverage score; HUFL, husk flag leaves score; LODG, occurrence of lodging; SMUT, occurrence of common smut (*Ustilago maydis*); BAST, occurrence of barren stalks; IFUS, incidence of ear rot (*Fusarium* ssp.); SFUS, ear rot severity; EALE, ear length; EADI, ear diameter; ROWS, number of rows per ear; KERO, number of kernels per row; THKW, thousand kernel weight; EDMC, ear dry matter content; GRYD, grain yield; KOIL, kernel oil content.
‡ 1 = good, 9 = poor
§ 1 = absent, 9 = pronounced
¶ Based on a single location (EWE: Eckartsweier; HOH: Hohenheim high N; HOL: Hohenheim low N; KLH: Kleinhohenheim; OLI: Oberer Lindenhof)

Table 2. Average, minimum and maximum adjusted mean of the genotypes evaluated in the genome wide association analysis.

Trait†	Mean	Minimum	Maximum	Heritability
EMER [1-9]‡	3.63	1.52	8.37	0.89
GERM [%]	60.51	6.80	77.29	0.92
CHLO [1-9]§	3.63	1.63	7.59	0.82
$EVIG_4$ [1-9]‡	4.54	1.23	7.48	0.92
$EVIG_8$ [1-9]‡	4.21	1.42	7.17	0.92
$EFMA_4$ [g m-2]	85.83	52.25	146.80	0.49
$EFMA_6$ [g m-2]	93.03	26.94	208.03	0.82
$EFMA_8$ [g m-2]	276.46	159.62	505.42	0.81
$EPHT_4$ [cm]	8.82	4.94	15.15	0.86
$EPHT_6$ [cm]	14.61	7.83	24.23	0.93
$EPHT_8$ [cm]	23.80	13.55	34.95	0.92
$TFMA_{8;OLI}$ [kg $plot^{-1}$]¶	1.82	0.16	3.79	-
$MAPL_{8;OLI}$ [g $plant^{-1}$]¶	74.41	13.16	155.25	-
REGR [x10-3 GDD-1]	17.58	14.85	20.27	0.75
$REGR_{EWE}$ [$x10^{-3}$ GDD^{-1}]¶	15.40	3.52	19.75	-
$REGR_{HOH}$ [$x10^{-3}$ GDD^{-1}]¶	17.49	6.59	19.77	-
$REGR_{HOL}$ [x10-3 GDD-1]¶	16.35	13.06	19.54	-
$REGR_{KLH}$ [$x10^{-3}$ GDD^{-1}]¶	20.19	14.68	22.62	-
$REGR_{OLI}$ [$x10^{-3}$ GDD^{-1}]¶	18.14	10.82	21.67	-
FFLO [GDD]	663.28	539.51	789.93	0.96
MFLO [GDD]	629.66	530.64	777.72	0.97
ASIN [GDD]	33.35	-13.26	106.04	0.90
SPAD [SPAD unit]	50.34	36.97	61.45	0.86
PLHT [cm]	149.20	80.08	207.87	0.97
EAHT [cm]	50.35	20.09	86.79	0.93
EASH [1-9]§	4.28	1.60	8.59	0.93
HUCO [1-9]‡	2.12	0.96	8.61	0.95
HUFL [1-9]§	2.09	0.93	7.75	0.93
LODG [%]	8.83	0.00	86.53	0.78
SMUT [%]	4.62	0.00	52.75	0.79
BAST [%]	3.74	0.00	48.28	0.67
IFUS [%]	27.70	1.64	100.00	0.88
SFUS [%]	3.15	0.00	60.20	0.80
EALE [cm]	127.01	76.10	188.09	0.92
EADI [cm]	34.58	23.25	44.75	0.93
ROWS [#]	12.22	7.14	17.71	0.97
KERO [#]	19.71	0.80	30.02	0.87
THKW [g]	217.20	125.90	328.69	0.92
EDMC [%]	57.64	27.73	74.40	0.93
GRYD [g]	50.73	4.63	84.77	0.83
$KOIL_{EWE}$ [%]	4.27	2.96	6.19	-

Genotypic variances within population were highest for the LR-DH in almost all instances (Annex 2). The differences were especially striking for common smut incidence, length of husk flag leaves, and the occurrence of barren stalks and lodging, where genotypic variance was almost absent in the EU-F and EU-D material. Estimates of $\sigma^2_{g\times e}$ were particularly high (>50% of σ^2_g) for early growth parameters in all populations, for the occurrence of lodging and barren stalks in the LR-DH material, for grain yield in the EU-D population, as well as for ear rot severity in the EU-D and LR-DH materials.

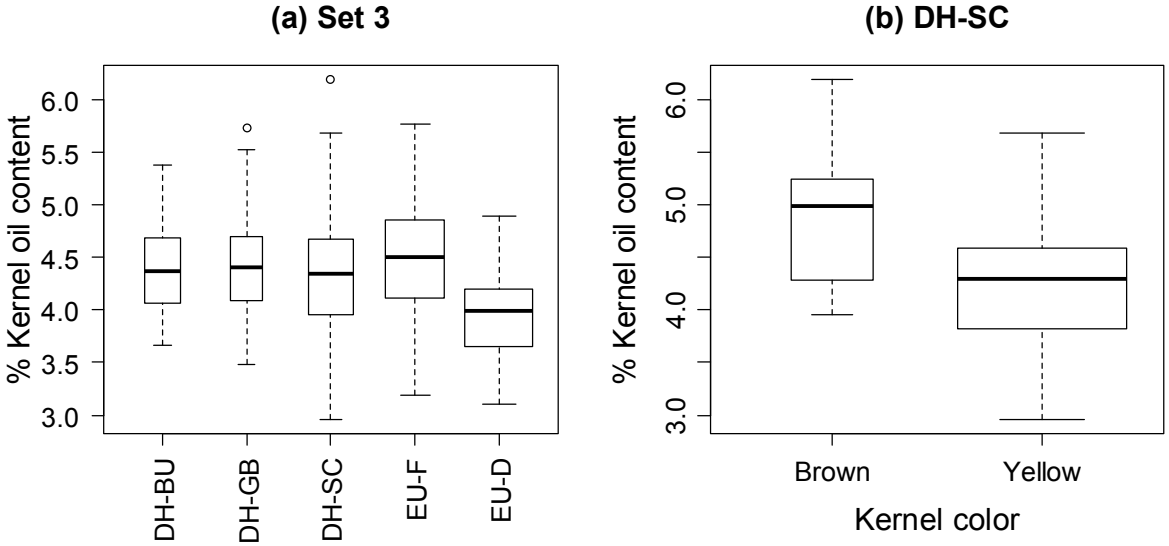


Figure 1. Kernel oil content of (a) Set 3 composed of doubled haploid (DH) lines derived from the European landraces *Bugard* (BU), *Gelber Badischer* (GB) and *Schindelmeiser* (SC), as well as of elite European flint (EU-F) and dent (EU-D) inbred lines, (b) DH lines derived from SC with brown or yellow kernels.

Marker distribution and population structure

A total of 29'279 polymorphic SNPs were retained in Set 1 after quality check (Table 1). Including the EU-F inbred lines in the mapping population while keeping the MAF threshold at 0.05 (Set 2) resulted in 1'432 more polymorphic SNPs (+4.9%) that can be considered as fixed in the LR-DH. Including the EU-D inbred lines with a MAF threshold of 0.05 (Set 3) resulted in 3'426 more polymorphic SNPs (+11.2%) that can be considered as fixed within

the flint heterotic pool. Reducing the MAF to 0.025 in the whole mapping population (Set 4) resulted in additional 2'191 polymorphic SNPs (+6.4%). These rare alleles occurred in 9 to 18 genotypes of Set 4. Linkage disequilibrium (r^2) dropped on average below the threshold of 0.1 within 0.725 Mbp in Set 3, whereas it stretched over more than 5 Mbp in the EU-D population. The LD decay within DH-BU, DH-GB, DH-SC, and EU-F was intermediate with values ranging from 0.275 Mbp for DH-GB to 3.875 Mbp for EU-F (Strigens et al., 2013b).

Principal coordinate analysis of the whole mapping population (Set 3) revealed four main clusters corresponding to the EU-D, EU-F, DH-BU, and a common group composed of DH-GB and DH-SC (Figure 2a). The proportion of variance among genotypes explained by the first and second principal coordinates was 17.3% and 7.9%, respectively. Limiting the PCoA to Set 2 resulted in three main clusters corresponding to EU-F, DH-BU, and a common cluster composed of DH-GB and DH-SC (Figure 2b). The amount of variance among genotypes explained by the first and second principal coordinates was 14.1% and 8.3%, respectively. Limiting the PCoA to Set 1 resulted in three main clusters corresponding to DH-BU, DH-GB and DH-SC (Figure 2c). The amount of variance among genotypes explained by the first and second principal coordinates was 16.7% and 5.3%, respectively.

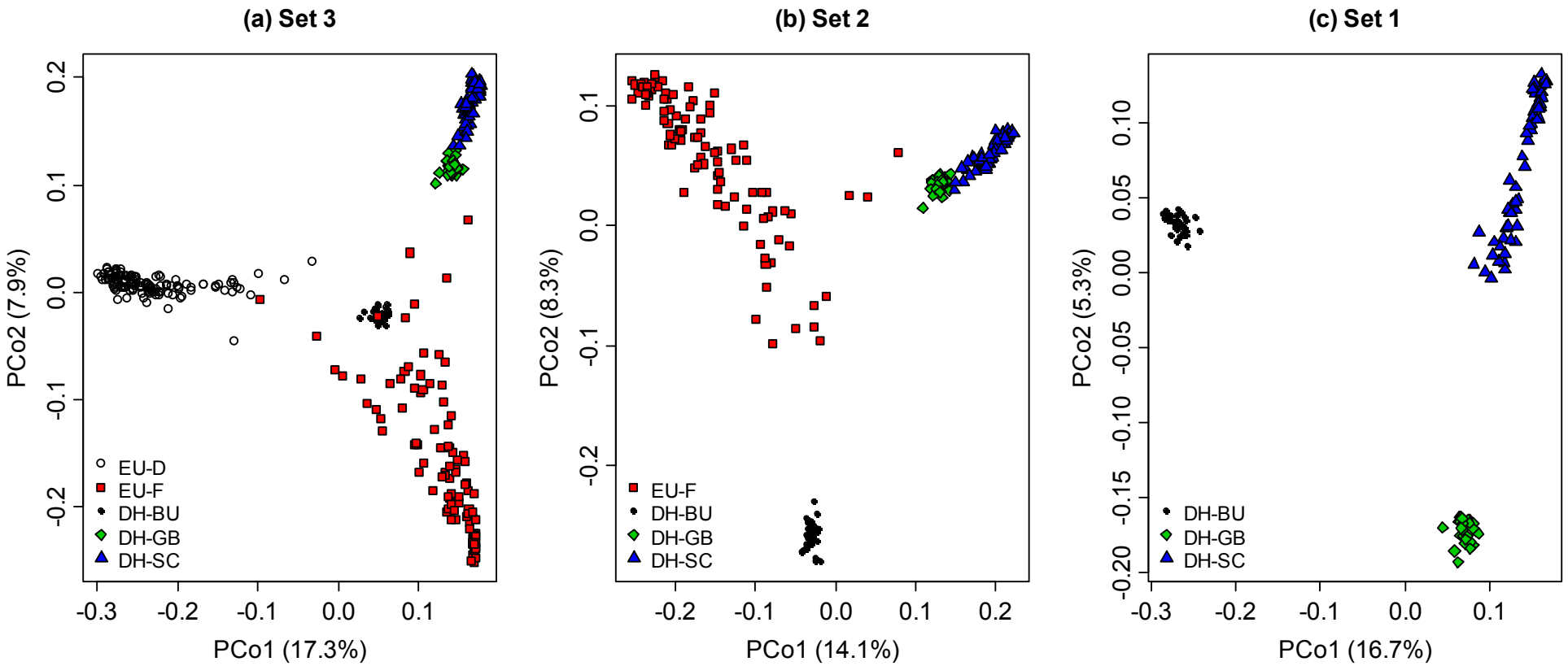


Figure 2. Biplot of the two first principal coordinates of (a) Set 3 composed of doubled haploid (DH) lines derived from European flint landraces *Bugard* (DH-BU), *Gelber Badischer* (DH-GB) and *Schndelmeiser* (DH-SC) as well as elite European flint (EU-F) and dent (EU-D) inbred lines, (b) Set 2 composed of DH-BU, DH-GB, DH-SC and EU-F, and (c) Set 1 composed only of DH-BU, DH-GB and DH-SC.

Genome wide association mapping

A total of 204 significant trait×SNP association were detected with the thirteen different GWA scenarios for 27 of the 41 measured traits. They corresponded to 69 unique SNPs and 49 QTL distributed across all chromosomes except chromosome 9 (Table 3). Results for scenario 11 are shown in Annex 3.The maximal distance between SNPs within single QTL reached from 0.006 to 725.3 kbp. Gene models could be associated to 42 of the SNPs, whereby more than one candidate gene was associated with QTL 5, 31, 39, 40, and 45 (Annex 4). Prevalence of the positive allele at each marker was varying between the three germplasm groups (Annex 5). Alleles associated with superior early growth performance were nearly fixed or had higher frequencies in the LR-DH and EU-F, whereas alleles associated with low incidence of lodging were almost fixed in the elite material. Alleles associated with higher oil content were distributed across all populations and none of them was specific to the DH-SC lines with brown kernels and high oil concentration. Most of these DH-SC lines carried all three alleles increasing the oil content.

For all sets, the number of QTL decreased with increasing numbers of principal coordinates included in the GWA models (Table 1). At the same level of correction for population structure, the number of QTL was generally higher in Set 3 composed of LR-DH, EU-F and EU-D materials than in Set 2 composed of LR-DH and EU-F material only. The lowest number of QTL was detected in Set 1 composed of LR-DH only. Fourteen QTL were identified in Set 3 only, whereas twelve QTL were detected in Set 2 only and three QTL in Set 1 only. QTL 8, associated with lodging, and QTL 31, associated with kernel oil content, were detected with all models in Set 2, Set 3 and Set 4, but not in Set 1.

A reduction of the MAF level from 0.05 to 0.025 resulted in the detection of six additional QTL in Set 4 compared to Set 3. The six associated SNPs were present in less than nineteen

genotypes (Annex 5), and, thus, below the MAF threshold of 0.05 used in Set 3. Each of these SNPs had a MAF >0.05 in at least one of the populations under study. QTL 18, associated with the length of the husk flag leaves, was also detected in Set 2 with all models, and in Set 1 with two models.

Candidate genes

Several highly plausible candidate genes could be identified in the vicinity of the significant trait × SNP associations (Annex 4). The gene *Rough sheath2* (*rs2*, GRMZM2G403620) was found within QTL 5 associated with germination and RGR in Oberer Lindenhof. An aldehyde oxidase (GRMZM2G141473), similar (57%) to the one overexpressed in the *Arabidopsis thaliana* mutant *superroot1* (Seo et al., 1998), was identified within QTL 8 associated with lodging. A β-amylase (GRMZM2G462258) and a pectinesterase (GRMZM2G162333) were identified within QTL11 associated with fresh weight at the four-leaf-stage. The diacylglycerol acyltransferase (*dgat1-2,* GRMZM2G169089) involved in the lipid pathway (Zheng et al., 2008) was found within QTL 31 associated with oil content. This QTL covers also the location of the *Linoleic acid 1* (*ln1*) locus and co-locates with several oil content and quality QTL identified in previous studies (Wassom et al., 2008; Yang et al., 2010; Cook et al., 2012).

Table 3. Detection of quantitative trait loci (QTL) identified by genome wide association analysis in thirteen scenarios differing in mapping populations composition (LR-DH: doubled haploid lines derived from European landrace; EU-F: elite European flint inbred lines; EU-D: elite European dent inbred lines), minor alleles frequency in % (MAF), and level of correction for population structure (K: kinship matrix; Q_i: *i* first components of principal coordinates matrix).

		LR-DH				**LR-DH & EU-F**				**LR-DH & EU-F & EU-D**					
		MAF5				**MAF5**				**MAF5**				**MAF2.5**	
Chr	**QTL**	**$K+Q_0$**	**$K+Q_3$**	**$K+Q_5$**	**$K+Q_{10}$**	**$K+Q_0$**	**$K+Q_3$**	**$K+Q_5$**	**$K+Q_{10}$**	**$K+Q_0$**	**$K+Q_3$**	**$K+Q_5$**	**$K+Q_{10}$**	**$K+Q_5$**	**Traits†**
1	1						■								KERO
	2													■	SFUS
	3									■					$EVIG_4$
	4					■									EMER, GERM, $REGR_{OLI}$
	5	■	■	■		■									EMER, $REGR_{OLI}$
	6					■	■	■							EAHT
	7									■					ROWS
	8					■	■	■	■	■	■	■	■	■	LODG
2	9	■													$EPHT_8$
	10									■					CHLO
	11									■					$EFMA_4$
	12									■				■	THKW
	13									■	■	■	■	■	KOIL
3	14					■	■	■		■	■	■		■	SFUS
	15										■				GRYD
	16					■	■	■							SFUS
	17													■	SMUT
	18		■	■		■	■	■	■					■	HUFL
	19				■										$EVIG_8$
4	20					■	■								$REGR_{HOH}$
	21					■	■			■					$REGR_{HOH}$
	22						■								$REGR_{HOH}$
	23													■	HUFL
	24													■	LODG
	25									■					LODG
5	26	■				■	■			■					HUCO
	27					■	■	■		■	■	■	■	■	LODG
	28					■	■			■	■	■	■	■	$TFMA_{8;OLI}$, $EFMA_6$, $EFMA_8$, $MAPL_{8;OLI}$, $EPHT_4$, $EPHT_6$, $EPHT_8$
	29					■									ROWS
	30													■	GERM

† For traits description see table 2

Table 3 (continued).

		LR-DH				LR-DH & EU-F				LR-DH & EU-F & EU-D					
		MAF5				MAF5				MAF5				MAF2.5	
Chr	**QTL**	**K+Q$_0$**	**K+Q$_3$**	**K+Q$_5$**	**K+Q$_{10}$**	**K+Q$_0$**	**K+Q$_3$**	**K+Q$_5$**	**K+Q$_{10}$**	**K+Q$_0$**	**K+Q$_3$**	**K+Q$_5$**	**K+Q$_{10}$**	**K+Q$_5$**	**Traits†**
6	31					■	■	■	■	■	■	■	■	■	KOIL
	32									■					LODG
7	33									■	■	■	■	■	LODG
	34	■													EMER, $REGR_{OLI}$
	35					■	■	■							SFUS
	36							■							HUFL
	37									■	■				SMUT
	38								■						KOIL
	39									■	■	■	■	■	LODG, $REGR_{HOH}$
	40									■	■	■		■	$REGR_{HOH}$
	41									■	■	■		■	EMER, $EVIG_8$, $EFMA_8$, $EPHT_8$
8	42									■					MFLO
	43									■					MFLO
	44									■					ROWS
	45						■								$REGR_{HOH}$
	46									■					MFLO
	47					■	■								$REGR_{HOH}$
10	48												■		$EVIG_4$
	49									■					ROWS

† For traits description see table 2

DISCUSSION

In agreement with previous studies on landraces, we could observe a huge phenotypic and genotypic diversity in the set of DH lines derived from landraces for all traits measured. As expected, there was also a large variance for unwanted properties within this material, as shown by the higher means and genotypic variance for the occurrence of barren stalks, lodging and common smut in DH lines derived from landraces compared to elite material. Because barren stalks, lodging and common smut can be regarded as a sign of low stress and concurrence tolerance (Betran et al., 2003; Duvick, 2005), this also reflects the absence of selection for high planting density within the landraces. This illustrates the part of the genetic burden of the landraces that was not removed during the DH production (Strigens et al., 2013b). Introgression of DH lines derived from landraces into elite materials to broaden its genetic diversity runs, thus, still the risk of re-introducing traits selected against during the past decades. A precise identification of the responsible genes would greatly help to select the best recombinants.

With regard to the large phenotypic and genotypic variances in our mapping panel composed of elite lines and DH lines derived from landraces, we expected to detect numerous QTL underlying the measured traits. Yet, only a relatively low number of QTL was identified across all sets. This can be due to several factors: population size, degree of polymorphism in the population, LD decay, desired significance level (Yan et al., 2011), population structure (Mezmouk et al., 2011), and nature of the traits (Riedelsheimer et al., 2012b).

Influence of population size on QTL detection

As expected, the number of QTL detected for a given MAF, population structure correction and significance level increased from Set 1 to Set 3 with the number of genotypes included in the GWA scan. At the one hand, the population size directly improved the power of the

performed score test, while at the other hand, additional polymorphic markers were included. The number of genotypes was certainly the main cause of increased number of QTL detected in Set 2 compared to Set 1, while the inclusion of additional polymorphisms from the dent material in Set 3 was certainly as important as the increased number of genotypes compared to Set 2. The importance of the number of polymorphism included in the mapping population was underlined by the additional QTL detected in Set 4, because the reduction of the MAF level to 0.025 included additional SNPs without affecting population size.

Mapping populations of larger size must, therefore, be composed with a strong focus on their diversity or, more precisely, on their effective population size (Riedelsheimer et al., 2012a). This might in particular be a challenge for mapping populations mainly composed of elite breeding material. Such panels might actually have a low effective population size despite large numbers of genotypes due to the ongoing inbreeding within such breeding population (Geiger and Gordillo, 2010).

Influence of population structure on QTL detection

Joining different mapping panels to increase both the size and the diversity of the mapping population, as done here, is a practical solution, but may result in strong population structures within the mapping panel. Several SNP were detected in Set 1 and/or Set 2 but not in Set 3 despite of much larger population size and increased number of polymorphic SNPs. Yet, the PCoA performed within the different sets of material (Set 1, 2, and 3) showed that the two first principal coordinates of the respective PCoAs explained similar proportions of the total genetic variance in each set. Therefore, the grouping of mapping populations had a negative impact on QTL detection despite the proportion of variance accounted for by the fix effects in the GWA model did not change. The non-detection of QTL identified in the smaller sets might be due to epistasis (Van Inghelandt et al., 2012) and/or to differences in the correlation

between population structure and trait expression in the larger mapping population. Phenotypic differences between flint and dent material, such as flowering time, plant height and early vigor were probably accounted for by the first principal coordinate (and even the K matrix) in Set 3, but not in Set 1 and 2. Therefore, it appears important to evaluate the population structure of the examined material on a genetic basis as well as on a phenotypic basis prior to grouping different association panels in joint GWA analyses.

Influence of minor allele frequency on QTL detection

As illustrated by QTL 18, the detection of QTL in smaller populations might also be partially explained by the presence of rare alleles that fall below the MAF threshold in larger mapping populations. This might especially be critical when working with very diverse material such as landraces in which rare alleles are expected and looked for. Strong support for considering these alleles with low frequencies as real rare alleles instead of genotyping errors was their non-random distribution pattern across the populations. Given that the probability of a genotyping mistake with the MaizeSNP50 platform was estimated to be below 1% in technical replicates and analyses of parent-F1 triplets (Ganal et al., 2011) and that such genotyping errors may rather follow a Poisson distribution with low λ values than a binomial distribution with $\pi = 0.05$ (the commonly used threshold for MAF), an adaptation of the MAF in large mapping population, as done for Set 4, is recommended.

Gene mining in doubled haploids derived from landraces

Many associations pointed to genes of unknown function or to no gene at all. Before interpreting the first as newly discovered genes of yet undiscovered function, it would be advisable to confirm those QTL in further populations and independent panels. Some of the QTL pointing to no gene might be false positives despite the stringent significance level correction used. However, these associations might also indicate some *cis* acting elements

(Van Inghelandt et al., 2012). Further, regarding the long range of LD in the EU-D and EU-F, candidate genes might be located in a wider window, beyond the gene × marker associations reported for the MaizeSNP50 chip (Ganal et al., 2011; Strigens et al., 2013a).

Several QTL were associated with well-known genes (e.g., *dgat1-2*, *rs2*) identified in previous studies or with proteins being plausible candidate genes owing to their expected function, confirming the power of association mapping to detect QTL in very diverse panels. Interestingly, none of the QTL identified for oil content could explain alone the very high oil content of the DH-SC lines with brown kernels. The combination of the three QTL identified for oil content (QTL 13, 31, 38) explained the observed phenotype. However, there might be more alleles than the two captured by the single SNPs or epistatic genes involved in this trait expression, because a few lines carrying the positive allele at all three QTL still had yellow kernels. Sequencing of the identified candidate genes in the DH-SC lines with brown kernels or a haplotype based approach of GWA might provide further insights in the control of oil content in maize kernels. In general, this illustrates well the limitations of GWA methods to explain complex traits involving from a few to many interacting genes (Riedelsheimer et al., 2012c; b). It shows also that the landraces carry properties or alleles combinations that are not present in the elite material, and, thus, underlines the great value of landraces as source of new alleles and haplotypes.

If most of the SNP identified in this study were already segregating in the elite material, many of them could only be detected in the combined analysis of elite material and LR-DH lines, because they would have fallen below the MAF threshold in the elite material alone. Inclusion of unselected material derived from landraces was, therefore, valuable to identify rare, often negative alleles that were certainly strongly selected against in elite material (*e.g.*, QTL 8, 24, 27, 32, 39 for lodging, QTL 18, 23 for husk flag leaves length). Screening for these alleles during the introgression of material derived from landraces or from other exotic sources

would allow the selection of the best recombinants and facilitate the use of landraces as genetic resources.

CONCLUSION

We showed that the composition of the mapping population, the choice of the MAF and the level of correction for population structure are tightly interconnected. Therefore, each mapping population should be investigated with different approaches, knowing the limitations of each. Associations detected with several models and levels of correction for population structure might be the most promising ones, but those correlated with the population structure will be omitted (Mezmouk et al., 2011). Conception of mapping panels breaking the co-linearity of trait expression and population structure like the nested association mapping (NAM) population is useful (Yu et al., 2008), but the range of the included material and, thus, the effective population size is limited. The combination of elite material and DH lines derived from landraces in our study strongly increased the number of haplotypes included and allowed high resolution mapping of QTL by GWA. However, the combination of strongly differentiated heterotic pools increased effects of the population structure. Performing GWA in a larger set of DH lines derived from landraces might overcome all these limitations. The larger phenotypic variation within landraces than between landraces will disrupt the co-linearity between trait expression and population structure, while the large genetic diversity will ensure a high effective population size. Consequently, such populations would represent a perfect tool to perform gene mining and identify new genes and alleles.

REFERENCES

Aulchenko, Y.S., S. Ripke, A. Isaacs, and C.M. van Duijn. 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics (Oxford, England) 23: 1294–1296.

Betran, F., D. Beck, M. Bänziger, and G. Edmeades. 2003. Secondary traits in parental inbreds and hybrids under stress and non-stress environments in tropical maize. Field Crops Research 83: 51–65.

Butler, D., B. Cullis, A. Gilmour, and B. Gogel. 2007. Analysis of mixed models for S language environments. ASReml-R reference manual. 2.0. The State of Queensland, Department of Primary Industries and Fisheries, Brisbane.

Chen, W.-M., and G.R. Abecasis. 2007. Family-based association tests for genomewide association scans. American journal of human genetics 81: 913–926.

CIMMYT. 2005. Laboratory protocols: CIMMYT applied molecular genetics laboratory. 3rd ed. CIMMYT.

Cook, J.P., M.D. Mcmullen, J.B. Holland, F. Tian, P.J. Bradbury, J. Ross-ibarra, E.S. Buckler, and S.A. Flint-Garcia. 2012. Genetic Architecture of Maize Kernel Composition in the Nested Association Mapping and Inbred Association Panels. Plant Physiology 158: 824–834.

Duvick, D.N. 2005. Genetic progress in yield of united states maize (*Zea mays* L .). Maydica 50: 193–202.

Eding, H., and T.H.E. Meuwissen. 2001. Marker based estimates of between and within population kinships for the conservation of genetic diversity. Journal of Animal Breeding and Genetics 118: 141–159.

Ganal, M.W., G. Durstewitz, A. Polley, A. Bérard, E.S. Buckler, A. Charcosset, J.D. Clarke, E.-M. Graner, M. Hansen, J. Joets, M.-C. Le Paslier, M.D. McMullen, P. Montalent, M. Rose, C.-C. Schön, Q. Sun, H. Walter, O.C. Martin, and M. Falque. 2011. A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PloS one 6: e28334.

Geiger, H.H., and G.A. Gordillo. 2010. Doubled haploids in hybrid maize breeding. Maydica 54: 485–499.

Gouesnard, B., J. Dallard, A. Panouillé, and A. Boyat. 1997. Classification of French maize populations based on morphological traits. Agronomie 17: 491–498.

Gower, J. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325–338.

Hallauer, A., M.J. Carena, and J.B. Miranda. 2010. Quantitative genetics in maize breeding. Springer Science and Business Media LLC, New York,NY.

Hill, W.G., and A. Robertson. 1966. Linkage Disequilibrium in Finite Populations. Theoretical and Applied Genetics 38: 226–231.

Hoisington, D., M. Khairallah, T. Reeves, J.-M. Ribaut, B. Skovmand, S. Taba, and M.L. Warburton. 1999. Plant genetic resources: what can they contribute toward increased crop productivity? Proceedings of the National Academy of Sciences 96: 5937–5943.

Mezmouk, S., P. Dubreuil, M. Bosio, L. Décousset, A. Charcosset, S. Praud, and B. Mangin. 2011. Effect of population structure corrections on the results of association mapping tests in complex maize diversity panels. Theoretical and Applied Genetics 122: 1149–60.

Montes, J.M., F. Technow, B.S. Dhillon, F. Mauch, and A.E. Melchinger. 2011. High-throughput non-destructive biomass determination during early plant development in maize under field conditions. Field Crops Research 121: 268–273.

Piepho, H.-P., J. Möhring, T. Schulz-Streeck, and J.O. Ogutu. 2012. A stage-wise approach for the analysis of multi-environment trials. Biometrical journal. Biometrische Zeitschrift 54: 844–60.

Prigge, V., R. Babu, B. Das, M.H. Rodriguez, G.N. Atlin, and A.E. Melchinger. 2012. Doubled haploids in tropical maize: II. Quantitative genetic parameters for testcross performance. Euphytica 185: 453–463.

R Development Core Team. 2011. R: a language and environment for statistical computing. Available at http://www.r-project.org.

Reif, J.C., S. Hamrit, M. Heckenberger, W. Schipprack, H. Peter Maurer, M. Bohn, and A.E. Melchinger. 2005. Genetic structure and diversity of European flint maize populations determined with SSR analyses of individuals and bulks. Theoretical and Applied Genetics 111: 906–913.

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

Riedelsheimer, C., J. Lisec, A. Czedik-Eysenberg, R. Sulpice, A. Flis, C. Grieder, T. Altmann, M. Stitt, L. Willmitzer, and A.E. Melchinger. 2012b. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. Proceedings of the National Academy of Sciences 109:8872–8877.

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

Röber, F.K., G.A. Gordillo, and H.H. Geiger. 2005. In vivo haploid induction in maize - performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 50: 275–283.

Seo, M., S. Akaba, T. Oritani, M. Delarue, C. Bellini, M. Caboche, and T. Koshiba. 1998. Higher Activity of an Aldehyde Oxidase in the Auxin-Overproducing superroot1 Mutant of Arabidopsis thaliana 1. : 687–693.

Stich, B., A.E. Melchinger, H.-P. Piepho, M. Heckenberger, H.P. Maurer, and J.C. Reif. 2006. A new test for family-based association mapping with inbred lines from plant breeding programs. Theoretical and Applied Genetics 113: 1121–30.

Stich, B., J. Möhring, H.-P. Piepho, M. Heckenberger, E.S. Buckler, and A.E. Melchinger. 2008. Comparison of mixed-model approaches for association mapping. Genetics 178: 1745–54.

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

Strigens, A., C. Grieder, B.I. Haussmann, and A.E. Melchinger. 2012. Genetic variation among inbred lines and testcrosses of maize for early growth parameters and their relationship to final dry matter yield. Crop Science 52: 1084–1092.

Strigens, A., W. Schipprack, J.C. Reif, and A.E. Melchinger. 2013b. Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding. PloS one 8: e57234.

Tenaillon, M.I., and A. Charcosset. 2011. A European perspective on maize history. Comptes Rendus Biologies 334: 221–228.

Troyer, A.F., and E.J. Wellin. 2009. Heterosis Decreasing in Hybrids: Yield Test Inbreds. Crop Science 49: 1969–1976.

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

Vigouroux, Y., J.C. Glaubitz, Y. Matsuoka, M.M. Goodman, J. Sánchez G, and J.F. Doebley. 2008. Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. American journal of botany 95: 1240–53.

Wassom, J.J., V. Mikkelineni, M.O. Bohn, and T.R. Rocheford. 2008. QTL for Fatty Acid Composition of Maize Kernel Oil in Illinois High Oil × B73 Backcross-Derived Lines. Crop Science 48: 69–78.

Yan, J., M.L. Warburton, and J.H. Crouch. 2011. Association mapping for enhancing maize (*Zea mays* L .) genetic improvement. Crop Science 51: 433–449.

Yang, X., Y. Guo, J. Yan, J. Zhang, T. Song, T. Rocheford, and J.-S. Li. 2010. Major and minor QTL and epistasis contribute to fatty acid compositions and oil concentration in high-oil maize. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik 120: 665–78.

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: 539–51.

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

Zheng, P., W.B. Allen, K. Roesler, M.E. Williams, S. Zhang, J. Li, K. Glassman, J. Ranch, D. Nubel, W. Solawetz, D. Bhattramakki, V. Llaca, S. Deschamps, G.-Y. Zhong, M.C. Tarczynski, and B. Shen. 2008. A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nature genetics 40: 367–72.

Zhu, C., M.A. Gore, E.S. Buckler, and J. Yu. 2008. Status and prospects of association mapping in plants. The Plant Genome Journal 1: 5–20.

Chapter 6

General discussion

The value of maize landraces as source of genetic diversity and of specific adaptations has long been recognized and discussed in previous studies (Gouesnard et al., 2005; Reif et al., 2005; Dubreuil et al., 2006). They were described for morphological properties, for their tolerance and resistance to abiotic (Peter et al., 2009a; b; Schneider et al., 2011) and biotic stress (Malvar et al., 2007). Moreover, their usefulness to improve the current breeding material was assessed (Reif et al., 2005; Revilla et al., 2006; Prigge et al., 2012). However, their use in breeding is still limited (Hoisington et al., 1999), mostly due to the presence of undesirable traits and deleterious genes in these materials not adapted to modern maize cropping, as well as to the large performance gap between landraces and modern hybrid varieties (Wilde et al., 2010).

In the present study, our aim was to use the advantages offered by the DH technology to get access to the phenotypic and genetic richness of landraces and make it available for research and breeding purposes. In the following, we will discuss, how far the availability of DH lines derived from landraces facilitates the exploitation of these genetic resources. We will mainly focus on two aspects: (i) the potential of DH lines derived from landraces to perform gene mining and (ii) their potential to improve the genetic diversity and performance of current elite European Flint germplasm.

Gene mining in DH lines derived from landraces

The advantages of DH lines derived from landraces for GWA mapping

As discussed by Strigens et al. (2013b), the production of libraries of inbred lines derived from landraces by selfing is very tedious. With the advent of the DH technology, a major breakthrough was achieved for the fast and efficient production of inbred lines from diverse materials (Deimling et al., 1997; Schmidt, 2003; Röber et al., 2005; Geiger and Gordillo, 2010; Prigge and Melchinger, 2012). It allowed us to produce between 31 and 65 DH lines from each of three landraces, whereas only few founder inbred lines were developed from a few landraces in the beginning of hybrid breeding in Europe. Yet, the even distribution of genetic distances among DH lines derived from single landrace suggested that additional DH lines might have been produced without re-sampling of the same haplotypes, underlining the huge genetic diversity available in the landraces (Strigens et al., 2013b). Further improvement of the DH technology (Melchinger et al., 2013) will certainly make it possible to produce hundreds of DH lines derived from landraces despite the lower rate of success for DH production in these materials (W. Schipprack, personal communication) and, thus, fix most of the diversity present in landraces in immortal homozygous lines.

Owing to the fact that the landraces underwent only moderate artificial selection over the centuries, DH lines derived from landraces should represent a random sample of rather unselected genes, except for the recessive lethal alleles lost during the haploid stage (Prigge et al., 2012) or those fixed by natural selection. Indeed, the huge phenotypic diversity observed among the DH lines derived from the three landraces suggested that we were able to recover a great part of their diversity in our populations of DH lines (Strigens et al., 2013b). In contrast to elite breeding germplasm, they harbored traits and properties eliminated during the past decades of modern maize breeding (Lauer et al., 2012) and showed all kind of extreme, unwanted or desired phenotypes (Chapter 5). Therefore, the contrast between genotypes may

have been as strong as in biparental mapping populations created with extreme parents, with the additional advantages of having a diverse genetic background and a faster decline of LD (Zhu et al., 2008; Strigens et al., 2013b).

This increased greatly the power of QTL detection of our GWA mapping approach in comparison to mapping in elite material and allowed us to identify numerous QTL and candidate genes for several agronomical traits (Chapter 5). Additionally, the higher resolution of GWA approaches in comparison to linkage mapping approaches increased the plausibility of the candidate genes identified (Strigens et al., 2013a). Further improvements in the marker coverage or re-sequencing approaches might enhance the resolution of GWA down to the causative mutation. Nevertheless, only cloning, silencing or over-expression studies would be able to confirm the validity of the proposed candidate genes, despite several of them co-located with previously reported QTL.

Use and limitations of GWA mapping in DH lines derived from landraces

To further increase the QTL detection power of GWA analysis, mapping populations larger than the present one would be required. Joint analysis of mapping panels, as done here, is a practical solution, but the positive effect of additional diversity might be counterbalanced by the strong population structure resulting from the admixture of different populations and heterotic pools (Chapter 5). Developing more DH lines from additional landraces would allow performing GWA in a single heterotic pool and, thus, eventually overcome the problems of population structure. In particular, larger mapping populations may allow for detection of rare QTL or such with smaller effects and, thus, mapping of highly polygenic traits. However, the practical use of such small effect QTL for marker assisted selection (MAS) might be limited, because breeders are rather interested in large effect QTL or, on the opposite, in direct assessment of the genotypic value of new lines by genomic prediction, taking into account all QTL effects (Meuwissen et al., 2001).

Nevertheless, performing GWA analyses in elite materials or libraries of DH lines derived from landraces allowed discovering new QTL alleles, as well as a better understanding of trait expression. Performing GWA with phenotypic data obtained in controlled environments or well monitored field conditions allowed us to detect interactions between QTL and environments, and gave us insights in the control of stress tolerance, early growth and plant morphology (Strigens et al., 2013a; Chapter 5). Especially, it revealed that genetic adaptation to environmental stresses can be achieved in different ways and that the resulting high genotype-by-environment interactions were partially explained by the frequent involvement of controlling and signaling genes in these responses (Strigens et al., 2013a). Further, it showed that the morphology of the plants was controlled by several distinct genes that lead to the same phenotype (Chapter 5). This might be the result of homologous genes with slightly different expression pattern (Kuusk et al., 2006; Danilevskaya et al., 2008), as commonly observed in maize (*e.g.*, plant coloration) or of epistatic interactions. Understanding of these mechanisms and identification of the key genes involved in trait expression can, thus, help selecting genotypes with the highest stress tolerance even without the necessity of tedious testing under controlled or field environments. Taking into account the redundancy of the maize genome or epistatic effects when performing MAS or genomic prediction can certainly improve the predictive power of such approaches.

Further prospects of GWA mapping in DH lines derived from landraces

It can be expected that traits not evaluated in this study may show a diversity of similar magnitude and that many additional useful properties might still be slumbering in our library of DH lines. It is, therefore, a great advantage to dispose of a collection of immortal homozygous lines that fix the phenotypic and genetic diversity of the original landrace (Reif et al., 2005; Strigens et al., 2013b). Individual genotypes can be evaluated for new traits, at different locations and under different conditions, without any changes in the genetic

composition of the studied subject. In comparison, open-pollinated landraces would give rise to new genotypes and allele combinations in each generation, and many interesting properties might remain hidden in the heterozygous plants. For example, the superior oil content observed in several DH lines derived from *Schindelmeiser* (Chapter 5), was not observed in the landrace itself despite of targeted selection for higher oil content (W. Schmidt, personal communication).

The current development of high throughput phenotyping platforms (Granier et al., 2006; Montes et al., 2011; Busemeyer et al., 2013), will greatly facilitate the evaluation of numerous traits and genotypes in diverse environments (Strigens et al., 2012, 2013a) and may reveal unexpected properties of the landraces. Development of databases for storage of all the morphological and physiological properties of the DH lines derived from landraces, would allow to dispatch the workload among institutes and phenotyping platforms, and to collect a very large spectrum of information on them. Access to this information for researchers and breeders would allow an efficient mining of information and might dramatically increase the use of the landraces, or DH lines derived from them, as genetic resources, because the lack of information on these materials would be overcome.

In summary, libraries of DH lines derived from landraces are a very powerful tool to identify new properties as well as new alleles and genes, owing to the large phenotypic and genotypic variation captured. Development of DH lines from additional landraces would be of great use to solve both the problems of population size and population structure, and allow very precise mapping of new genes by GWA analysis.

Broadening the genetic base of the European Flint germplasm

In addition to the advantages for GWA mapping described above, the DH lines derived from landraces are precious sources of genetic diversity that can be used to broaden the genetic base of the elite materials (Reif et al., 2005). The low LD within landraces and the even distribution of genetic distances between DH derived from the same landraces suggested a high effective population size (N_e) in our libraries of DH lines (Strigens et al., 2013b). First explorative approaches using the relation between N_e, LD and recombination rate described by (Sved, 1971) and successfully implemented in laying hens and cattle for estimation of N_e (Qanbari et al., 2010a; b) suggested that the N_e of the used landraces was much larger than that of the elite Flint population of the University of Hohenheim (data not shown). Consequently, introducing germplasm from European Flint landraces into elite breeding populations will definitely broaden the genetic base of the elite European Flint breeding material.

Broadening the genetic diversity by intogression of DH lines derived from landrace into the elite material instead of the landrace itself bears many advantages. First, owing to their complete homozygosis, superior DH lines or such ones carrying interesting QTL could be identified and directly used for breeding purposes (Strigens et al., 2013b; Chapter 5). Second, the production of DH lines from landraces should eliminate recessive lethal alleles, even if not directly observed at the phenotypic level (Strigens et al., 2013b). The precise identification of QTL and underlying genes by high resolution GWA analysis further allows targeted introgression of the desired properties by MAS or in combination with genomic prediction approaches. Known QTL might for example be introduced as fixed factors in the prediction models.

Further, introgression of DH lines derived from landraces adapted to the climatic conditions prevailing in Europe and showing for example superior early growth (Strigens et al., 2013b)

might be more efficient than introducing unadapted tropical or U.S. germplasm (Stamp, 1987; Reif et al., 2010). Owing to the relatively large yield gap between the elite material and the best DH lines derived from landraces (Strigens et al., 2013b), several backcrosses might be needed to bridge the performance gap. Classical selfing might then be preferred to DH production for line development in that case, to allow for more genetic recombination This bears the risk of breaking positive linkage groups selected in the elite material over the past decades, but it may also allow breaking of negative correlations such as the one between chilling tolerance and flowering time (Strigens et al., 2012, 2013a).

Conclusion

Several questions remain concerning the use of the DH method to produce lines from landraces: What is the effect of the DH method on the recovered diversity? How random is the selection of gametes that are surviving to the haploid and doubled haploid stage? Are there specific selective sweeps around genes responsible for a good aptitude to haploid induction and recognition? How large is the effect of the purging of lethal alleles on the recovered diversity? Are there long haplotypes blocks around the eliminated alleles? We could neither answer these questions on a phenotypic basis nor on a genetic basis, because no systematic morphological differences were observed between the original landraces and synthetic landraces produced by intermating the corresponding DH lines, and no genetic data was available for the original landraces (Strigens et al., 2013b). Nevertheless, with the advent of next generation sequencing method allowing the fast sequencing of pooled genotypes, efficient genotyping of the landraces themselves would become possible. This would allow estimation of allele frequencies in a large set of individuals from each landrace and, thus, quantification of changes in allele frequencies in DH lines derived from them, which would allow to monitor the purge of lethal alleles occurring at the haploid stage. Additionally, SNPs specific to the Flint germplasm and omitted in the construction of the MaizeSNP50 chip might be discovered. Especially rare alleles might be better represented with such approaches and the ascertainment bias of the MaizeSNP50 chip may be overcome (Frascaroli et al., 2012).

Nevertheless, the availability of DH lines derived from landraces greatly facilitates the selection of material or genes from the landraces that could be introduced into elite germplasm. It gives access to tremendous sources of new properties and allele combinations, allowing an efficient broadening of the genetic base of the elite material for future breeding success.

REFERENCES

Busemeyer, L., D. Mentrup, K. Möller, E. Wunder, K. Alheit, V. Hahn, H.P. Maurer, J.C. Reif, T. Würschum, J. Müller, F. Rahe, and A. Ruckelshausen. 2013. BreedVision--a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors (Basel, Switzerland) 13: 2830–2847.

Danilevskaya, O.N., X. Meng, Z. Hou, E. V Ananiev, and C.R. Simmons. 2008. A genomic and expression compendium of the expanded PEBP gene family from maize. Plant physiology 146: 250–264.

Deimling, S., F.K. Röber, and H.H. Geiger. 1997. Methodik und Genetik der Haploiden-Induktion bei Mais. Vortr. Pflanzenzüchtung 38: 203–224.

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.

Frascaroli, E., T. a Schrag, and A.E. Melchinger. 2012. 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. Theoretical and applied genetics.

Geiger, H.H., and G.A. Gordillo. 2010. Doubled haploids in hybrid maize breeding. Maydica 54: 485–499.

Gouesnard, B., J. Dallard, P. Bertin, A. Boyat, and A. Charcosset. 2005. European maize landraces: genetic diversity, core collection definition and methodology of use. Maydica 50: 115–234.

Granier, C., L. Aguirrezabal, K. Chenu, S.J. Cookson, M. Dauzat, P. Hamard, J. Thioux, G. Rolland, S. Bouchier-combaud, A. Lebaudy, B. Muller, T. Simonneau, and F. Tardieu. 2006. PHENOPSIS , an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. New Phytologist 169: 623–635.

Hoisington, D., M. Khairallah, T. Reeves, J.-M. Ribaut, B. Skovmand, S. Taba, and M.L. Warburton. 1999. Plant genetic resources: what can they contribute toward increased crop productivity? Proceedings of the National Academy of Sciences 96: 5937–5943

Kuusk, S., J.J. Sohlberg, D.M. Eklund, and E. Sundberg. 2006. Functionally redundant SHI family genes regulate *Arabidopsis* gynoecium development in a dose-dependent manner. The Plant Journal 47: 99–111.

Lauer, S., B.D. Hall, E. Mulaosmanovic, S.R. Anderson, B.K. Nelson, and S. Smith. 2012. Morphological changes in parental lines of Pioneer brand maize hybrids in the U. S. Central Corn Belt. Crop Science 52: 1033–1043.

Malvar, R.A., A. Butrón, A. Álvarez, G. Padilla, M. Cartea, P. Revilla, and A. Ordás. 2007. Yield performance of the European Union Maize Landrace Core Collection under multiple corn borer infestations. Crop Protection 26: 775–781.

Melchinger, A.E., W. Schipprack, T. Würschum, S. Chen, and F. Technow. 2013. Rapid and accurate identification of *in-vivo* induced haploid seeds based on oil content in maize. Scientific Reports 3: 2129, 1-5.

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.

Montes, J.M., F. Technow, B.S. Dhillon, F. Mauch, and A.E. Melchinger. 2011. High-throughput non-destructive biomass determination during early plant development in maize under field conditions. Field Crops Research 121: 268–273.

Peter, R., T.W. Eschholz, P. Stamp, and M. Liedgens. 2009a. Early growth of flint maize landraces under cool conditions. Crop Science 49: 169–178.

Peter, R., T.W. Eschholz, P. Stamp, and M. Liedgens. 2009b. Swiss Flint maize landraces—A rich pool of variability for early vigour in cool environments. Field Crops Research 110: 157–166.

Prigge, V., R. Babu, B. Das, M.H. Rodriguez, G.N. Atlin, and A.E. Melchinger. 2012. Doubled haploids in tropical maize: II. Quantitative genetic parameters for testcross performance. Euphytica 185: 453–463.

Prigge, V., and A.E. Melchinger. 2012. Production of haploids and doubled haploids in maize. *In* Loyola-Vargas, V., Ochoa-Alejo, N. (eds.), Plant cell culture protocols. 3rd ed. Humana Press - Springer Verlag, Totowa, New Jersey.

Qanbari, S., M. Hansen, S. Weigend, R. Preisinger, and H. Simianer. 2010a. Linkage disequilibrium reveals different demographic history in egg laying chickens. BMC Genetics 11: 103.

Qanbari, S., E.C.G. Pimentel, J. Tetens, G. Thaller, P. Lichtner, a R. Sharifi, and H. Simianer. 2010b. The pattern of linkage disequilibrium in German Holstein cattle. Animal genetics 41: 346–356.

Reif, J.C., S. Fischer, T.A. Schrag, K.R. Lamkey, D. Klein, B.S. Dhillon, H.F. Utz, and A.E. Melchinger. 2010. Broadening the genetic base of European maize heterotic pools with US Cornbelt germplasm using field and molecular marker data. Theoretical and Applied Genetics 120: 301–310.

Reif, J.C., S. Hamrit, M. Heckenberger, W. Schipprack, H. Peter Maurer, M. Bohn, and A.E. Melchinger. 2005. Genetic structure and diversity of European flint maize populations determined with SSR analyses of individuals and bulks. Theoretical and Applied Genetics 111: 906–913.

Revilla, P., A. Boyat, A. Álvarez, B. Gouesnard, B. Ordás, V.M. Rodríguez, A. Ordás, and R.A. Malvar. 2006. Contribution of autochthonous maize populations for adaptation to European conditions. Euphytica 152: 275–282.

Röber, F.K., G.A. Gordillo, and H.H. Geiger. 2005. *In vivo* haploid induction in maize - performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 50: 275–283.

Schmidt, W. 2003. Hybridmaiszüchtung bei der KWS SAAT AG. Bericht über die 54. Tagung 2003 der Vereinigung der Pflanzenzüchter und Saatgutkaufleute Österreichs BAL Gumpenstein: 1–6.

Schneider, D.N., N.M. Freitag, M. Liedgens, B. Feil, and P. Stamp. 2011. Early growth of field-grown swiss flint maize landraces. Maydica 56: 1702.

Stamp, P. 1987. Seedling Development of Adapted and Exotic Maize Genotypes at Severe Chilling Stress. Journal of Experimental Botany 38: 1336–1342.

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

Strigens, A., C. Grieder, B.I. Haussmann, and A.E. Melchinger. 2012. Genetic variation among inbred lines and testcrosses of maize for early growth parameters and their relationship to final dry matter yield. Crop Science 52: 1084–1092.

Strigens, A., W. Schipprack, J.C. Reif, and A.E. Melchinger. 2013b. Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding. PloS one 8: e57234.

Sved, J. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theoretical population biology 2: 125–141.

Wilde, K., H. Burger, V. Prigge, T. Presterl, W. Schmidt, M. Ouzunova, and H.H. Geiger. 2010. Testcross performance of doubled-haploid lines developed from European flint maize landraces. Plant Breeding 129: 181–185.

Zhu, C., M.A. Gore, E.S. Buckler, and J. Yu. 2008. Status and prospects of association mapping in plants. The Plant Genome Journal 1: 5–20.

Chapter 7

Summary

Since the introduction of maize into Europe by Colombus in 1492 and further discoverers in the 16th century, open-pollinated varieties of flint maize were cultivated across the continent. Natural selection promoted adaptation to the climatic conditions prevailing in the different regions. With the advent of hybrid breeding in Europe during the 1950's, some of the genes and alleles responsible for the specific adaptations of the landraces to abiotic and biotic stress were captured in the first developed inbred lines, but most of their genetic diversity is still untapped. Development of inbred lines out of this material by recurrent selfing is very tedious due to strong inbreeding depression. In contrast, the doubled-haploid (DH) technology allows producing fully homozygous lines out of landraces in only one step. This allows their precise characterization in replicated trials and identification of new genes by genome wide association (GWA) mapping.

In this study we genotyped a set of 132 DH lines derived from European Flint landraces and 364 elite European flint (EU-F), European dent (EU-D) and North-American dent (NA-D) inbred lines with 56,110 single nucleotide polymorphism (SNP) markers. The lines were evaluated in field trials for morphologic and agronomic traits and GWA mapping was performed to identify underlying quantitative trait loci (QTL). In particular, our objectives were to (1) develop a robust method for quantifying early growth with a non-destructive remote-sensing platform, (2) evaluate the importance of early growth performance of inbred lines with regard to their testcross performance, (3) determine the potential of GWA mapping to identify genes underlying early growth and cold tolerance related traits, (4) evaluate the phenotypic and genotypic diversity recovered in the DH lines derived from the landraces, (5) estimate the effect of the DH method on the recovered genetic diversity, (6) identify new genes by GWA mapping in the DH lines derived from landraces, and (8) discuss the potential of DH lines derived from landraces to improve the genetic diversity and performance of elite maize germplasm.

A phenotyping platform using spectral reflectance and light curtains was used to perform repeated measurements of biomass and estimate relative growth rates (RGR) of the DH and inbred lines, as well as of two testcrosses of 300 dent inbred lines. Heritability (h^2) of RGR was high ($h^2 = 0.88$) for line *per se* performance and moderate ($h^2 = 0.79$) for testcross performance in 2008 and 2009,

and somewhat lower (h^2 = 0.70) for line *per se* performance in 2010. The DH lines derived from the landraces *Schindelmeiser* and *Gelber Badischer* had the highest RGR followed by EU-F lines, DH lines derived from *Bugard,* EU-D lines and, finally, NA-D lines. For inbred lines, whole plant dry matter yield (DMY) was positively correlated with RGR (r_g = 0.49), whereas this relation was weaker in the testcrosses (r_g = 0.29). RGR of the inbred lines correlated with RGR of their testcrosses (r_g = 0.42), but it had no influence on testcross DMY.

A set of 375 EU-F, EU-D and NA-D lines were further evaluated in growth chambers under chilling (16/13°C) and optimal (27/25°C) temperatures. Photosynthetic and early growth performance were estimated for each treatment and an adaptation index (AI) built as the chilling to optimal performance ratio. In EU-D and EU-F lines, RGR was correlated with leaf area, shoot and leaf dry weight measured under chilling temperatures. Nineteen QTL were identified by GWA mapping for trait performance, calculated AI and RGR. Candidate genes involved in ethylene signaling, brassinolide, and lignin biosynthesis were found in their vicinity. Several QTL for photosynthetic performance co-located with previously reported QTL and the QTL identified for shoot dry wieght under optimal conditions co-located with a QTL for RGR. The frequent involvement of candidate genes into signaling or regulation underlines the complex response of photosynthetic performance and early growth to climatic conditions, and supports pleiotropism as a major cause of QTL co-locations.

Comparison of the DH lines derived from landraces with the EU-F lines showed that genotypic variances in single DH populations were greater than in the EU-F breeding population. A high average genetic distance among the DH lines derived from the same landrace as well as a rapid decay of linkage disequilibrium suggests a high effective population size of the landraces. Because no systematic phenotypic differences were observed between the landraces and synthetic landraces obtained by intermating the corresponding DH lines, the expected purge of lethal recessive alleles during the DH production did neither improve grain yield performance nor affect the recovered genetic diversity. Performing GWA in the DH lines derived from landraces as well as the EU-F, and EU-D lines allowed the identification of 49 QTL for 27 traits. A larger set of DH lines derived from more landraces might solve problems arising from population structure and allow a much higher power for the detection of new alleles.

In conclusion, the introgression of DH lines derived from landraces into the elite breeding material would strongly broaden its genetic base. However, grain yield performance was 22% higher in EU-F lines than in the DH lines derived from landraces. Selection of the best DH lines would allow partially bridging this yield gap and marker-assisted selection may allow introgression of positive QTL without introducing negative features by linkage drag.

Chapter 7

Zusammenfassung

Seit der Einfuhr von Mais aus der „neuen“ Welt nach Europa durch Kolumbus im Jahr 1492 und weitere Entdecker im 16. Jahrhundert, wurden offen abblühende Flint-Mais Populationen auf dem gesamten Kontinent angebaut. Durch natürliche Selektion passten sich diese Landsorten an die verschiedenen Klimate des Kontinents an. In den Anfängen der Hybridzüchtung während der 1950er Jahre wurden Gene und Allele, die für diese spezifische Anpassung an biotische und abiotische Stressfaktoren verantwortlich sind, in den ersten Inzuchtlinien nur teilweise fixiert. Der Grossteil der genetischen Vielfalt der Landsorten blieb jedoch ungenutzt, da die Entwicklung von Inzuchtlinien aus diesem Material wegen besonders starker Inzuchtdepression sehr mühsam ist. Demgegenüber erlaubt es die seit etwa 10 Jahre eingesetzte Methode der Erzeugung von Doppel-Haploiden (DH), vollständig homozygote Linien aus Landsorten in einem einzigen Schritt zu entwickeln. Diese DH-Linien können in wiederholten Feldversuchen sehr präzise evaluiert werden. Dies vereinfacht die Kartierung von Genen mithilfe der Genom-weiten Assoziations-Kartierung (GWA) enorm.

In der vorliegenden Studie wurden 132 DH-Linien aus europäischen Landsorten, 364 Inzucht-linien aus Nordamerikanischem Dent (NA-D), europäischem Flint (EU-F) und europäischem Dent (EU-D) Zuchtmaterial mit 56110 genetischen Markern genotypisiert. Agronomische Eigenschaften der DH-Linien und Elite-Inzuchtlinien wurden in Feldversuchen evaluiert und mittels GWA kartiert, um vorteilhafte Gene zu identifizieren. Zu unseren Zielen gehörten insbesondere (1) die Entwicklung einer robusten, nicht-destruktiven Methode zur Erfassung der Jugendentwicklung mittels Sensoren, (2) die Untersuchung des Zusammenhangs zwischen der Jugendentwicklung der Linien *per se* und deren Testkreuzungen, (3) die Erforschung von GWA zur Identifikation von Kühletoleranz- und Jugendentwicklungs-Genen in Elite-Inzuchtlinien, (4) die Evaluierung der aus den Landsorten mittels der DH-Methode geborgene phänotypische und genetische Vielfalt, (5) die Abschätzung eines möglichen Einfluss der DH-Methode auf der genetischen Vielfalt der DH-Linien, (6) die Entdeckung neuer Gene in den DH-Linien aus Landsorten mittels GWA, und (7) die Ermittlung des Potentials von DH-Linien aus Landsorten, um die Leistung und genetische Diversität des modernen Zuchtmaterials zu verbessern.

Die Biomasse und relative Wachstumsrate (RGR) der DH-Linien und Elite-Inzuchtlinien sowie je zwei Testkreuzungen von 300 Dent Inzuchtlinien wurden mit Lichtschranken und spektraler Reflektion geschätzt. Die Heritabilität (h^2) von RGR war hoch (h^2 = 0.88) für die *per se* Leistung der Linien und moderat (h^2 = 0.79) für die Testkreuzungsleistung in drei-ortigen Feldexperimenten in den Jahren 2008 und 2009. Etwas tiefer war diese für *per se* Leistung der Linien (h^2 = 0.70) in fünf-ortigen Feldexperimenten im Jahr 2010. Die DH-Linien aus den Landsorten *Schindelmeiser* und *Gelber Badischer* wiesen die höchste RGR auf, gefolgt von EU-F Linien, DH-Linien aus *Bugard*, EU-D Linien und zuletzt NA-D Linien. Die Gesamttrockenmasse der Linien war mit deren RGR positiv korreliert (r_g = 0.49), während diese Korrelationen für die Testkreuzungen schwächer ausfiel (r_g = 0.29). Die RGR der Linien korrelierte mit der RGR der Testkreuzungen (r_g = 0.42), hatte jedoch keinen Einfluss auf deren Gesamttrockenmasse.

Ein Satz von 375 EU-F, EU-D und NA-D Linien wurde unter kühlen (16/13°C) und optimalen (27/25°C) Temperaturen in Klimakammern untersucht. Die photosynthetische Leistung und die Jugendentwicklung wurden für jedes Verfahren gemessen. Aus dem Verhältnis der Leistungen unter kühlen und optimalen Bedingungen wurde ein Adaptations-Index (AI) berechnet. Für EU-F und EU-D Linien korrelierten Blattfläche, Blatt- und Sprossmasse unter kühlen Bedingungen mit RGR auf dem Feld. Neunzehn Genorte (QTL = qantitative trait loci) wurden für photosynthetische Leistung, AI und RGR mittels GWA identifiziert. Gene mit Beteiligung in der Äthylen-Signalkette, Brassinolid- und Lignin-Biosynthese wurden als Kandidaten identifiziert. Mehrere QTL für photosynthetische Leistung co-lokalisierten mit bereits beschriebenen QTL. Die häufige Beteiligung der Kandidatengene in Signalketten und Regulierung unterstreicht die Komplexität der Anpassung photosynthetischer Leistung und Jugendentwicklung an die Temperatur. Dies unterstützt die Hypothese von Pleiotropie als eine der Hauptursachen der Kolokalisierung von QTL.

Der Vergleich der genetischen Varianzen zeigte, dass diese innerhalb der einzelnen Landsorten grösser ist als innerhalb des EU-F Zuchtmaterials. Sowohl die hohe mittlere genetische Distanz zwischen den DH-Linien einer Landsorte, als auch das rasch abfallende Kopplungs-ungleichgewicht innerhalb der Landsorten deuten auf eine grosse Effektive Populationsgrösse hin. Die erwartete Eliminierung von rezessiven letalen Allelen durch die DH-Methode konnte den Ertrag synthetischer Landsorten nicht erhöhen und hatte auch keinen grossen Einfluss auf die genetische Diversität, da keine systematischen phänotypischen Änderungen zwischen den Landsorten und re-synthetisierten Landsorten zu beobachten waren. Mittels GWA Analyse in den DH-Linien aus Landsorten und in Elite-Inzuchtlinien konnten 49 QTL für 27 Merkmale kartiert werden. Eine grössere Anzahl von DH-Linien aus Landsorten würde es erlauben, die durch

Populationsstruktur verursachten Artefakte zu beseitigen und somit die Wahrscheinlichkeit, neue Allele zu entdecken, stark erhöhen.

Zusammengefasst kann die genetische Diversität des Zuchtmaterials durch die Einkreuzung von DH-Linien aus Landsorten stark erhöht werden. Der grosse Abstand zwischen der Leistung des Zuchtmaterials und den DH-Linien aus Landsorten (22%) kann durch Selektion der besten DH-Linien teilweise ausgeglichen werden. Marker-gestützte Selektion könnte das Einkreuzen von positiven QTL ohne Introgression von unerwünschten negativen Eigenschaften erleichtern.

Acknowledgements

First of all, I want to thank my academic supervisor Prof. Dr. A.E. Melchinger, who gave me the opportunity to do my PhD on such a great topic. I am very grateful for his patience, advice and, especially, his great support during the writing of the publications and the present thesis.

Many thanks also go to Prof. Dr. J.C. Reif for the inspiring discussions and great advice, to Prof. Dr. H.F. Utz for his support on statistical issues, and to Dr. B.I.G. Haussmann, Dr. T Schrag and Dr. U.K. Posselt for their corrections and suggestions on the manuscripts.

Special thanks go to Dr. E. Orsini and Dr. R. Messmer, who initiated and supervised the fructuous collaboration between the Group of Plant Breeding of the University of Hohenheim, Stuttgart, and the Group for Crop Sciences of the Federal Institute of Technology, Zürich, for the research on early growth. In this regard, many thanks also go to Dr. N.M. Freitag and X. Gilbert for the tedious phenotyping in the growth chambers!

I am greatly indebted to Dr. T. Presterl and KWS SAAT AG, Einbeck, who provided the heart of this research: the DH lines derived from landraces; as well as to Dr. W. Schipprack, T. Schmid, R. Volkhausen, and R. Lutz for the lines multiplication, building of the synthetics, field trial supervision and logistical support in Eckartsweier. Warm thanks go also to J. Jesse, F. Mauch, H. Pöschel and all the staff of the research stations for the realization of the field trials and their technical support.

I thank D. Neumeister, V. Klotz, N. Münch and K. Hütter for their contribution to this thesis during their Diploma, Master and Bachelor thesis. They were, together with the multitude of Hiwis and trainees, a great help for the phenotyping of the field trials and of the ~15500 collected ears. Thanks also to all my colleagues at the Institute who also came for support when the Hiwis were gone! In general, thanks to all the people on the 1st floor, 2nd floor and 3rd floor, as well as from other institutes for the great discussions, ideas and contributions to this work, for the discovery of Stuttgart and surroundings, for the fun at sports and apéros. Special thanks go to Dr. C. Grieder, for his precious collaboration on publication 1, and for being a very good sparring partner. And not the least, thanks to H. Kösling, S. Meyer, M. Lieb, B. Devezi-Savula for their precious support for the administrative stuff, as well as to S. Boger and B. Schilling for the lab work.

I thank the German Research Foundation (DFG; RE2254/3-1) as well as the Bundesministerium für Bildung und Forschung (BMBF; FK 0315045B, FK 03115461A, FK 0315528D) for the financial support of this work and data acquisition in the frame of several projects.

Finally I thank F. Müller and my family, who encouraged me all the time and endured these four years of maize research from the other side.

Curriculum Vitae

Name:	Alexander Carl Georg Strigens
Date and place of birth:	11 August 1983, Wiesbaden, Germany

School Education

1990 - 1995	Ecole primaire de Blonay, Switzerland
1995 - 1999	Ecole secondaire de la Tour de Peilz, Switzerland
1999 - 2002	Gymnase de Burier, Biology & Chemistry, Switzerland

University Education

2003 - 2008	B.Sc. in Agricultural sciences, Eidgenössische Technische Hochschule, Zürich, Switzerland
2008 - 2009	M.Sc. in Agricultural sciences, Eidgenössische Technische Hochschule, Zürich, Switzerland

Employment record

09/2009 – 09/2012	Doctorate candidate in Applied Genetics and Plant Breeding (Prof. Dr. A. E. Melchinger), University of Hohenheim, Stuttgart, Germany
Since 10/2012	Maize Breeder, Delley Semences et Plantes SA, Delley, Switzerland

_____

Alexander Carl Georg Strigens

Hohenheim, 31.10.2013

Erklärung

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

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

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

_____

Alexander Carl Georg Strigens

Hohenheim, 31.10.2013

Annexes

Annex 1. Name, heterotic pool and population of the genotypes evaluated in Chapter 5.

Annex 2. Mean, genetic variance (σ^2_g), genotype-by-environment interaction variance ($\sigma^2_{g\times e}$), and residual variance (σ^2_ε) within elite European dent (EU-D) and flint (EU-F) inbred lines as well as within the set of 132 doubled-haploid (DH) lines derived from three landraces (LR-DH) for each trait measured on at least four locations in 2010.

Annex 3. Genome wide association scans for single nucleotide polymorphism (SNP) × trait associations detected in Set 3 with a model correcting for population structure using the kinship matrix and five first principal coordinates from the principal coordinate analysis performed on the marker data. Left hand: The $-\log_{10}(P)$ values from the genome wide scan are plotted against the SNP position on the physical map of each chromosome, for each trait × treatment combination. Right hand: QQ-plot of expected against observed *P* values for SNP × trait associations, and corresponding inflation factor λ. The horizontal line shows the significance threshold ($\alpha = 0.05$) after Bonferroni-correction for multiple comparison.

Annex 4. Position within chromosome (Chr) and QTL assignment of single nucleotide polymorphism (SNP) significantly associated with trait expression in a mapping population composed of 132 doubled-haploid (DH) lines derived from three landraces and elite European dent and flint inbred lines, as well as gene model and putative functions associated to each SNP.

Annex 5. Frequency within population of the positive allele at each marker within quantitative trait loci detected for agronomic and morphological traits in the mapping panel composed of elite European dent (EU-D) and flint (EU-F) inbred lines as well as of doubled haploid (DH) lines derived from the landraces *Bugard* (DH-BU), *Gelber Badischer* (DH-GB), and *Schindelmeiser* (DH-SC).

Annex 1. Name, heterotic pool and population of the genotypes evaluated in Chapter 5.

Genotype	Heterotic pool	Population
A188	Dent	NA-D
BAREILLES-002	Flint	EU-F
BAREILLES-005	Flint	EU-F
BAREILLES-017	Flint	EU-F
BUGARD_A-DH003	Flint	DH-BU
BUGARD_A-DH005	Flint	DH-BU
BUGARD_A-DH006	Flint	DH-BU
BUGARD_A-DH011	Flint	DH-BU
BUGARD_A-DH013	Flint	DH-BU
BUGARD_A-DH014	Flint	DH-BU
BUGARD_A-DH015	Flint	DH-BU
BUGARD_A-DH017	Flint	DH-BU
BUGARD_A-DH019	Flint	DH-BU
BUGARD_A-DH023	Flint	DH-BU
BUGARD_A-DH024	Flint	DH-BU
BUGARD_A-DH028	Flint	DH-BU
BUGARD_A-DH032	Flint	DH-BU
BUGARD_A-DH034	Flint	DH-BU
BUGARD_A-DH036	Flint	DH-BU
BUGARD_A-DH037	Flint	DH-BU
BUGARD_A-DH040	Flint	DH-BU
BUGARD_A-DH042	Flint	DH-BU
BUGARD_A-DH043	Flint	DH-BU
BUGARD_A-DH046	Flint	DH-BU
BUGARD_A-DH048	Flint	DH-BU
BUGARD_A-DH050	Flint	DH-BU
BUGARD_A-DH058	Flint	DH-BU
BUGARD_A-DH059	Flint	DH-BU
BUGARD_A-DH062	Flint	DH-BU
BUGARD_A-DH063	Flint	DH-BU
BUGARD_A-DH064	Flint	DH-BU
BUGARD_A-DH065	Flint	DH-BU
BUGARD_A-DH068	Flint	DH-BU
BUGARD_A-DH070	Flint	DH-BU
BUGARD_A-DH073	Flint	DH-BU
BUGARD_A-DH074	Flint	DH-BU
BUGARD_A-DH075	Flint	DH-BU
BUGARD_A-DH077	Flint	DH-BU
BUGARD_A-DH083	Flint	DH-BU
BUGARD_A-DH084	Flint	DH-BU
CL30	Dent	NA-D
CM105	Dent	NA-D
Co125	Dent	NA-D
D102	Flint	EU-F
D107	Flint	EU-F
D118	Flint	EU-F

Annex 1 (continued).

Genotype	Heterotic pool	Population
D140	Flint	EU-F
D143	Flint	EU-F
D147	Flint	EU-F
D149	Flint	EU-F
D150	Flint	EU-F
D152	Flint	EU-F
D157	Flint	EU-F
D164	Flint	EU-F
D167	Flint	EU-F
D171	Flint	EU-F
D305	Flint	EU-F
D32	Dent	EU-D
D408	Dent	EU-D
D503	Flint	FLINT
D504	Flint	FLINT
D60	Dent	EU-D
D66	Dent	EU-D
D67	Dent	EU-D
D800	Flint	EU-F
DK105	Flint	EU-F
EP1	Flint	EU-F
F005	Flint	EU-F
F011	Flint	EU-F
F012	Flint	EU-F
F013	Flint	EU-F
F016	Flint	EU-F
F018	Flint	EU-F
F020	Flint	EU-F
F023	Flint	EU-F
F027	Flint	EU-F
F030	Flint	EU-F
F034	Flint	EU-F
F035	Flint	EU-F
F037	Flint	EU-F
F038	Flint	EU-F
F039	Flint	EU-F
F040	Flint	EU-F
F043	Flint	EU-F
F045	Flint	EU-F
F047	Flint	EU-F
F048	Flint	EU-F
F050	Flint	EU-F
F052	Flint	EU-F
F054	Flint	EU-F
F055	Flint	EU-F
F056	Flint	EU-F

Annex 1 (continued).

Genotype	Heterotic pool	Population
F057	Flint	EU-F
F058	Flint	EU-F
F059	Flint	EU-F
F060	Flint	EU-F
F061	Flint	EU-F
F062	Flint	EU-F
F066	Flint	EU-F
F068	Flint	EU-F
F070	Flint	EU-F
F072	Flint	EU-F
F073	Flint	EU-F
F074	Flint	EU-F
F077	Flint	EU-F
F082	Flint	EU-F
F084	Flint	EU-F
F087	Flint	EU-F
F088	Flint	EU-F
F090	Flint	EU-F
F093	Flint	EU-F
F094	Flint	EU-F
F096	Flint	EU-F
F098	Flint	EU-F
F099	Flint	EU-F
F101	Flint	EU-F
F103	Flint	EU-F
F104	Flint	EU-F
F105	Flint	EU-F
F106	Flint	EU-F
F108	Flint	EU-F
F109	Flint	EU-F
F110	Flint	EU-F
F124	Flint	EU-F
F2	Flint	EU-F
F7	Flint	EU-F
FF067-n-52-3-2-1-n	Flint	EU-F
FF067-n-7-1-1-1-n	Flint	EU-F
FF084-n-10-1-1-1	Flint	EU-F
FV271	Dent	EU-D
GELBER_BADISCHER-DH102	Flint	DH-GB
GELBER_BADISCHER-DH106	Flint	DH-GB
GELBER_BADISCHER-DH109	Flint	DH-GB
GELBER_BADISCHER-DH110	Flint	DH-GB
GELBER_BADISCHER-DH113	Flint	DH-GB
GELBER_BADISCHER-DH114	Flint	DH-GB
GELBER_BADISCHER-DH115	Flint	DH-GB
GELBER_BADISCHER-DH116	Flint	DH-GB

Annex 1 (continued).

Genotype	Heterotic pool	Population
GELBER_BADISCHER-DH119	Flint	DH-GB
GELBER_BADISCHER-DH120	Flint	DH-GB
GELBER_BADISCHER-DH121	Flint	DH-GB
GELBER_BADISCHER-DH122	Flint	DH-GB
GELBER_BADISCHER-DH123	Flint	DH-GB
GELBER_BADISCHER-DH124	Flint	DH-GB
GELBER_BADISCHER-DH125	Flint	DH-GB
GELBER_BADISCHER-DH127	Flint	DH-GB
GELBER_BADISCHER-DH130	Flint	DH-GB
GELBER_BADISCHER-DH131	Flint	DH-GB
GELBER_BADISCHER-DH203	Flint	DH-GB
GELBER_BADISCHER-DH204	Flint	DH-GB
GELBER_BADISCHER-DH206	Flint	DH-GB
GELBER_BADISCHER-DH209	Flint	DH-GB
GELBER_BADISCHER-DH210	Flint	DH-GB
GELBER_BADISCHER-DH211	Flint	DH-GB
GELBER_BADISCHER-DH212	Flint	DH-GB
GELBER_BADISCHER-DH213	Flint	DH-GB
GELBER_BADISCHER-DH215	Flint	DH-GB
GELBER_BADISCHER-DH216	Flint	DH-GB
GELBER_BADISCHER-DH217	Flint	DH-GB
GELBER_BADISCHER-DH219	Flint	DH-GB
GELBER_BADISCHER-DH220	Flint	DH-GB
L005	Flint	EU-F
L007	Flint	EU-F
L012	Flint	EU-F
L016	Flint	EU-F
L017	Flint	EU-F
L019	Flint	EU-F
L023	Flint	EU-F
L024	Flint	EU-F
L025	Flint	EU-F
L032	Flint	EU-F
L035	Flint	EU-F
L037	Flint	EU-F
L041	Flint	EU-F
L045	Flint	EU-F
L046	Flint	EU-F
L047	Flint	EU-F
L048	Flint	EU-F
L050	Flint	EU-F
L051	Flint	EU-F
L054	Flint	EU-F
L056	Flint	EU-F
L057	Flint	EU-F
L058	Flint	EU-F

Annex 1 (continued).

Genotype	Heterotic pool	Population
L059	Flint	EU-F
L060	Flint	EU-F
LACAUNE-002	Flint	EU-F
LACAUNE-004	Flint	EU-F
LACAUNE-005	Flint	EU-F
LACAUNE-006	Flint	EU-F
LAURENT_DE_NESTE-002	Flint	EU-F
M012	Dent	EU-D
P001	Dent	EU-D
P006	Dent	EU-D
P009	Dent	EU-D
P024	Dent	EU-D
P029	Dent	EU-D
P033	Dent	EU-D
P034	Dent	EU-D
P036	Dent	EU-D
P038	Dent	EU-D
P040	Dent	EU-D
P042	Dent	EU-D
P045	Dent	EU-D
P046	Dent	EU-D
P047	Dent	EU-D
P048	Dent	EU-D
P053	Dent	EU-D
P060	Dent	EU-D
P063	Dent	EU-D
P064	Dent	EU-D
P065	Dent	EU-D
P066	Dent	EU-D
P068	Dent	EU-D
P069	Dent	EU-D
P070	Dent	EU-D
P071	Dent	EU-D
P072	Dent	EU-D
P074	Dent	EU-D
P075	Dent	EU-D
P079	Dent	EU-D
P080	Dent	EU-D
P081	Dent	EU-D
P083	Dent	EU-D
P084	Dent	EU-D
P085	Dent	EU-D
P086	Dent	EU-D
P087	Dent	EU-D
P092	Dent	EU-D
P093	Dent	EU-D

Annex 1 (continued).

Genotype	Heterotic pool	Population
P094	Dent	EU-D
P095	Dent	EU-D
P096	Dent	EU-D
P097	Dent	EU-D
P099	Dent	EU-D
P100	Dent	EU-D
P101	Dent	EU-D
P102	Dent	EU-D
P104	Dent	EU-D
P105	Dent	EU-D
P106	Dent	EU-D
P107	Dent	EU-D
P108	Dent	EU-D
P110	Dent	EU-D
P111	Dent	EU-D
P113	Dent	EU-D
P115	Dent	EU-D
P118	Dent	EU-D
P120	Dent	EU-D
P122	Dent	EU-D
P123	Dent	EU-D
P127	Dent	EU-D
P128	Dent	EU-D
P129	Dent	EU-D
P130	Dent	EU-D
P131	Dent	EU-D
P133	Dent	EU-D
P135	Dent	EU-D
P136	Dent	EU-D
P140	Dent	EU-D
P148	Dent	EU-D
P149	Dent	EU-D
P150	Dent	EU-D
P154	Dent	EU-D
P159	Dent	EU-D
P165	Dent	EU-D
P167	Dent	EU-D
P182	Dent	EU-D
P184	Dent	EU-D
P188	Dent	EU-D
P194	Dent	EU-D
P197	Dent	EU-D
P202	Dent	EU-D
P204	Dent	EU-D
P206	Dent	EU-D
P209	Dent	EU-D

Annex 1 (continued).

Genotype	Heterotic pool	Population
P210	Dent	EU-D
P211	Dent	EU-D
PIED_DE_PORTE-001	Flint	EU-F
PIED_DE_PORTE-005	Flint	EU-F
PS065-2-2-3-2-2-n	Dent	EU-D
PS081-n-52-2-1-n	Dent	EU-D
S002	Dent	EU-D
S015	Dent	EU-D
S016	Dent	EU-D
S018	Dent	EU-D
S020	Dent	EU-D
S025	Dent	EU-D
S028	Dent	EU-D
S033	Dent	EU-D
S035	Dent	EU-D
S036	Dent	EU-D
S037	Dent	EU-D
S040	Dent	EU-D
S044	Dent	EU-D
S046	Dent	EU-D
S048	Dent	EU-D
S049	Dent	EU-D
S050	Dent	EU-D
S051	Dent	EU-D
S052	Dent	EU-D
S058	Dent	EU-D
S064	Dent	EU-D
S065	Dent	EU-D
S066	Dent	EU-D
S067	Dent	EU-D
S069	Dent	EU-D
S070	Dent	EU-D
S072	Dent	EU-D
S074	Dent	EU-D
S077	Dent	EU-D
SCHINDELMEISER-DH102	Flint	DH-SC
SCHINDELMEISER-DH103	Flint	DH-SC
SCHINDELMEISER-DH104	Flint	DH-SC
SCHINDELMEISER-DH105	Flint	DH-SC
SCHINDELMEISER-DH106	Flint	DH-SC
SCHINDELMEISER-DH107	Flint	DH-SC
SCHINDELMEISER-DH108	Flint	DH-SC
SCHINDELMEISER-DH109	Flint	DH-SC
SCHINDELMEISER-DH112	Flint	DH-SC
SCHINDELMEISER-DH113	Flint	DH-SC
SCHINDELMEISER-DH115	Flint	DH-SC

Annex 1 (continued).

Genotype	Heterotic pool	Population
SCHINDELMEISER-DH116	Flint	DH-SC
SCHINDELMEISER-DH117	Flint	DH-SC
SCHINDELMEISER-DH118	Flint	DH-SC
SCHINDELMEISER-DH119	Flint	DH-SC
SCHINDELMEISER-DH120	Flint	DH-SC
SCHINDELMEISER-DH121	Flint	DH-SC
SCHINDELMEISER-DH122	Flint	DH-SC
SCHINDELMEISER-DH124	Flint	DH-SC
SCHINDELMEISER-DH125	Flint	DH-SC
SCHINDELMEISER-DH126	Flint	DH-SC
SCHINDELMEISER-DH128	Flint	DH-SC
SCHINDELMEISER-DH129	Flint	DH-SC
SCHINDELMEISER-DH132	Flint	DH-SC
SCHINDELMEISER-DH133	Flint	DH-SC
SCHINDELMEISER-DH134	Flint	DH-SC
SCHINDELMEISER-DH136	Flint	DH-SC
SCHINDELMEISER-DH137	Flint	DH-SC
SCHINDELMEISER-DH138	Flint	DH-SC
SCHINDELMEISER-DH141	Flint	DH-SC
SCHINDELMEISER-DH142	Flint	DH-SC
SCHINDELMEISER-DH143	Flint	DH-SC
SCHINDELMEISER-DH144	Flint	DH-SC
SCHINDELMEISER-DH145	Flint	DH-SC
SCHINDELMEISER-DH146	Flint	DH-SC
SCHINDELMEISER-DH147	Flint	DH-SC
SCHINDELMEISER-DH152	Flint	DH-SC
SCHINDELMEISER-DH154	Flint	DH-SC
SCHINDELMEISER-DH156	Flint	DH-SC
SCHINDELMEISER-DH157	Flint	DH-SC
SCHINDELMEISER-DH158	Flint	DH-SC
SCHINDELMEISER-DH161	Flint	DH-SC
SCHINDELMEISER-DH163	Flint	DH-SC
SCHINDELMEISER-DH164	Flint	DH-SC
SCHINDELMEISER-DH166	Flint	DH-SC
SCHINDELMEISER-DH168	Flint	DH-SC
SCHINDELMEISER-DH169	Flint	DH-SC
SCHINDELMEISER-DH170	Flint	DH-SC
SCHINDELMEISER-DH203	Flint	DH-SC
SCHINDELMEISER-DH206	Flint	DH-SC
SCHINDELMEISER-DH208	Flint	DH-SC
SCHINDELMEISER-DH209	Flint	DH-SC
SCHINDELMEISER-DH212	Flint	DH-SC
SCHINDELMEISER-DH213	Flint	DH-SC
SCHINDELMEISER-DH215	Flint	DH-SC
SCHINDELMEISER-DH216	Flint	DH-SC
SCHINDELMEISER-DH220	Flint	DH-SC

Annex 1 (continued).

Genotype	Heterotic pool	Population
SCHINDELMEISER-DH221	Flint	DH-SC
SCHINDELMEISER-DH222	Flint	DH-SC
SCHINDELMEISER-DH223	Flint	DH-SC
SCHINDELMEISER-DH225	Flint	DH-SC
SCHINDELMEISER-DH228	Flint	DH-SC
SCHINDELMEISER-DH238	Flint	DH-SC
SCHINDELMEISER-DH244	Flint	DH-SC
SCHINDELMEISER-DH247	Flint	DH-SC
STRENZFELDER-001	Flint	EU-F
STRENZFELDER-002	Flint	EU-F
STRENZFELDER-005	Flint	EU-F
STRENZFELDER-007	Flint	EU-F
STRENZFELDER-008	Flint	EU-F
STRENZFELDER-011	Flint	EU-F
STRENZFELDER-013	Flint	EU-F
STRENZFELDER-016	Flint	EU-F
VACQUIERS-DH053	Flint	EU-F
VACQUIERS-DH065	Flint	EU-F
VIEY-001	Flint	EU-F
VIEY-003	Flint	EU-F

Annex 2. Mean, genetic variance (σ^2_g), genotype-by-environment interaction variance ($\sigma^2_{g\times e}$), and residual variance (σ^2_ε) within elite European dent (EU-D) and flint (EU-F) inbred lines as well as within the set of 132 doubled-haploid (DH) lines derived from three landraces (LR-DH) for each trait measured on at least four locations in 2010.

Trait†	**Population**	**Mean**‡	σ^2_g	$\sigma^2_{g\times e}$	σ^2_ε
EMER [1-9]§	EU-D	3.66b	0.95	0.28	0.55
	EU-F	3.29a	0.89	0.57	0.55
	LR-DH	3.87b	1.22	0.36	0.55
GERM [%]	EU-D	60.11b	121.81	44.16	48.88
	EU-F	64.63a	107.70	44.55	48.88
	LR-DH	57.76b	156.11	34.07	48.88
CHLO [1-9]¶	EU-D	4.20c	0.22	0.48	0.55
	EU-F	3.09a	0.31	0.28	0.55
	LR-DH	3.49b	0.72	0.57	0.55
$EVIG_4$ [1-9]§	EU-D	5.07c	0.41	0.14	0.45
	EU-F	4.51b	0.83	0.43	0.45
	LR-DH	4.08a	1.54	0.32	0.45
$EVIG_8$ [1-9]§	EU-D	4.81b	0.28	0.15	0.32
	EU-F	4.00a	0.60	0.34	0.32
	LR-DH	3.79a	1.21	0.35	0.32
$EFMA_4$ [g m^{-2}]	EU-D	80.49b	0.00	188.17	1106.10
	EU-F	84.01b	88.27	110.39	1106.10
	LR-DH	91.63a	218.89	227.39	1106.10
$EFMA_6$ [g m^{-2}]	EU-D	76.33c	106.64	472.08	938.93
	EU-F	92.35b	420.23	398.00	938.93
	LR-DH	108.54a	1435.86	756.23	938.93
$EFMA_8$ [g m^{-2}]	EU-D	239.06c	1194.15	4605.50	3151.06
	EU-F	282.94b	3676.60	6056.53	3151.06
	LR-DH	304.22a	7243.76	7967.91	3151.06
$EPHT_4$ [cm]	EU-D	7.72c	0.47	0.51	1.77
	EU-F	8.91b	1.21	0.56	1.77
	LR-DH	9.75a	4.41	2.25	1.77
$EPHT_6$ [cm]	EU-D	12.64c	2.21	0.96	2.99
	EU-F	14.78b	3.84	1.22	2.99
	LR-DH	16.33a	12.81	2.79	2.99
$EPHT_8$ [cm]	EU-D	20.78c	4.83	5.10	5.11
	EU-F	24.60b	10.43	6.42	5.11
	LR-DH	26.01a	24.65	5.52	5.11
REGR [$\times 10^{-3}$ GDD^{-1}]	EU-D	17.10c	0.15	0.93	1.34
	EU-F	17.60b	0.69	0.79	1.34
	LR-DH	17.98a	1.12	0.16	1.34

† For traits description see table 2.

‡ Values followed by different letters are significant different

§ 1 = good, 9 = poor

¶ 1 = absent, 9 = pronounced

Annex 2 (continued).

Trait†	Population	Mean‡	σ^2_g	$\sigma^2_{g\times e}$	σ^2_e
FFLO [GDD]	EU-D	691.71a	1135.95	100.75	206.17
	EU-F	639.06c	1923.54	243.22	206.17
	LR-DH	654.15b	1624.02	336.10	206.17
MFLO [GDD]	EU-D	669.98a	1018.31	111.32	138.49
	EU-F	604.54b	1177.90	141.82	138.49
	LR-DH	610.19b	1460.99	231.16	138.49
ASIN [GDD]	EU-D	22.02c	149.30	48.74	140.12
	EU-F	34.38b	393.90	84.59	140.12
	LR-DH	43.00a	446.50	162.83	140.12
SPAD [SPAD unit]	EU-D	51.51a	9.57	0.71	10.35
	EU-F	49.53b	11.07	3.20	10.35
	LR-DH	49.67b	18.29	2.39	10.35
PLHT [cm]	EU-D	156.80a	235.75	29.26	42.23
	EU-F	148.96b	305.90	23.59	42.23
	LR-DH	143.33b	468.51	44.49	42.23
EAHT [cm]	EU-D	53.27a	91.60	21.37	30.39
	EU-F	51.19a	75.36	11.42	30.39
	LR-DH	47.11b	103.27	17.06	30.39
EASH [1-9]¶	EU-D	4.71a	0.88	0.20	0.41
	EU-F	4.10b	0.75	0.09	0.41
	LR-DH	4.05b	1.35	0.30	0.41
HUCO [1-9]§	EU-D	2.54a	2.38	0.65	0.40
	EU-F	2.29a	1.49	0.40	0.40
	LR-DH	1.60b	1.49	0.04	0.40
HUFL [1-9]¶	EU-D	1.44a	0.20	0.00	0.47
	EU-F	1.62a	0.41	0.02	0.47
	LR-DH	3.03b	2.65	0.73	0.47
LODG [%]	EU-D	3.56a	0.00	0.00	37.59
	EU-F	5.60a	0.83	0.00	37.59
	LR-DH	15.82b	125.84	163.76	37.59
SMUT [%]	EU-D	2.51a	2.06	0.00	19.76
	EU-F	2.68a	10.18	0.00	19.76
	LR-DH	7.74b	73.86	34.53	19.76
BAST [%]	EU-D	1.69a	0.00	0.00	16.92
	EU-F	2.83a	2.78	0.12	16.92
	LR-DH	6.09b	13.23	22.18	16.92

† For traits description see table 2.

‡ Values followed by different letters are significant different

§ 1 = good, 9 = poor

¶ 1 = absent, 9 = pronounced

Annex 2 (continued).

Trait†	Population	Mean‡	σ^2_g	$\sigma^2_{g\times e}$	σ^2_e
IFUS [%]	EU-D	25.58a	347.23	165.64	279.33
	EU-F	21.15a	223.28	45.02	279.33
	LR-DH	33.71b	396.16	156.87	279.33
SFUS [%]	EU-D	2.46a	5.87	4.19	6.99
	EU-F	1.70a	1.31	0.00	6.99
	LR-DH	4.62b	15.93	17.83	6.99
EALE [cm]	EU-D	126.38a	135.25	21.46	103.31
	EU-F	133.24b	239.49	28.12	103.31
	LR-DH	123.20a	422.97	79.89	103.31
EADI [cm]	EU-D	36.48b	3.48	0.86	3.10
	EU-F	33.95a	4.34	0.33	3.10
	LR-DH	33.46a	10.45	1.45	3.10
ROWS [#]	EU-D	13.39a	1.70	0.13	0.50
	EU-F	12.93a	1.30	0.14	0.50
	LR-DH	10.70b	2.61	0.10	0.50
KERO [#]	EU-D	20.95a	5.02	2.90	7.10
	EU-F	21.61a	8.01	1.36	7.10
	LR-DH	17.13b	13.53	3.46	7.10
THKW [g]	EU-D	212.76a	765.76	173.49	355.73
	EU-F	212.20a	756.30	76.80	355.73
	LR-DH	227.21b	1189.48	370.36	355.73
EDMC [%]	EU-D	58.73a	35.79	4.76	7.12
	EU-F	59.31a	23.61	2.75	7.12
	LR-DH	55.33b	51.81	7.94	7.12
GRYD [g]	EU-D	56.65b	48.77	78.85	87.22
	EU-F	55.82b	78.35	33.89	87.22
	LR-DH	42.21a	93.69	24.34	87.22

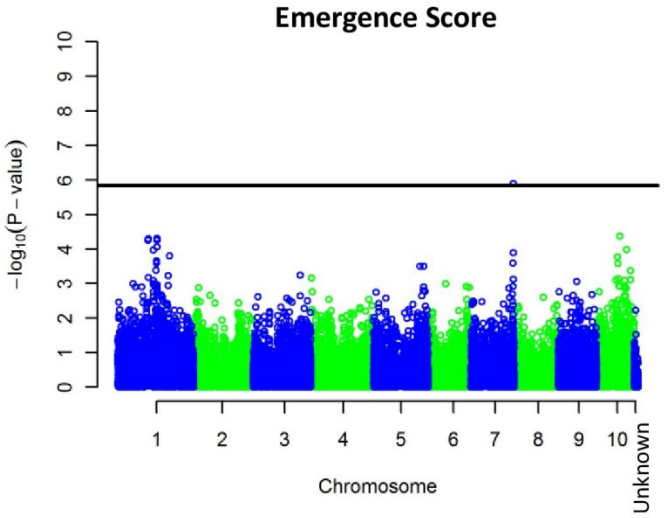
† For traits description see table 2.

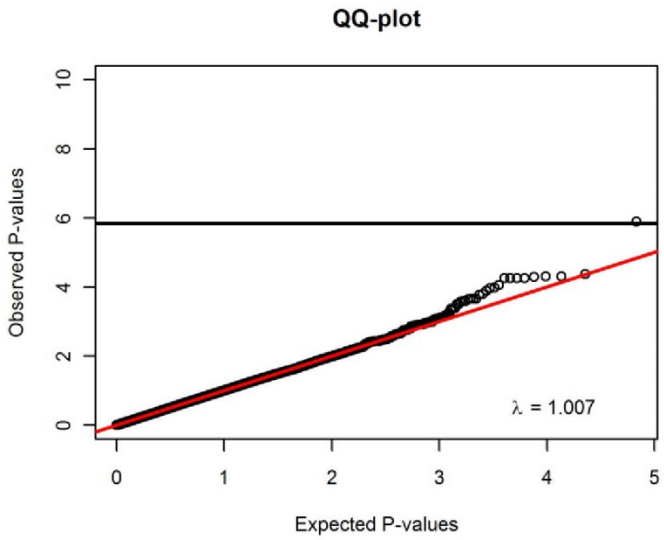
‡ Values followed by different letters are significant different

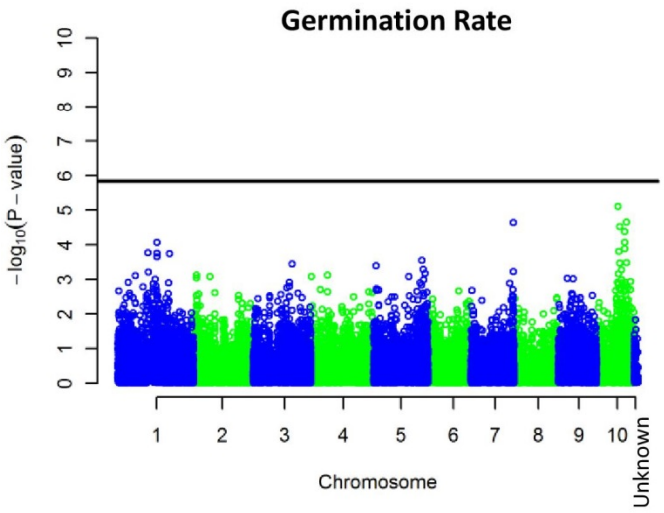
§ 1 = good, 9 = poor

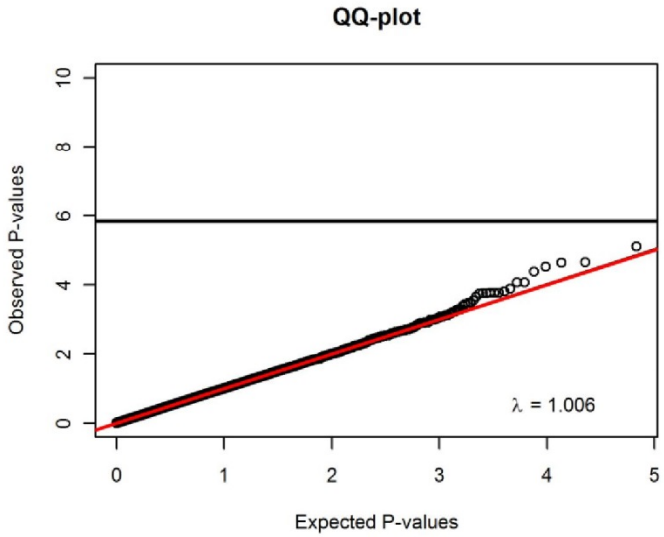
¶ 1 = absent, 9 = pronounced

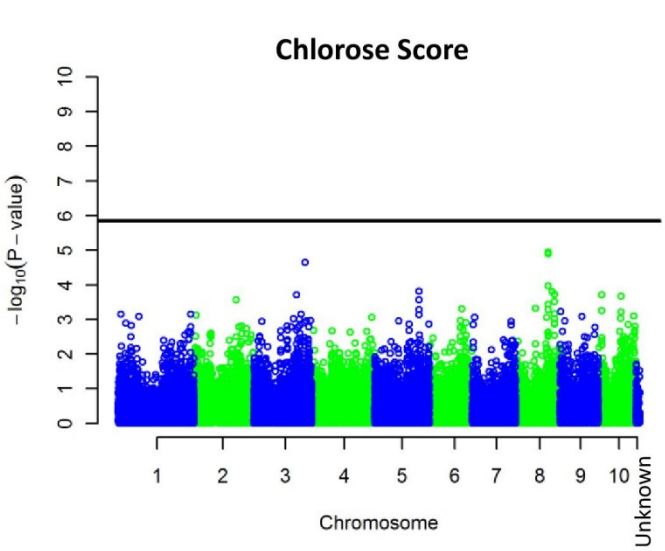
Annex 3. Genome wide association scans for single nucleotide polymorphism (SNP) × trait associations detected in Set 3 with a model correcting for population structure using the kinship matrix and five first principal coordinates from the principal coordinate analysis performed on the marker data. Left hand: The $-\log_{10}(P)$ values from the genome wide scan are plotted against the SNP position on the physical map of each chromosome, for each trait × treatment combination. Right hand: QQ-plot of expected against observed *P* values for SNP × trait associations, and corresponding inflation factor λ. The horizontal line shows the significance threshold ($\alpha = 0.05$) after Bonferroni-correction for multiple comparison.

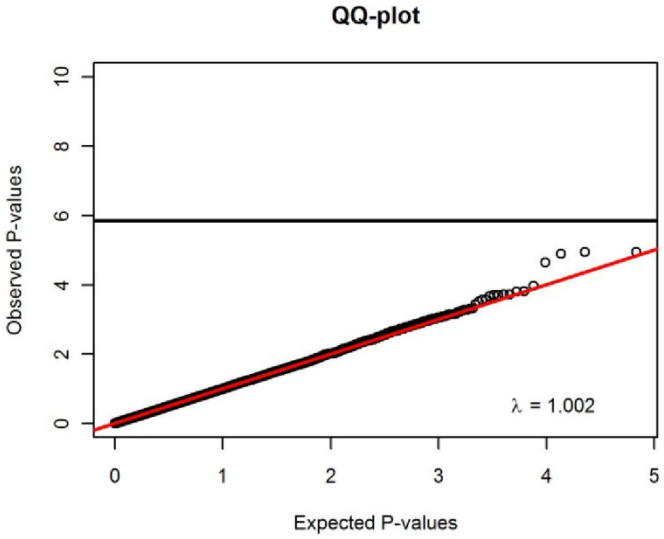

Emergence Score
−log10(P − value)
Chromosome
Unknown

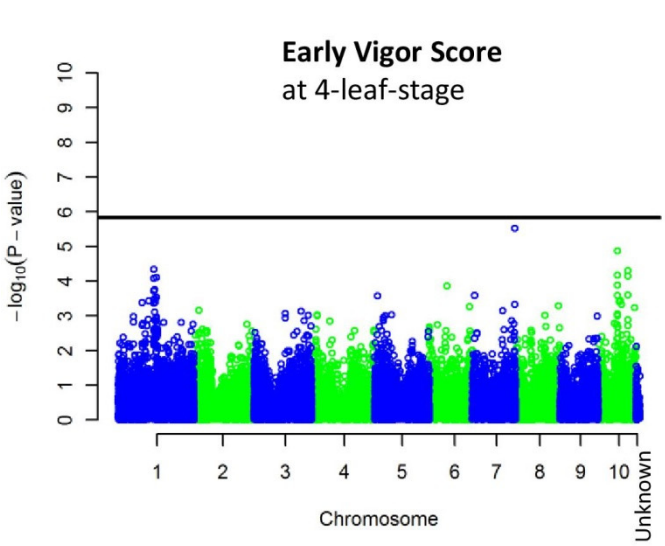

QQ-plot
Observed P-values
Expected P-values
λ = 1.007

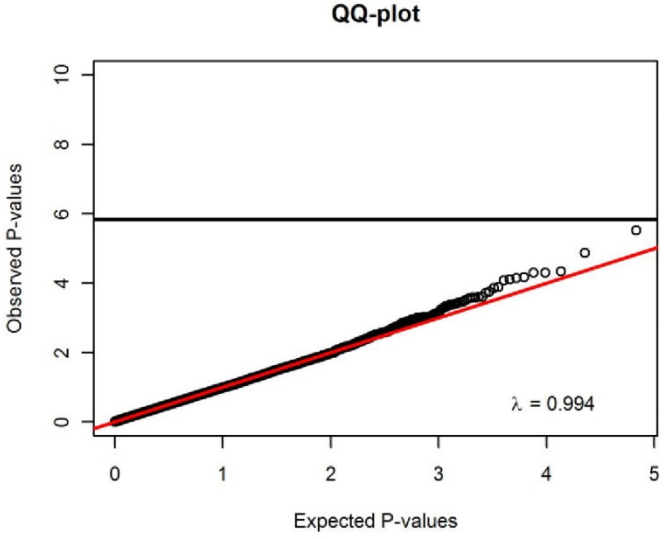

Germination Rate
−log10(P − value)
Chromosome
Unknown

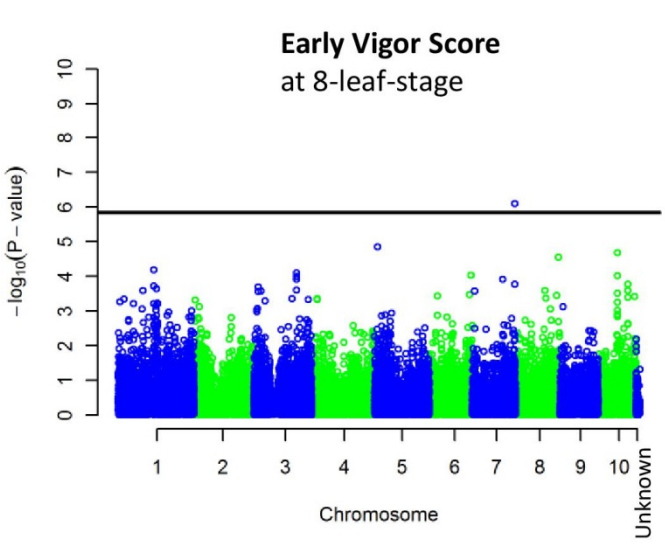

QQ-plot
Observed P-values
Expected P-values
λ = 1.006

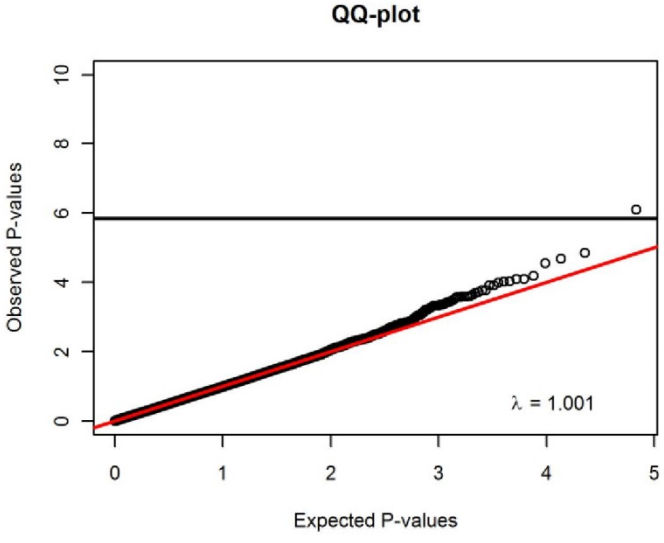

Chlorose Score
$-\log_{10}(P-value)$
Chromosome
Unknown

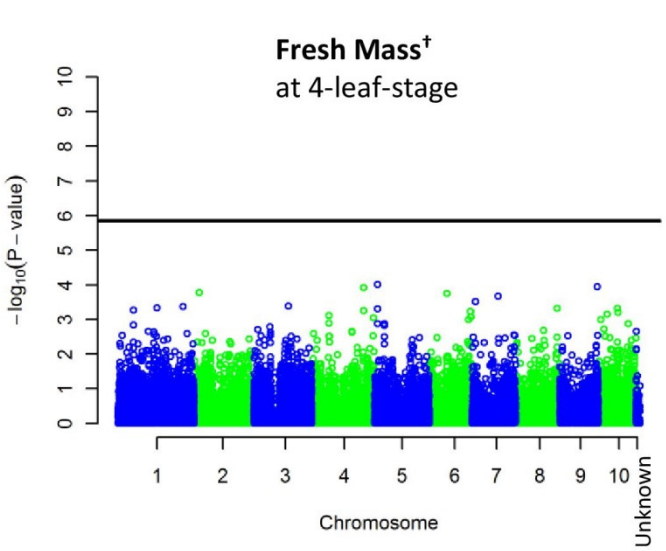

QQ-plot
Observed P-values
Expected P-values
λ = 1.002

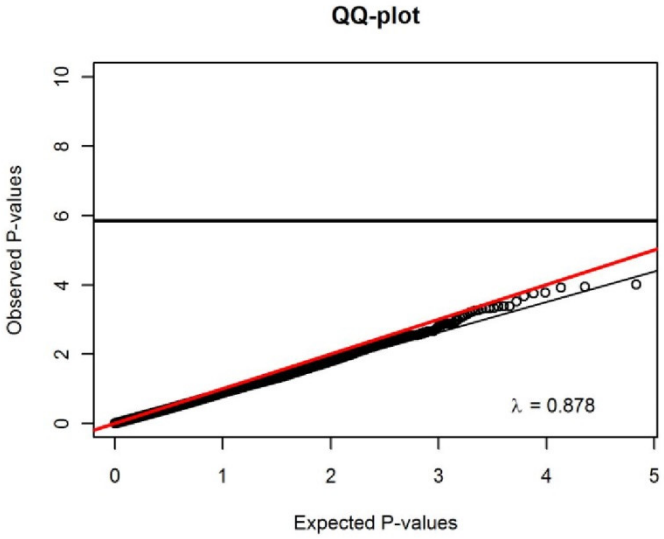

Early Vigor Score
at 4-leaf-stage
$-\log_{10}(P-value)$
Chromosome
Unknown

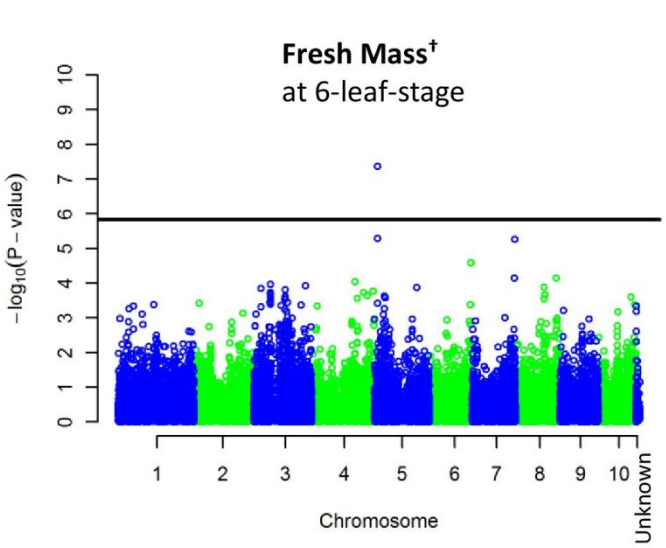

QQ-plot
Observed P-values
Expected P-values
λ = 0.994

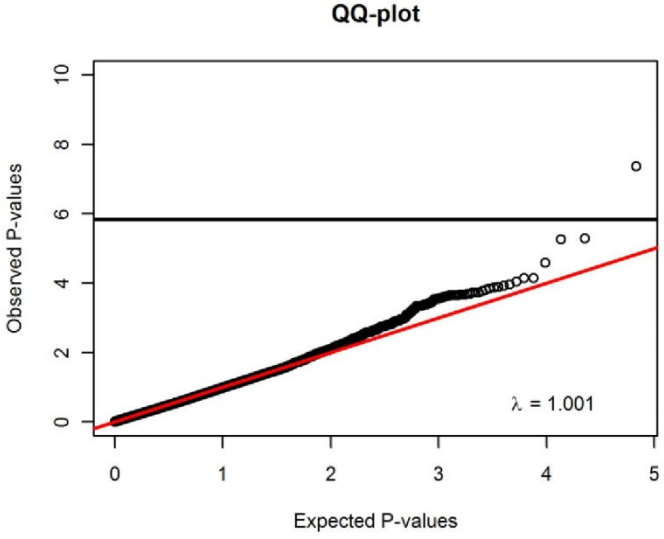

Early Vigor Score
at 8-leaf-stage
$-\log_{10}(P-value)$
Chromosome
Unknown

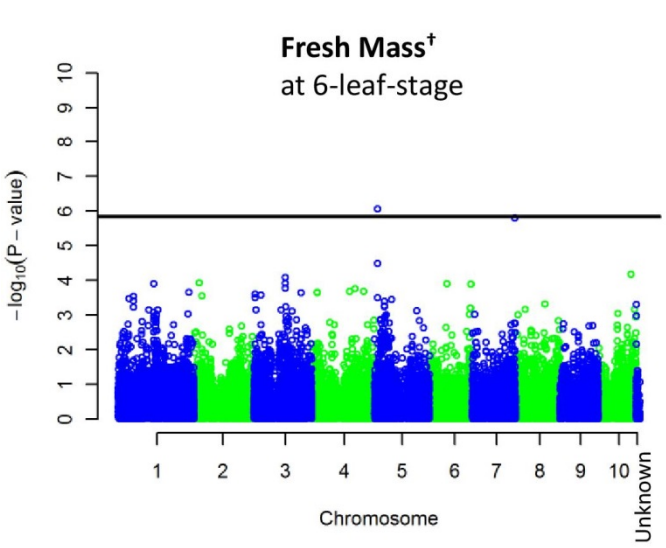

QQ-plot
Observed P-values
Expected P-values
λ = 1.001

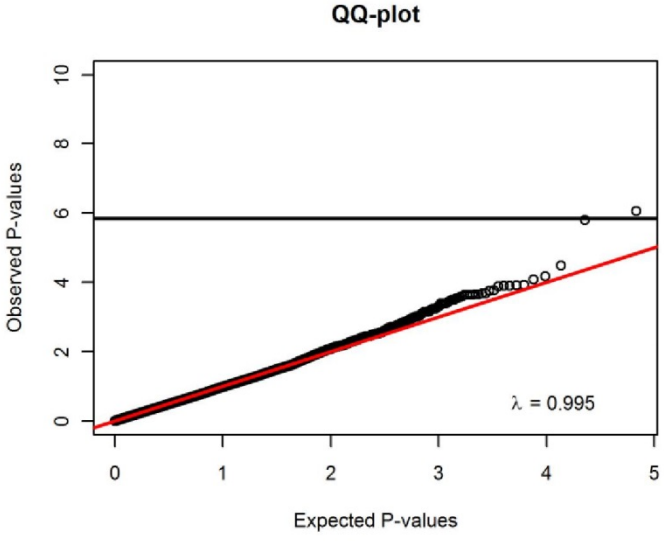

Fresh Mass†
at 4-leaf-stage
−log10(P − value)
Chromosome
Unknown

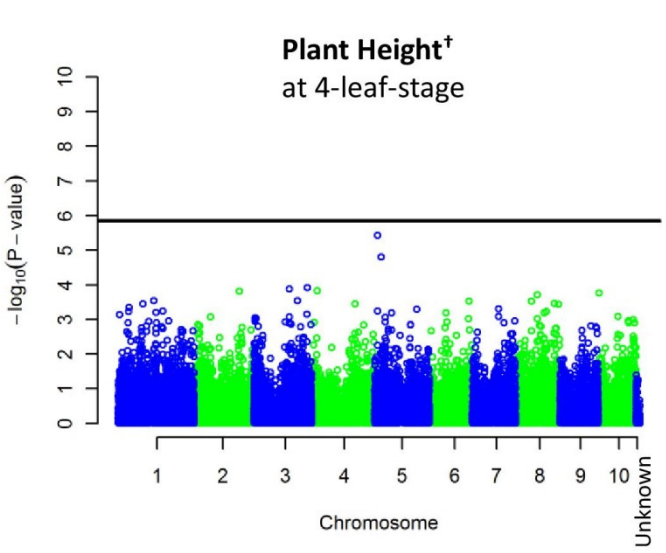

QQ-plot
Observed P-values
Expected P-values
λ = 0.878

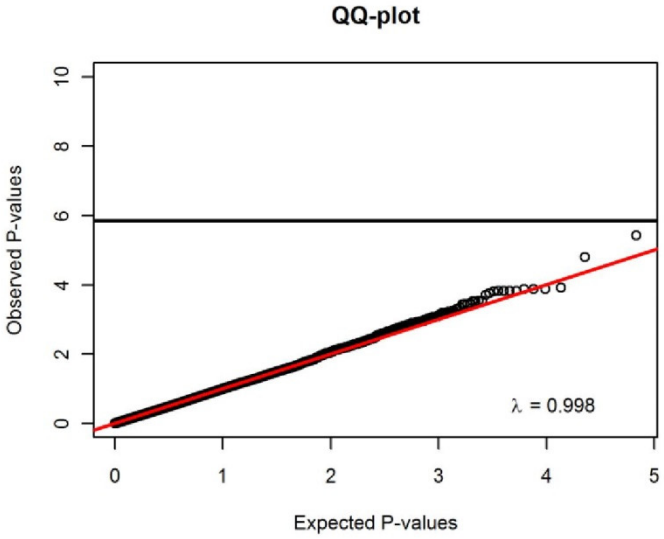

Fresh Mass†
at 6-leaf-stage
−log10(P − value)
Chromosome
Unknown

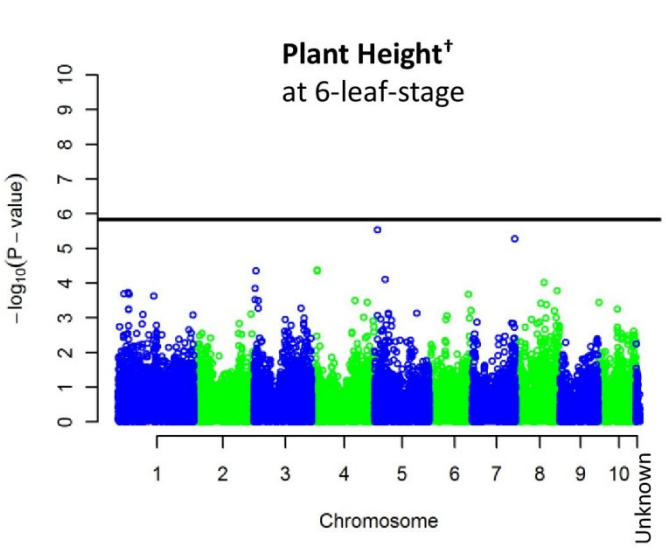

QQ-plot
Observed P-values
Expected P-values
λ = 1.001

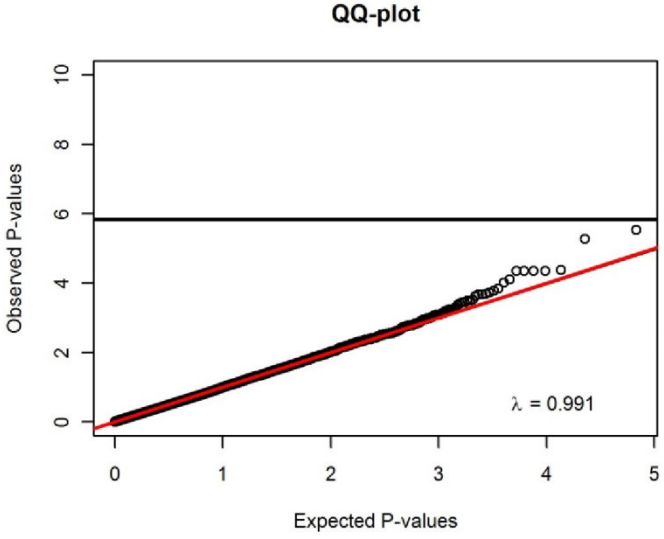

Fresh Mass†
at 6-leaf-stage
−log10(P − value)
Chromosome
Unknown

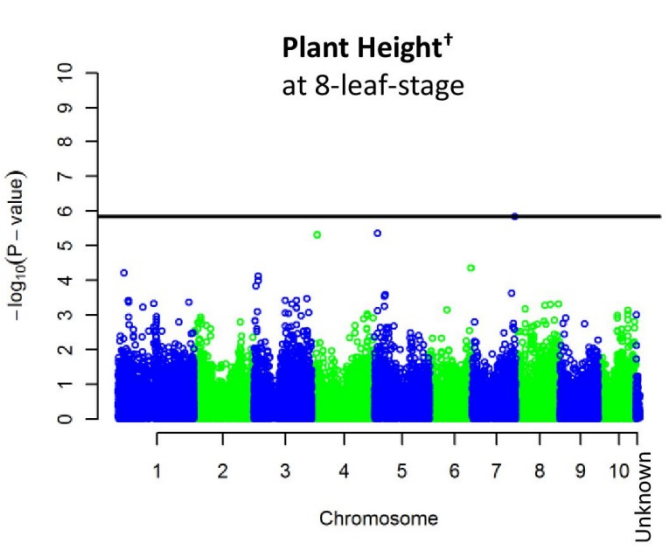

QQ-plot
Observed P-values
Expected P-values
λ = 0.995

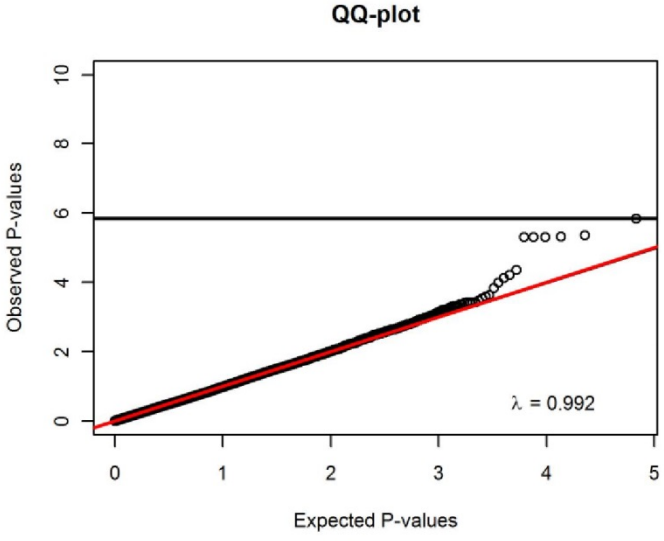

Plant Height†
at 4-leaf-stage
$-\log_{10}(P-value)$
Chromosome
Unknown

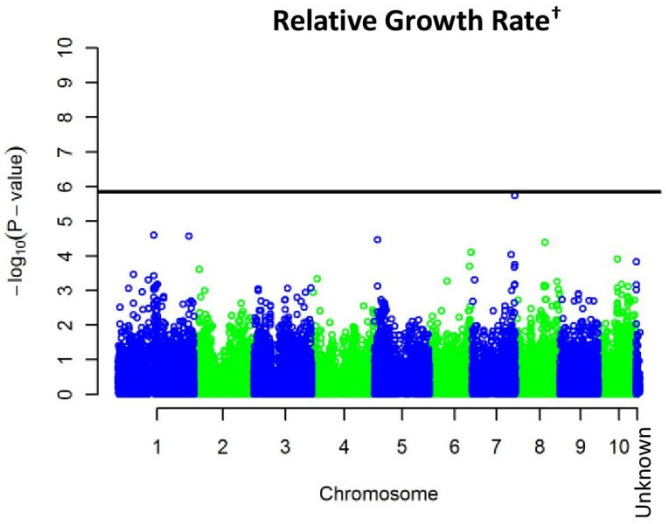

QQ-plot
Observed P-values
Expected P-values
λ = 0.998

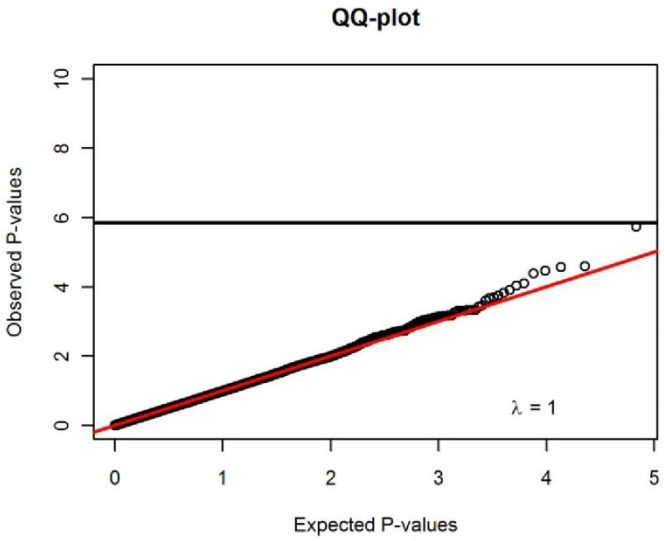

Plant Height†
at 6-leaf-stage
$-\log_{10}(P-value)$
Chromosome
Unknown

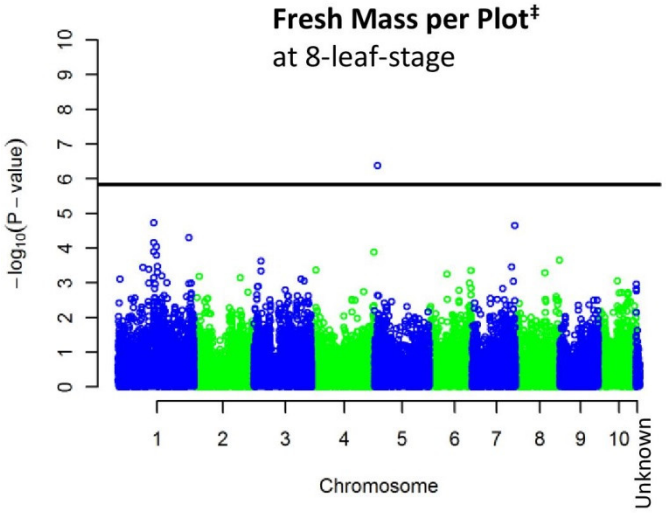

QQ-plot
Observed P-values
Expected P-values
λ = 0.991

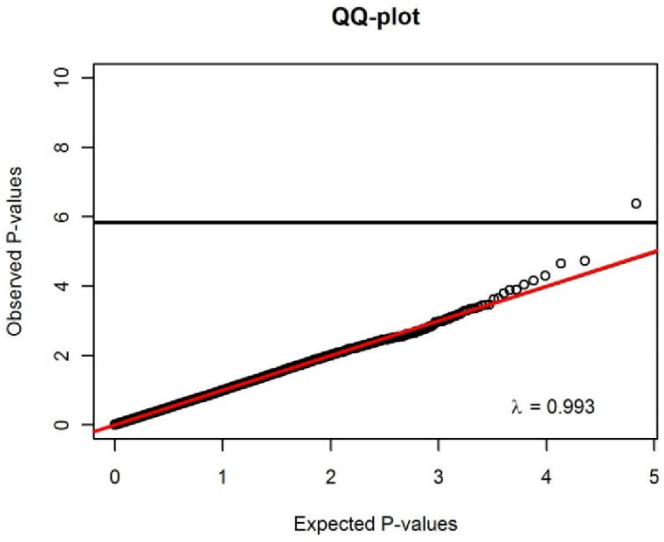

Plant Height†
at 8-leaf-stage
$-\log_{10}(P-value)$
Chromosome
Unknown

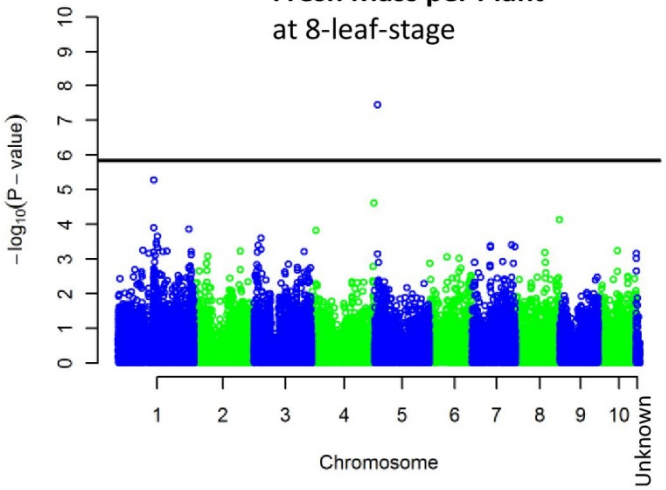

QQ-plot
Observed P-values
Expected P-values
λ = 0.992

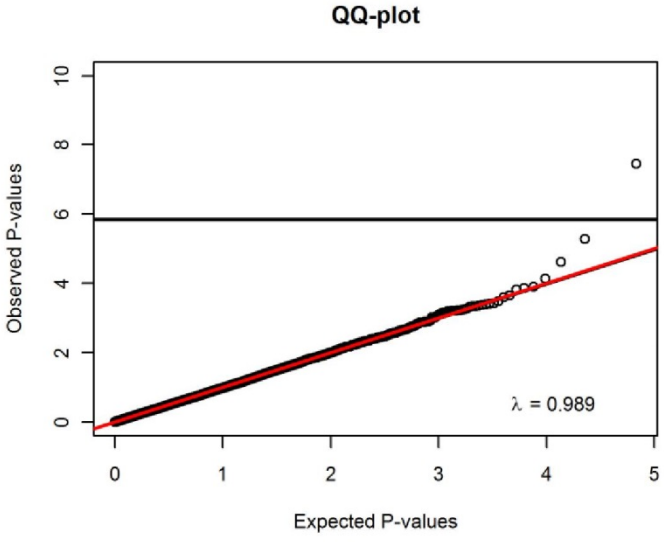

Relative Growth Rate†
−log10(P − value)
Chromosome
Unknown

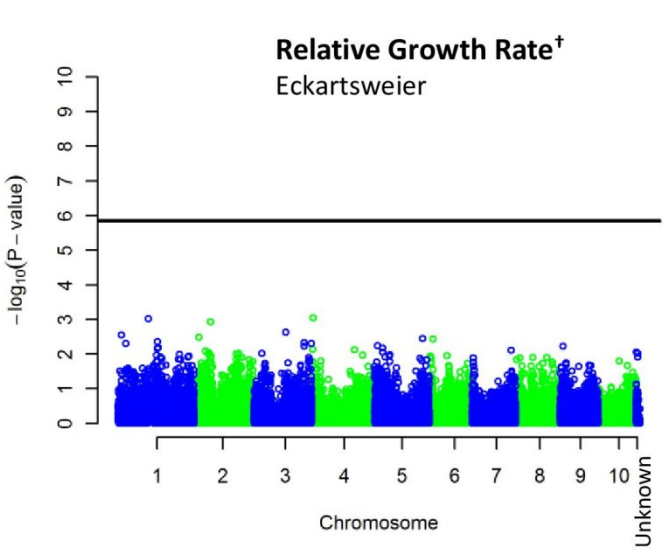

QQ-plot
Observed P-values
Expected P-values
λ = 1

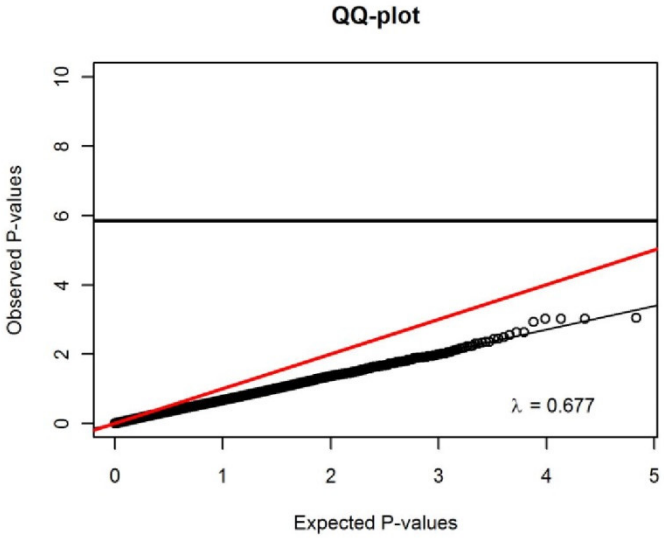

Fresh Mass per Plot‡
at 8-leaf-stage
−log10(P − value)
Chromosome
Unknown

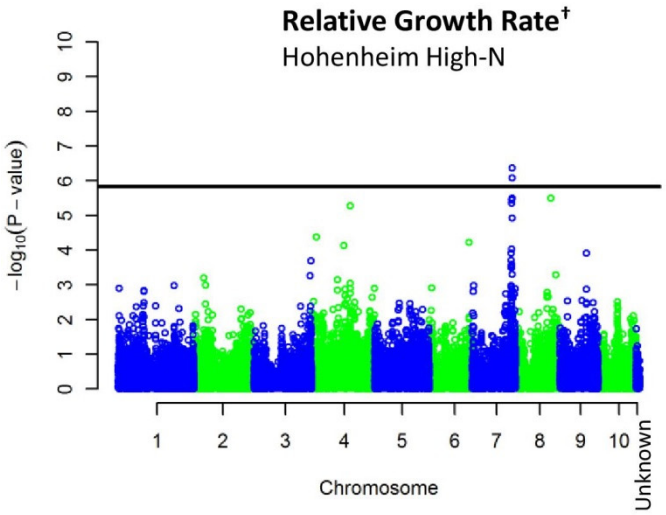

QQ-plot
Observed P-values
Expected P-values
λ = 0.993

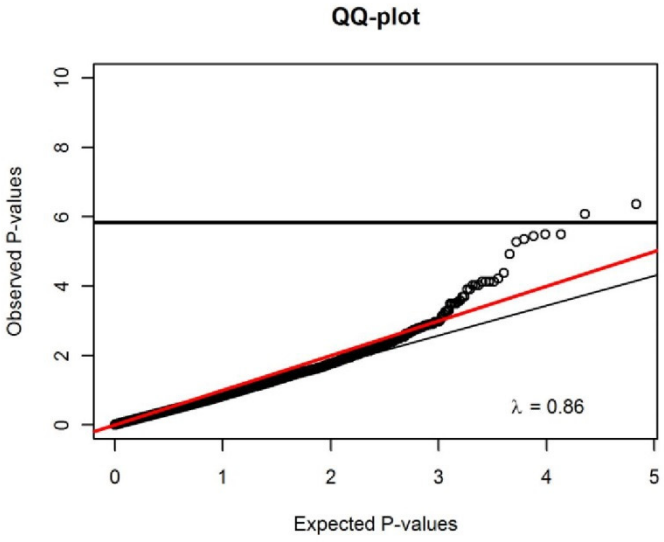

Fresh Mass per Plant‡
at 8-leaf-stage
−log10(P − value)
Chromosome
Unknown

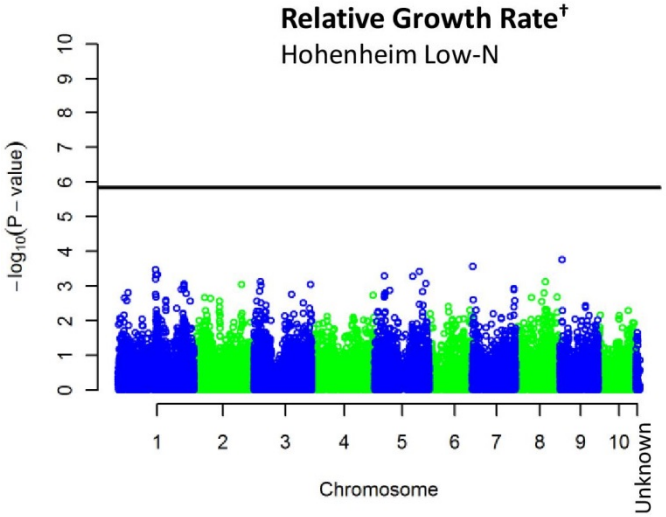

QQ-plot
Observed P-values
Expected P-values
λ = 0.989

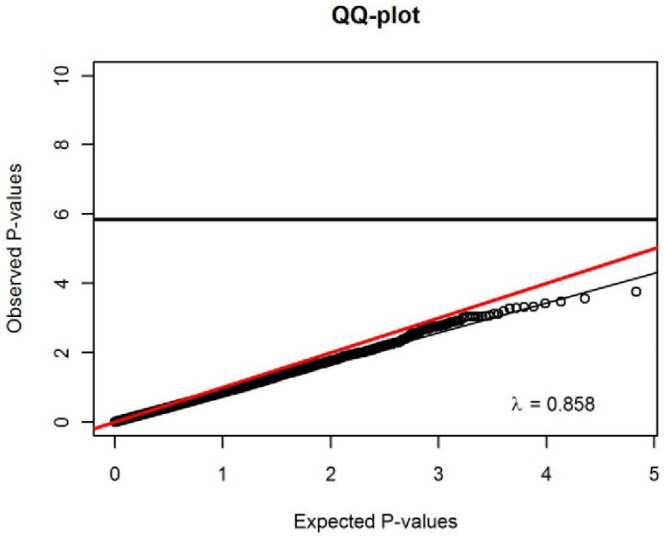

Relative Growth Rate†
Eckartsweier
−log10(P−value)
Chromosome
Unknown

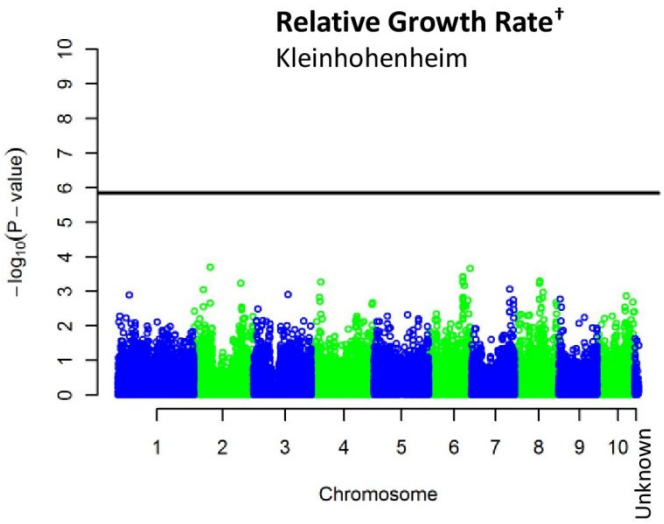

QQ-plot
Observed P-values
Expected P-values
λ = 0.677

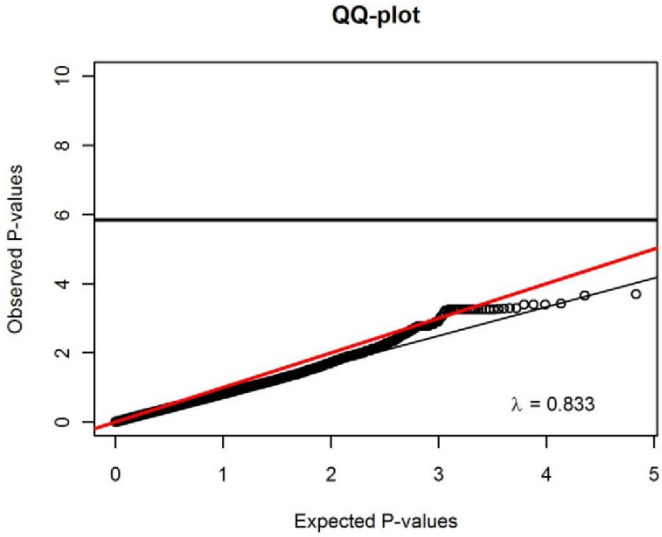

Relative Growth Rate†
Hohenheim High-N
−log10(P−value)
Chromosome
Unknown

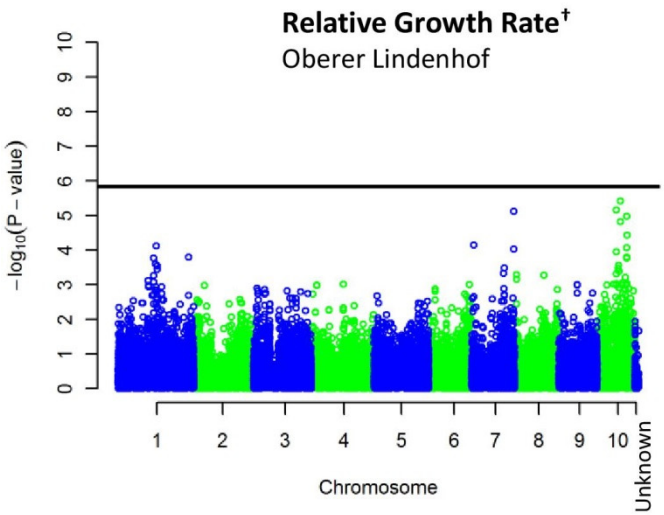

QQ-plot
Observed P-values
Expected P-values
λ = 0.86

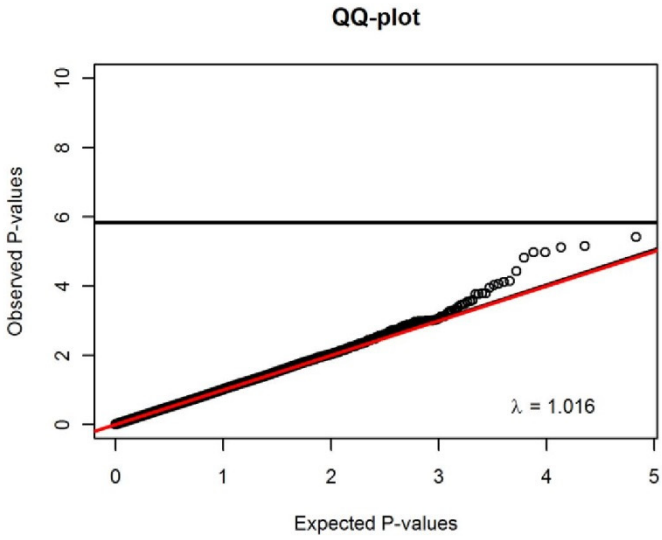

Relative Growth Rate†
Hohenheim Low-N
−log10(P−value)
Chromosome
Unknown

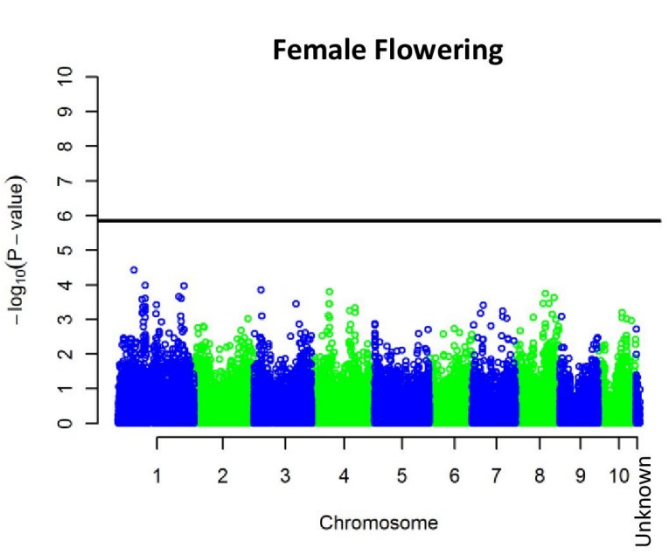

QQ-plot
Observed P-values
Expected P-values
λ = 0.858

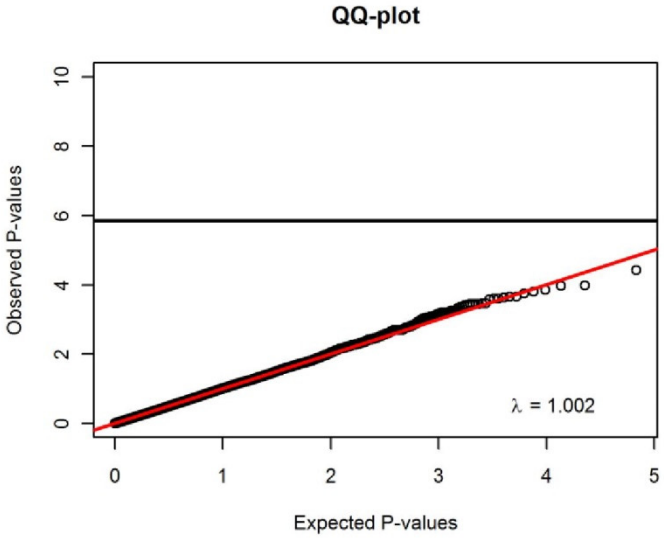

Relative Growth Rate†
Kleinhohenheim
−log10(P − value)
Chromosome
Unknown

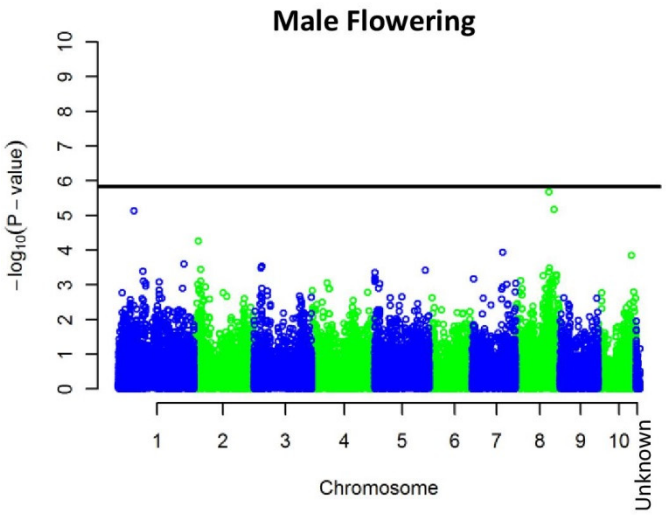

QQ-plot
Observed P-values
Expected P-values
λ = 0.833

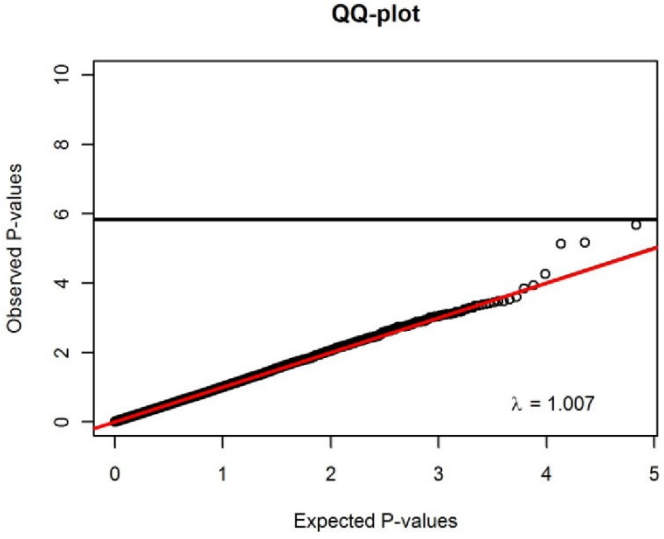

Relative Growth Rate†
Oberer Lindenhof
−log10(P − value)
Chromosome
Unknown

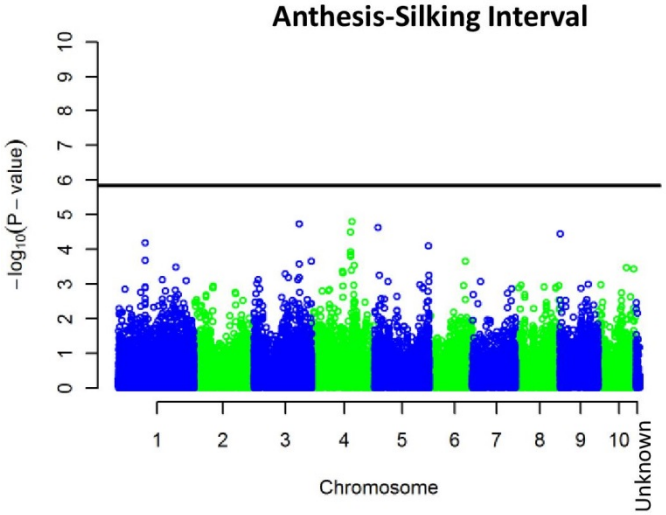

QQ-plot
Observed P-values
Expected P-values
λ = 1.016

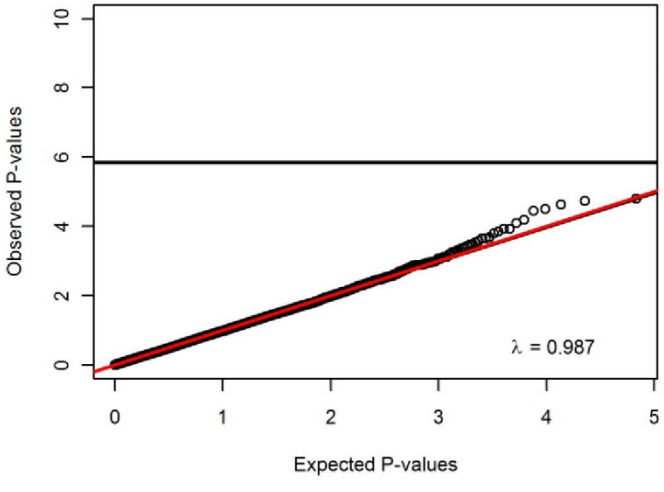

Female Flowering
$-\log_{10}(P-value)$
Chromosome
Unknown

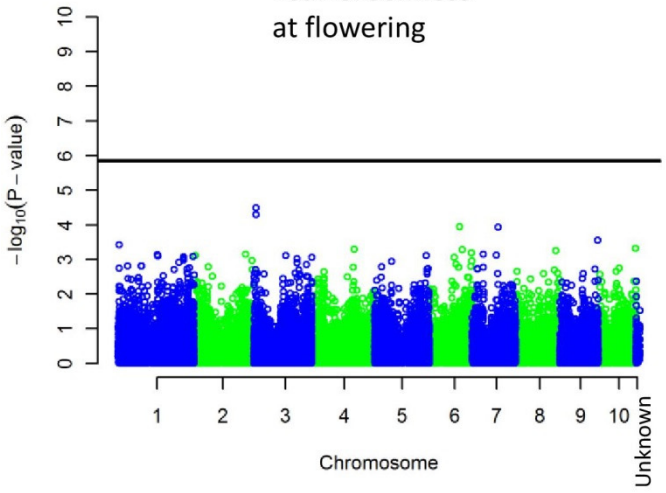

QQ-plot
Observed P-values
Expected P-values
λ = 1.002

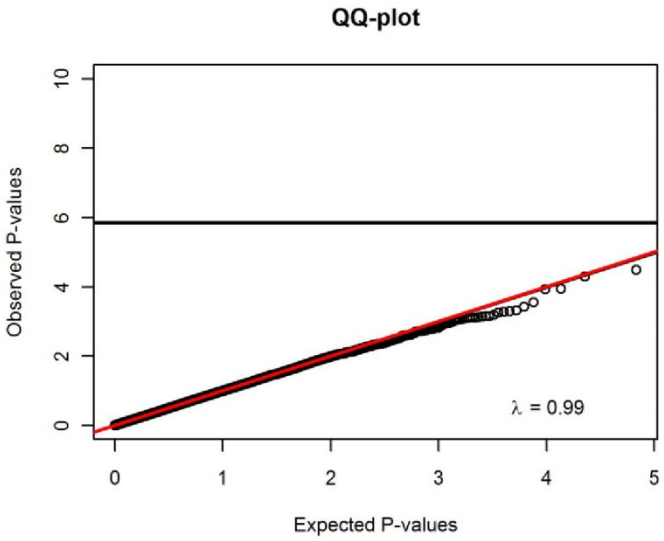

Male Flowering
$-\log_{10}(P-value)$
Chromosome
Unknown

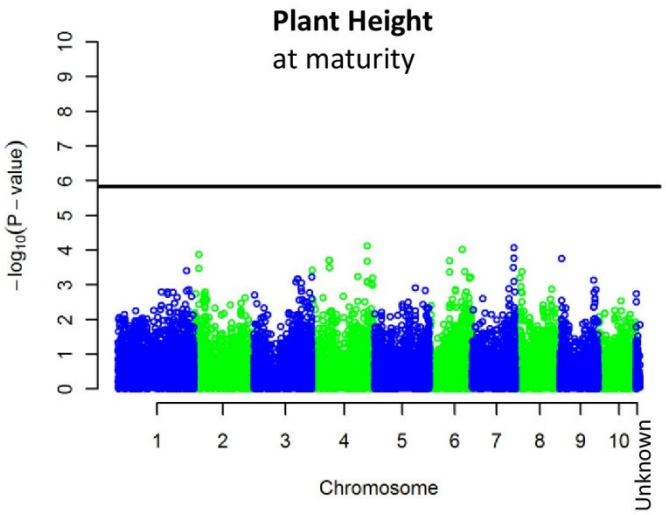

QQ-plot
Observed P-values
Expected P-values
λ = 1.007

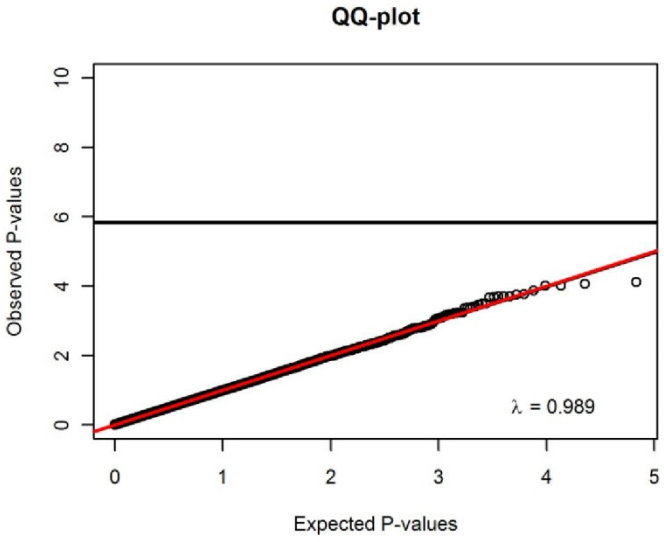

Anthesis-Silking Interval
$-\log_{10}(P-value)$
Chromosome
Unknown

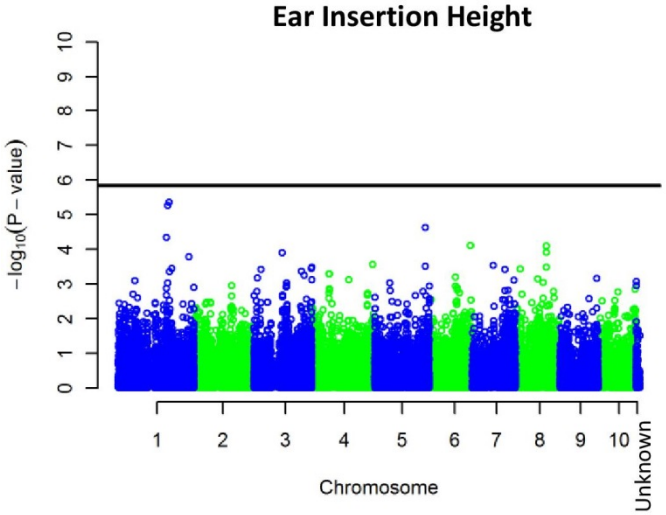

QQ-plot
Observed P-values
Expected P-values
λ = 0.987

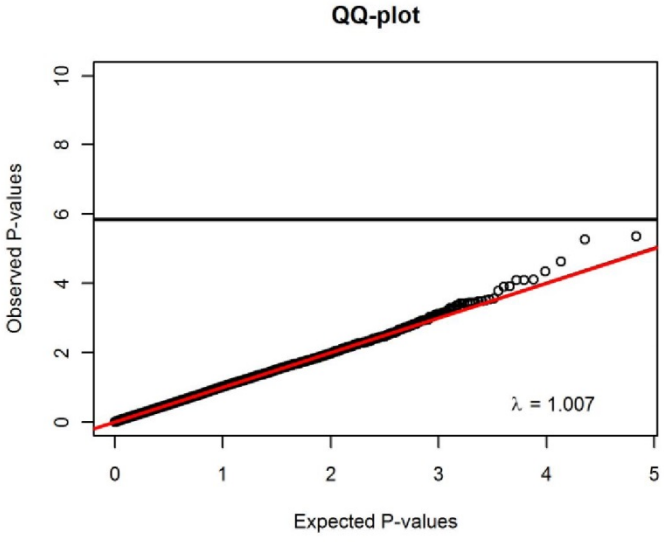

Leaf Greenness
at flowering
$-\log_{10}(P-value)$
Chromosome
Unknown

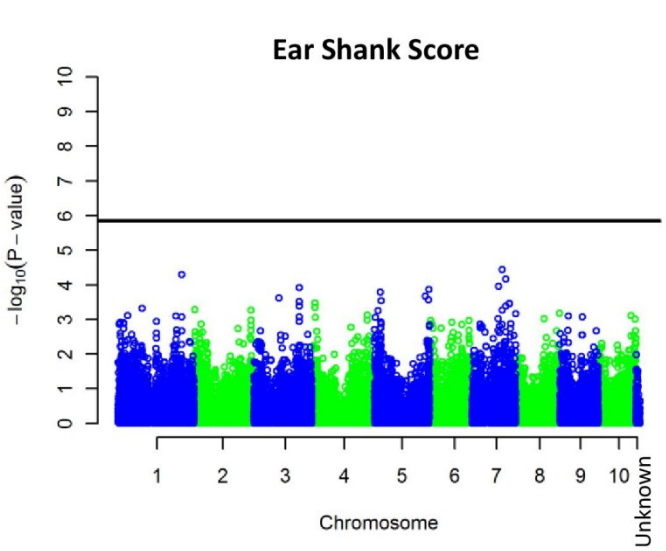

QQ-plot
Observed P-values
Expected P-values
λ = 0.99

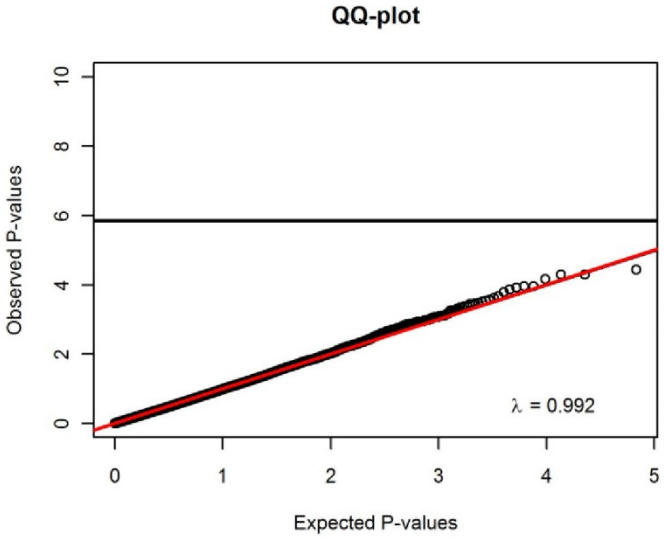

Plant Height
at maturity
$-\log_{10}(P-value)$
Chromosome
Unknown

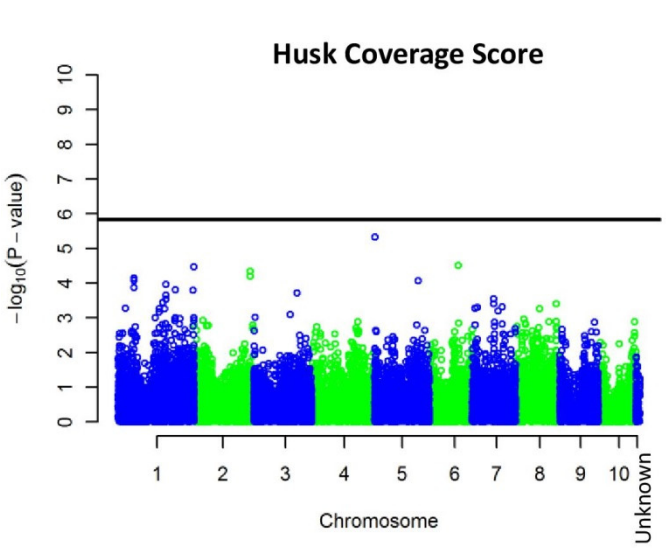

QQ-plot
Observed P-values
Expected P-values
λ = 0.989

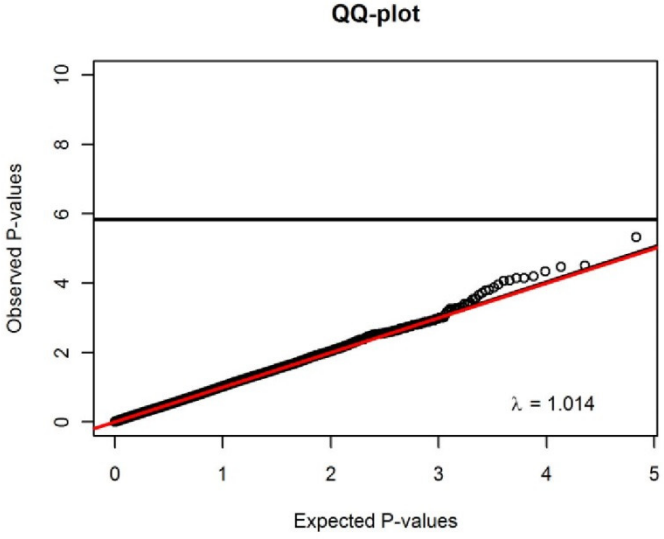

Ear Insertion Height
$-\log_{10}(P-value)$
Chromosome
Unknown

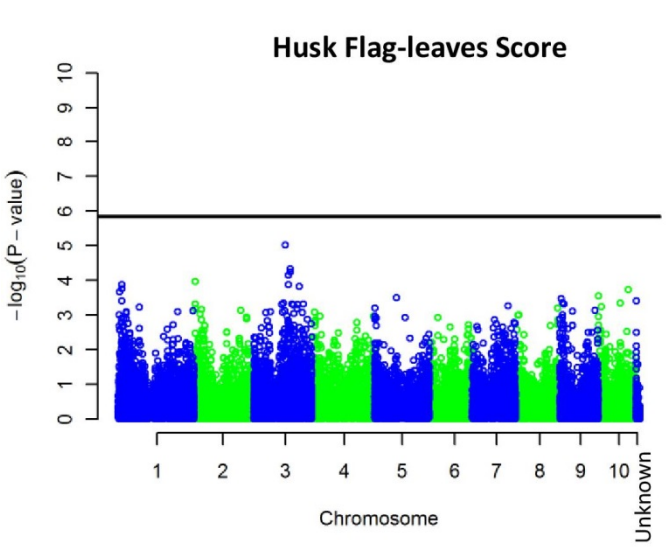

QQ-plot
Observed P-values
Expected P-values
λ = 1.007

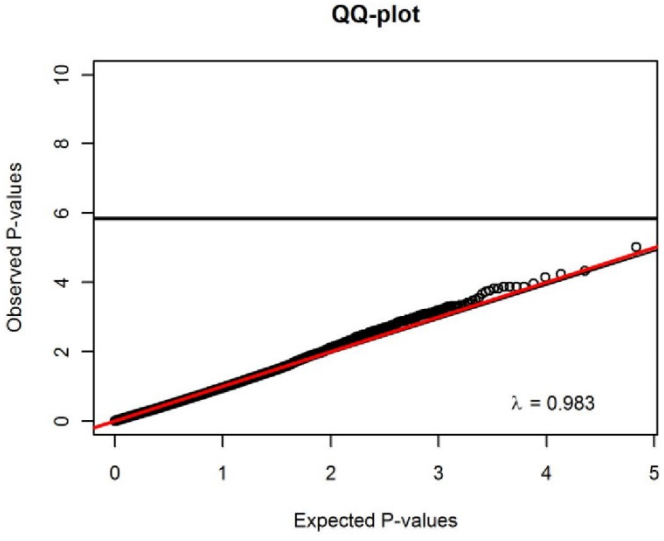

Ear Shank Score
−log10(P − value)
Chromosome
Unknown

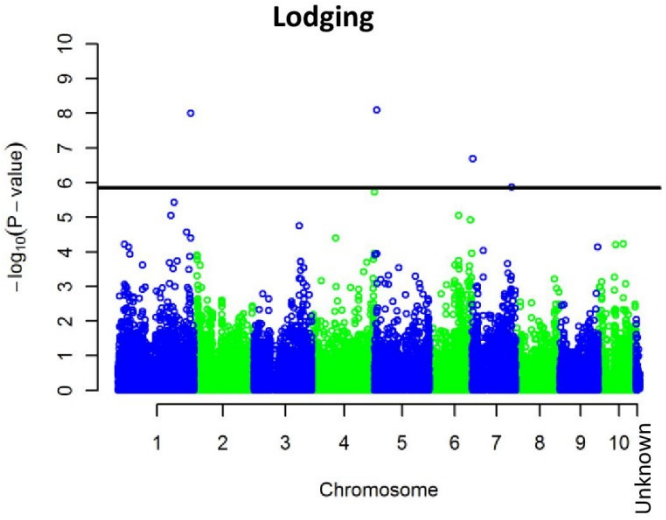

QQ-plot
Observed P-values
Expected P-values
λ = 0.992

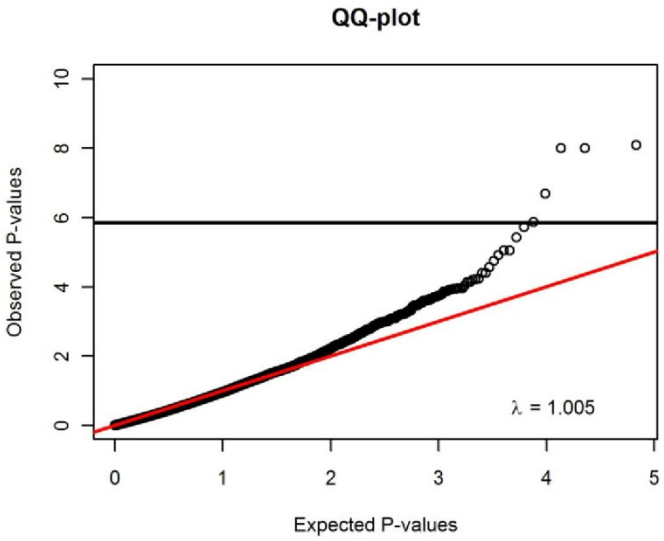

Husk Coverage Score
−log10(P − value)
Chromosome
Unknown

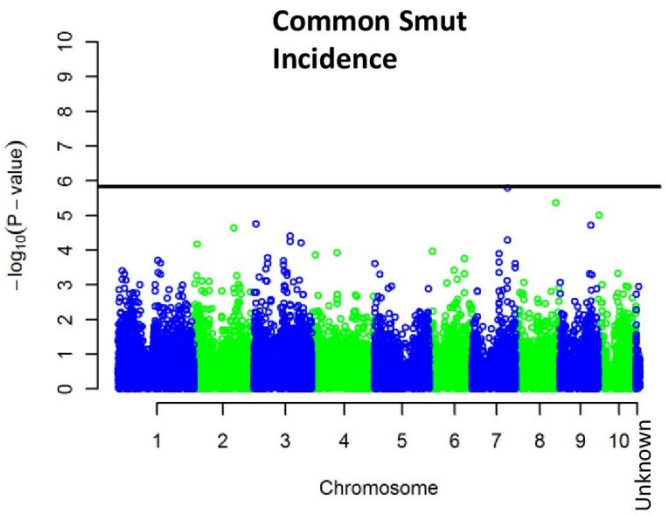

QQ-plot
Observed P-values
Expected P-values
λ = 1.014

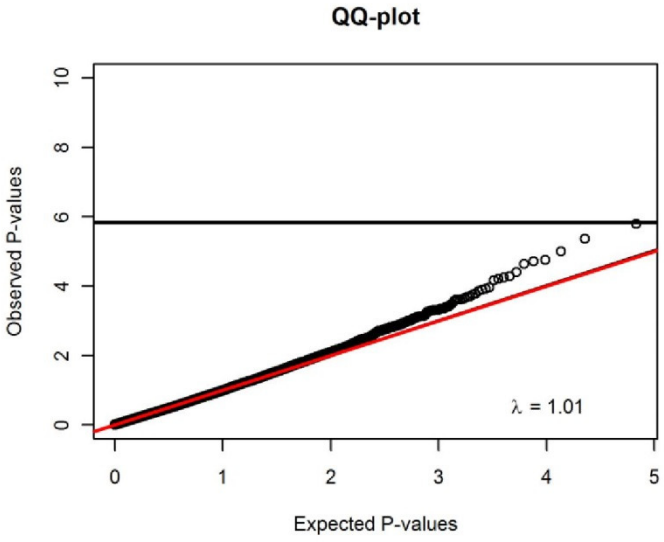

Husk Flag-leaves Score
−log10(P − value)
Chromosome
Unknown

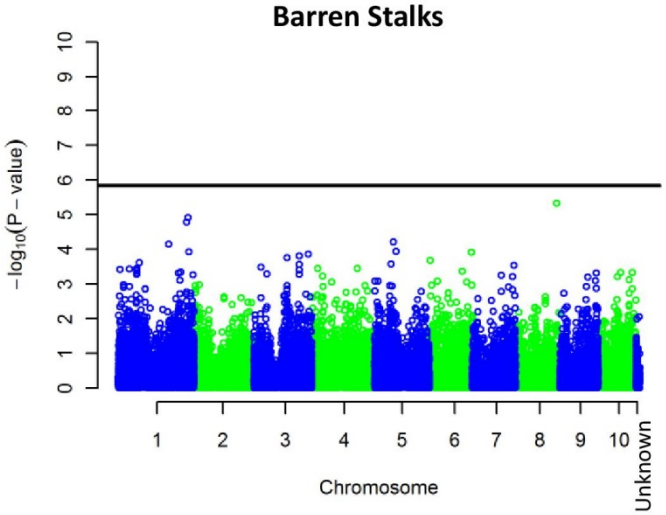

QQ-plot
Observed P-values
Expected P-values
λ = 0.983

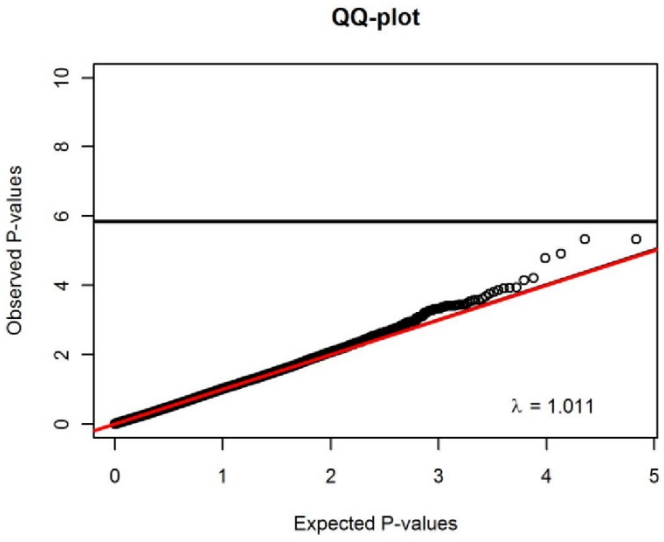

Lodging
$-\log_{10}(P-value)$
Chromosome
Unknown

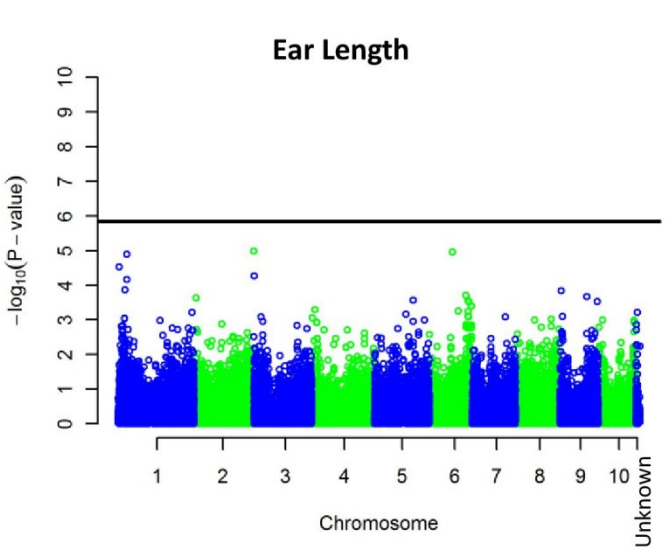

QQ-plot
Observed P-values
Expected P-values
λ = 1.005

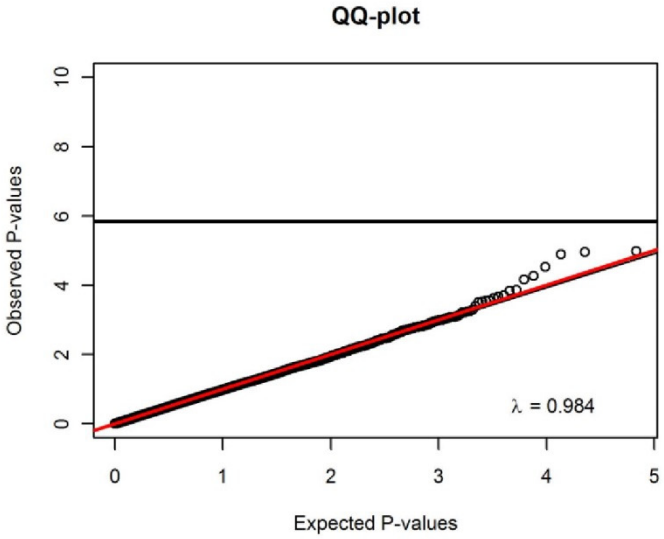

Common Smut Incidence
$-\log_{10}(P-value)$
Chromosome
Unknown

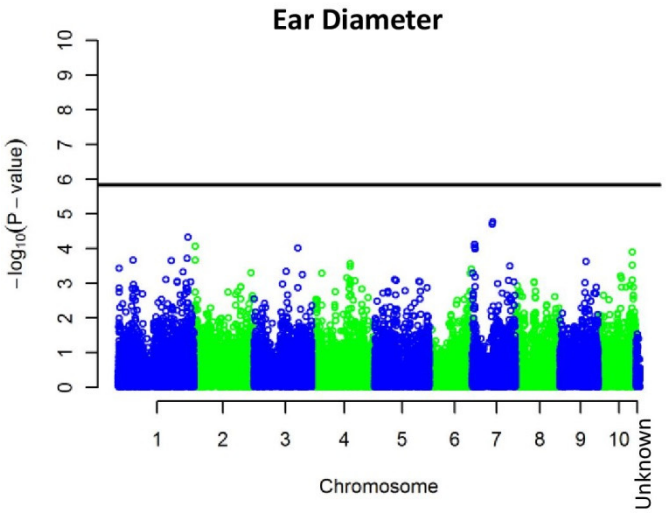

QQ-plot
Observed P-values
Expected P-values
λ = 1.01

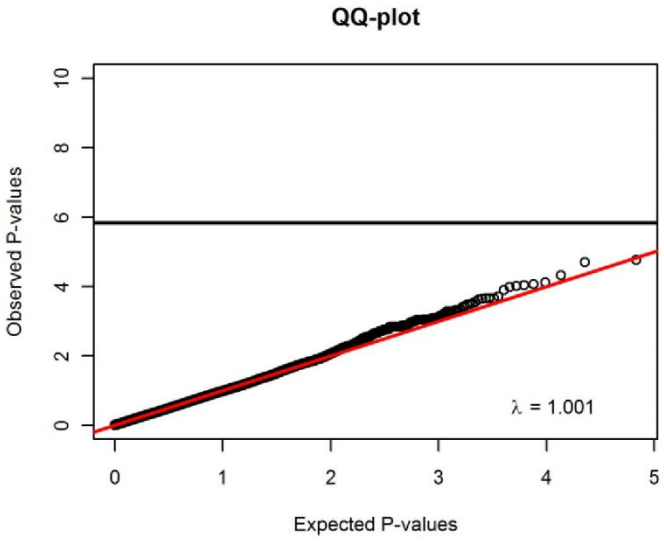

Barren Stalks
$-\log_{10}(P-value)$
Chromosome
Unknown

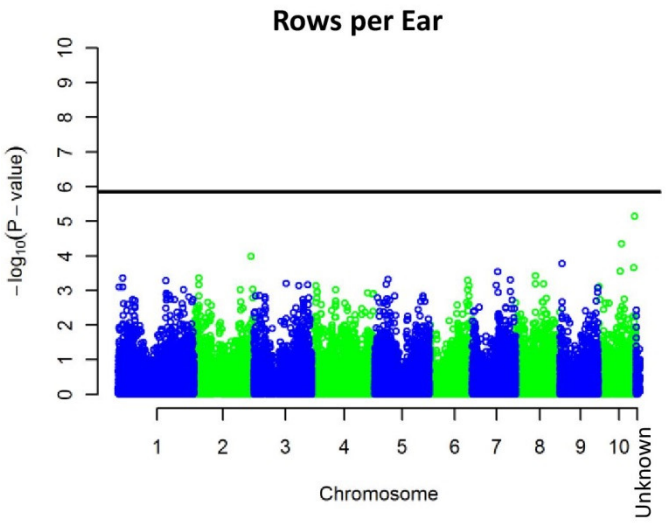

QQ-plot
Observed P-values
Expected P-values
λ = 1.011

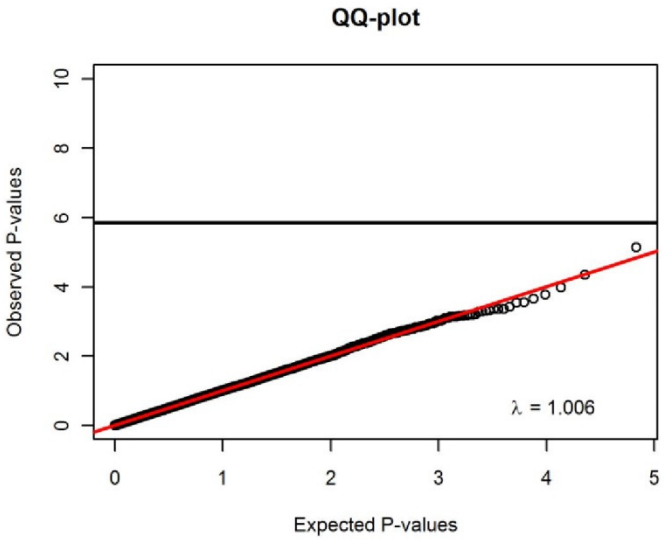

Ear Length
−log10(P − value)
Chromosome
Unknown

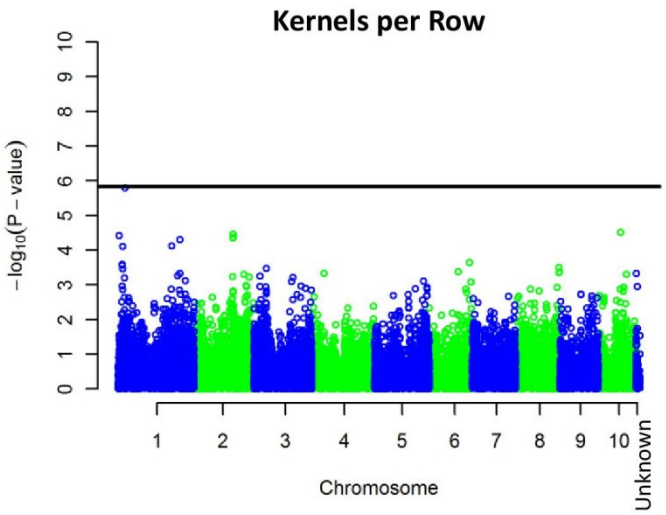

QQ-plot
Observed P-values
Expected P-values
λ = 0.984

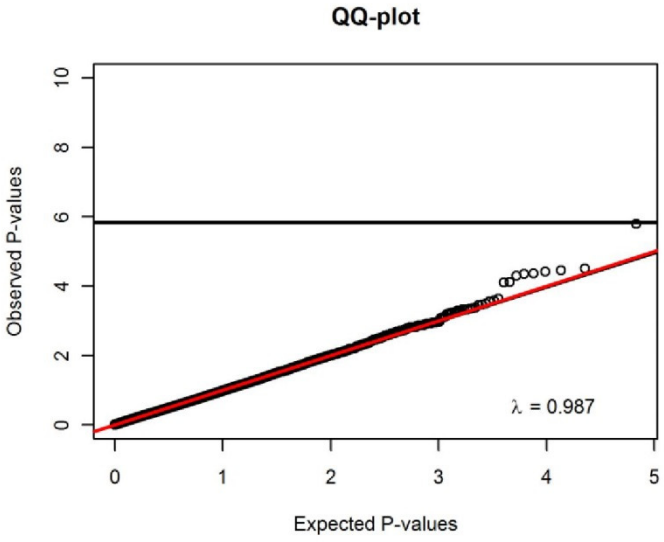

Ear Diameter
−log10(P − value)
Chromosome
Unknown

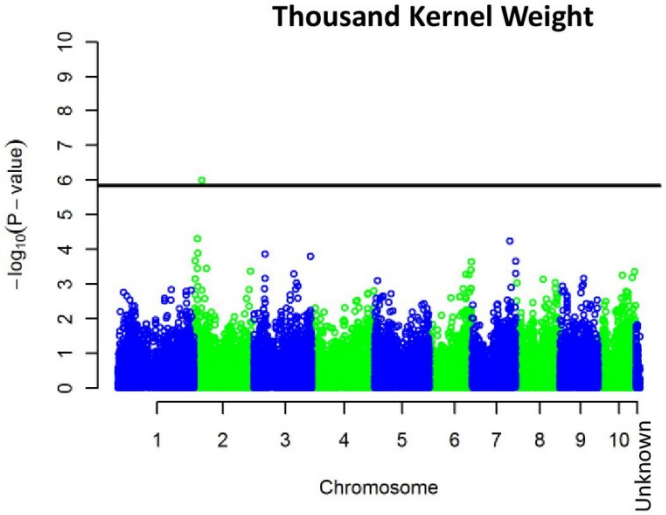

QQ-plot
Observed P-values
Expected P-values
λ = 1.001

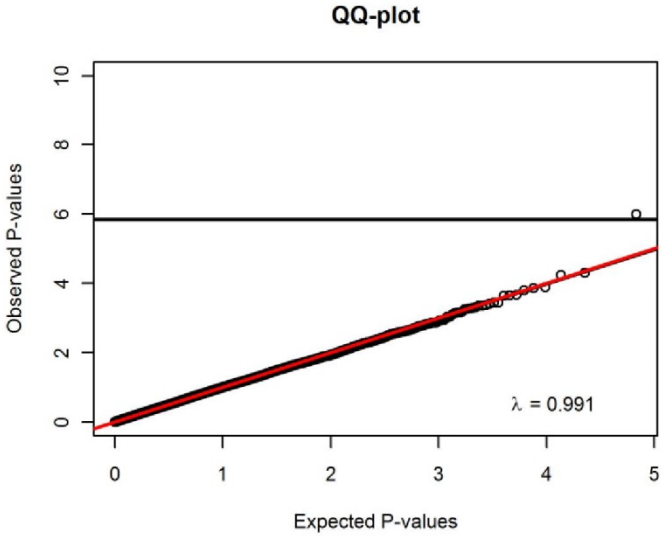

Rows per Ear
$-\log_{10}(P-value)$
Chromosome
Unknown

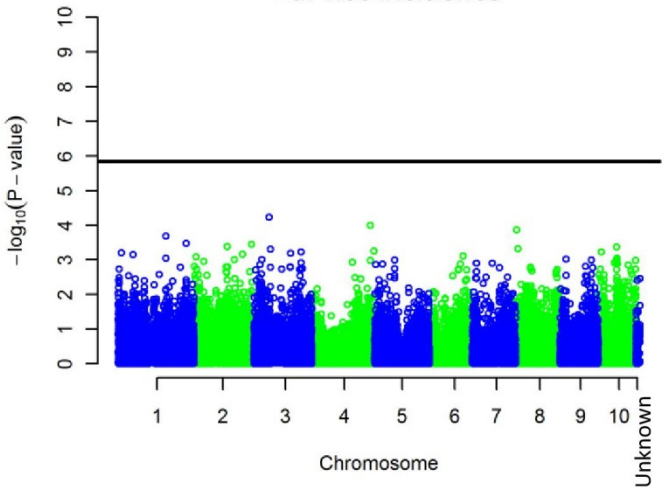

QQ-plot
Observed P-values
Expected P-values
λ = 1.006

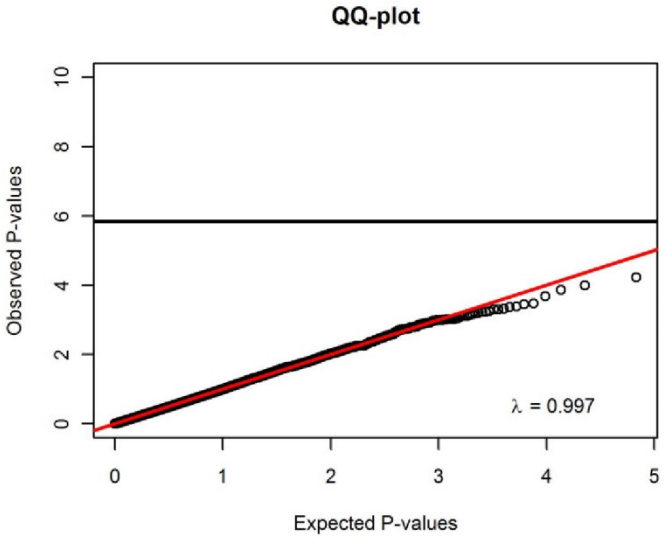

Kernels per Row
$-\log_{10}(P-value)$
Chromosome
Unknown

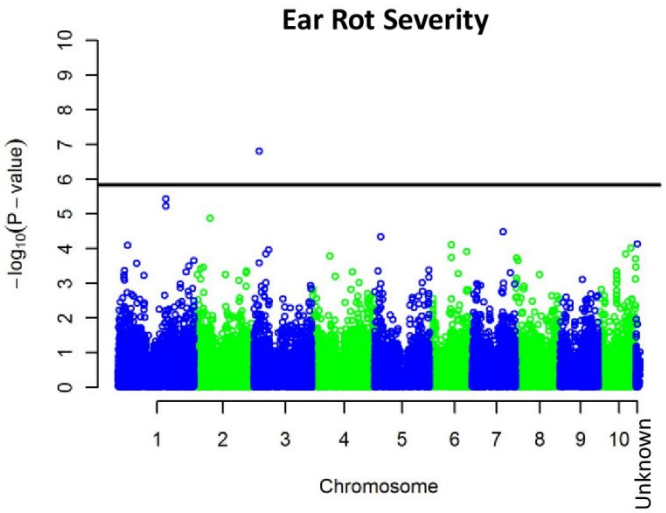

QQ-plot
Observed P-values
Expected P-values
λ = 0.987

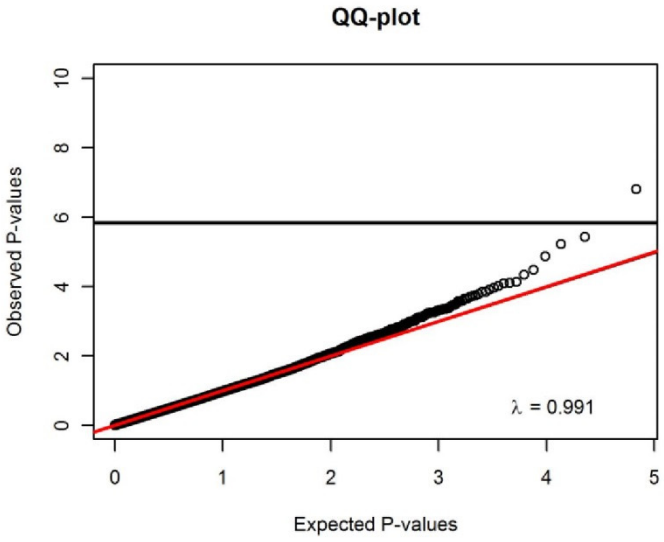

Thousand Kernel Weight
$-\log_{10}(P-value)$
Chromosome
Unknown

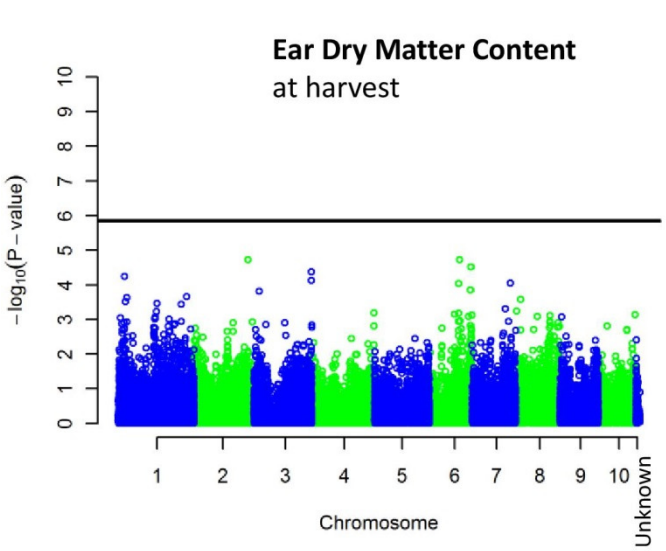

QQ-plot
Observed P-values
Expected P-values
λ = 0.991

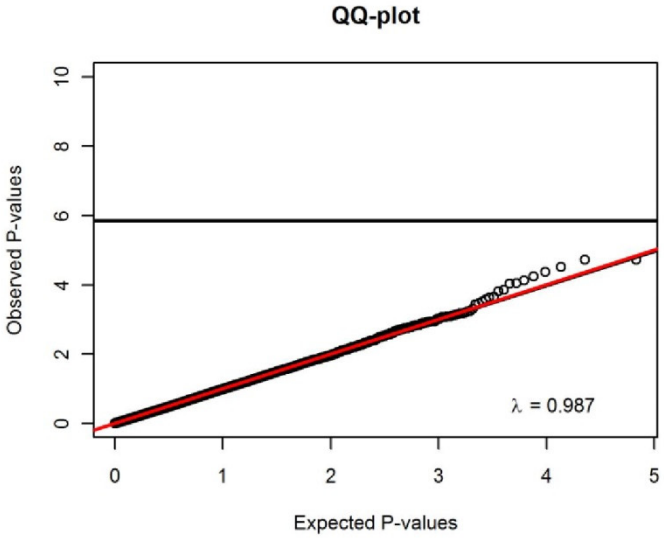

Ear Rot Incidence
−log10(P − value)
Chromosome
Unknown

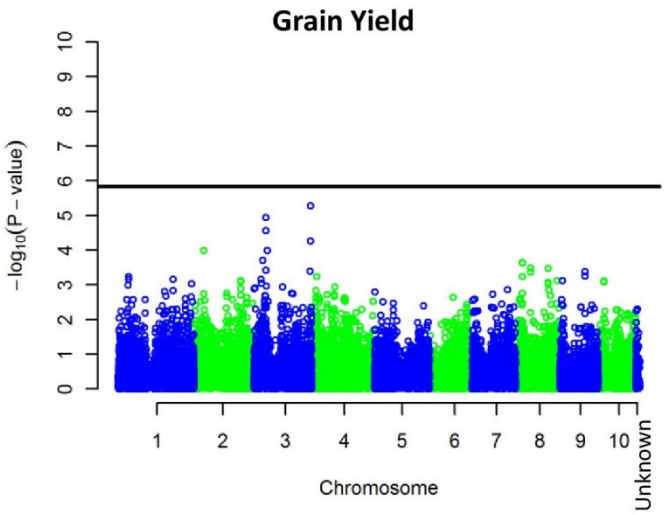

QQ-plot
Observed P-values
Expected P-values
λ = 0.997

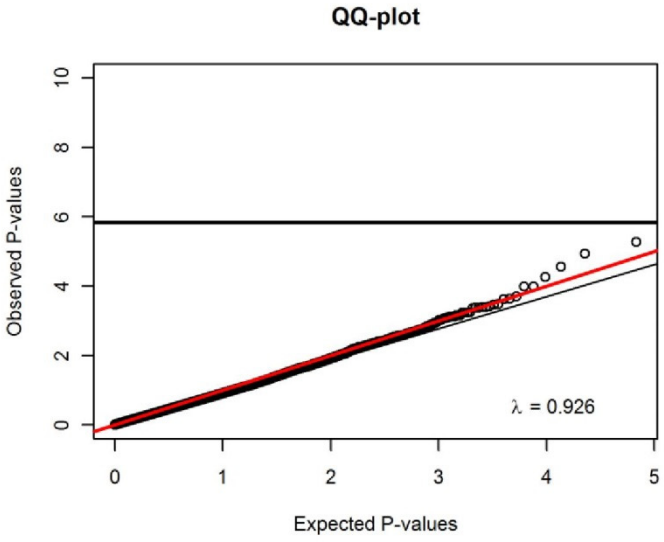

Ear Rot Severity
−log10(P − value)
Chromosome
Unknown

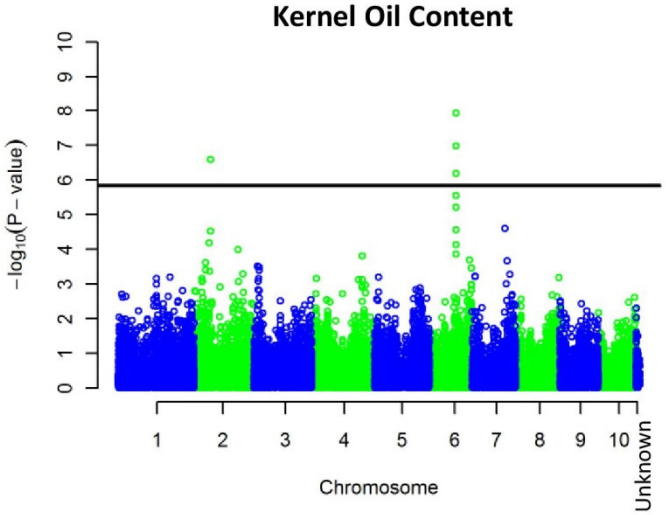

QQ-plot
Observed P-values
Expected P-values
λ = 0.991

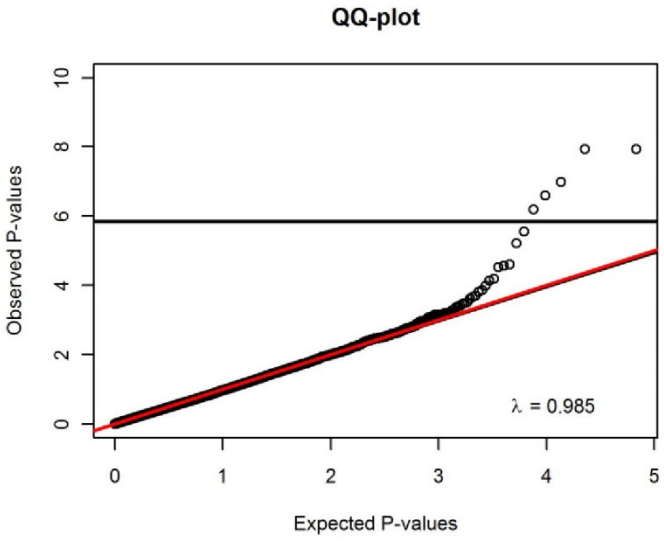

Ear Dry Matter Content
at harvest
$-\log_{10}(P-value)$
Chromosome
Unknown


QQ-plot
Observed P-values
Expected P-values
λ = 0.987


Grain Yield
$-\log_{10}(P-value)$
Chromosome
Unknown


QQ-plot
Observed P-values
Expected P-values
λ = 0.926


Kernel Oil Content
$-\log_{10}(P-value)$
Chromosome
Unknown


QQ-plot
Observed P-values
Expected P-values
λ = 0.985

Annex 4. Position within chromosome (Chr) and QTL assignment of single nucleotide polymorphism (SNP) significantly associated with trait expression in a mapping population composed of 132 doubled-haploid (DH) lines derived from three landraces and elite European dent and flint inbred lines, as well as gene model and putative functions associated to each SNP.

SNP Marker	Chr	Position	QTL	Trait	Putative Product/Function	Gene model
PZE-101039400	1	26788426	1	KERO		
PZE-101085145	1	73273384	2	SFUS		
PZE-101118338	1	144046242	3	$EVIG_4$		
PZE-101121877	1	150857195	4	GERM, EMER, $REGR_{OLI}$		
PZE-101123359	1	154551265	5	GERM, EMER, $REGR_{OLI}$		
PZE-101123390	1	154561075	5	GERM, EMER, $REGR_{OLI}$		
PZE-101123442	1	154727113	5	GERM, EMER, $REGR_{OLI}$		
PZE-101123501	1	154826777	5	GERM, EMER, $REGR_{OLI}$	-	GRMZM2G446047
PZE-101123504	1	154829048	5	GERM, EMER, $REGR_{OLI}$	-	GRMZM2G446047
PZE-101123613	1	154997091	5	GERM, EMER, $REGR_{OLI}$	*rough sheath 2*	GRMZM2G403620
PZE-101158364	1	200449666	6	EAHT	LOC100272686	GRMZM2G136443
SYN2527	1	266032842	7	ROWS	-	GRMZM2G041770
PZB01394.4	1	285421736	8	LODG	Aldehyde oxydase	GRMZM2G141473
PZE-101239111	1	285421905	8	LODG	Aldehyde oxydase	GRMZM2G141473
PZE-102000197	2	387250	9	$EPHT_8$		
PZE-102016498	2	7074136	10	CHLO	-	GRMZM2G379758
PZE-102039123	2	19173678	11	$EFMA_4$	Beta-amylase	GRMZM2G462258
PZE-102052836	2	30793901	12	THKW		
PZE-102080077	2	63546600	13	KOIL	Transcription factor	GRMZM2G436533
PZE-102080077	2	63546600	13	KOIL	LOC100382918	GRMZM2G136412
PZE-103029038	3	21401580	14	SFUS	Beta-lactamase	GRMZM2G150866
PZE-103046561	3	48521587	15	GRYD		
PZE-103053062	3	59474003	16	SFUS		
PZE-103075286	3	120914632	17	SMUT	Cyclin dependent kinase	GRMZM2G018372
PZE-103075286	3	120914632	17	SMUT	-	GRMZM2G018527

SNP Marker	Chr	Position	QTL	Trait	Putative Product/Function	Gene model
PZB01183.1	3	148380841	18	HUFL		
SYN23238	3	182924975	19	$EVIG_8$		
PZE-104017875	4	17692224	20	$REGR_{HOH}$	LOC100279220	GRMZM2G066304
SYN4724	4	139101848	21	$REGR_{HOH}$	LOC100280267	GRMZM2G044882
PZE-104070042	4	139212042	21	$REGR_{HOH}$		
SYN2191	4	158962743	22	$REGR_{HOH}$	LOC100272809	GRMZM2G076631
PZE-104110372	4	185721002	23	HUFL	Auxin efflux carrier component	GRMZM2G171702
PZE-104152153	4	243110299	24	LODG	Plant calmodulin-binding protein-related	GRMZM2G166044
PZE-104157783	4	246144837	25	LODG	-	GRMZM2G037128
PUT-163a-71762647-3467	5	938473	26	HUCO	-	GRMZM2G030858
PZE-105019535	5	9278091	27	LODG	-	GRMZM2G122863
PZE-105024245	5	11918607	28	$TFMA_8$, $MAPL_8$, $EFMA_6$, $EFMA_8$, $EPHT_4$, $EPHT_6$, $EPHT_8$	Extensin-like protein	GRMZM2G157202
PUT-163a-74237711-3635	5	141715894	29	ROWS	-	GRMZM2G159759
PZE-105136017	5	190886598	30	GERM		
SYN12088	6	104659659	31	KOIL	-	GRMZM2G159744
PZE-106054182	6	105013351	31	KOIL	Cation transporter	GRMZM2G169114
PZE-106054189	6	105019334	31	KOIL	*Diacylglycerol acyltransferase 1-2* (DGAT1-2)	GRMZM2G169089
PZE-106054245	6	105119390	31	KOIL	Transmembrane amino acid transporter	GRMZM2G331283
SYN35140	6	115431274	32	LODG	Dynamin-like 3	GRMZM2G157462
PZE-107000845	7	998928	33	LODG	-	GRMZM2G090744
SYNGENTA6482	7	8587447	34	EMER, $REGR_{OLI}$	Frigida-like protein	GRMZM2G011742
SYNGENTA6495	7	8587453	34	EMER, $REGR_{OLI}$	Frigida-like protein	GRMZM2G011742
SYN13685	7	115298441	35	SFUS	Auxin-responsive SAUR family member	GRMZM2G011463
PZE-107071640	7	122113632	36	HUCO	Glucose-6-phosphate isomerase	GRMZM2G140614
PZE-107089380	7	138727405	37	SMUT		
PZE-107097762	7	147529458	38	KOIL	LOC100382369	GRMZM2G449709
SYN13846	7	153910163	39	$REGR_{HOH}$	Endoribonuclease L-PSP	GRMZM2G158452

SNP Marker	Chr	Position	QTL	Trait	Putative Product/Function	Gene model
PZE-107109512	7	154594779	39	LODG, $REGR_{HOH}$	Glucan endo-1,3-beta-glucosidase 4	GRMZM2G072526
PZE-107109652	7	154635539	39	LODG	LOC100276126	GRMZM2G063420
PZE-107113339	7	156452408	40	$REGR_{HOH}$	-	GRMZM2G057260
PZE-107113482	7	156523539	40	$REGR_{HOH}$	Nucleotide-sugar transporter family protein	GRMZM2G089630
PZE-107113712	7	156647570	40	$REGR_{HOH}$	Nuclear transport factor 2	GRMZM2G167932
PZE-107113723	7	156648103	40	$REGR_{HOH}$	Nuclear transport factor 2	GRMZM2G167932
PZE-107130789	7	166217773	41	GERM, $EVIG_8$, $EFMA_8$, $EPHT_8$	LOC100384249	GRMZM2G333433
PZE-108034742	8	44361425	42	MFLO		
PZE-108036458	8	52679949	43	MFLO		
SYN2640	8	69564074	44	ROWS	Elongation factor Tu GTP binding domain	GRMZM2G158024
SYN2640	8	69564074	44	ROWS	Ribosomal RNA large subunit methyltransferase	GRMZM2G158091
SYN2640	8	69564074	44	ROWS	SANT/MYB transcription factor	GRMZM2G158117
PZE-108072761	8	125071835	45	$REGR_{HOH}$		AC210413
PZE-108072784	8	125129384	45	$REGR_{HOH}$		
PZE-108072786	8	125129519	45	$REGR_{HOH}$		
PZE-108072804	8	125165654	45	$REGR_{HOH}$		AC208327
PZE-108072805	8	125168213	45	$REGR_{HOH}$		AC208327
SYN17423	8	134818224	46	MFLO	Calmodulin binding protein	GRMZM2G100229
PZE-108083889	8	139580697	47	$REGR_{HOH}$		
SYN17921	10	75520755	48	$EVIG_4$		GRMZM2G099352
PZE-110054216	10	102859287	49	ROWS		

† For traits description see table 2.

Annex 5. Frequency within population of the positive allele at each marker within quantitative trait loci detected for agronomic and morphological traits in the mapping panel composed of elite European dent (EU-D) and flint (EU-F) inbred lines as well as of doubled haploid (DH) lines derived from the landraces *Bugard* (DH-BU), *Gelber Badischer* (DH-GB), and *Schindelmeiser* (DH-SC).

			Frequency of the positive QTL allele				
Marker	**QTL**	**Trait**	**EU-D**	**EU-F**	**DH-BU**	**DH-GB**	**DH-SC**
PZE-101039400	1	KERO	0.04	0.04	0.00	0.80	0.76
PZE-101085145	2	SFUS	0.99	0.97	0.89	0.90	0.63
PZE-101118338	3	$EVIG_4$	0.03	0.38	0.19	1.00	0.98
PZE-101121877	4	GERM, EMER, $REGR_{OLI}$	0.42	0.54	0.47	1.00	1.00
PZE-101123359	5	GERM, EMER, $REGR_{OLI}$	0.57	0.86	0.64	1.00	1.00
PZE-101123390	5	GERM, EMER, $REGR_{OLI}$	0.57	0.86	0.64	1.00	1.00
PZE-101123442	5	GERM, EMER, $REGR_{OLI}$	0.97	0.96	0.66	1.00	1.00
PZE-101123501	5	GERM, EMER, $REGR_{OLI}$	0.97	0.96	0.64	1.00	1.00
PZE-101123504	5	GERM, EMER, $REGR_{OLI}$	0.97	0.96	0.64	1.00	1.00
PZE-101123613	5	GERM, EMER, $REGR_{OLI}$	0.90	1.00	0.44	1.00	1.00
PZE-101158364	6	EAHT	0.98	0.92	0.86	1.00	0.98
SYN2527	7	ROWS	0.45	0.01	0.11	0.03	0.00
PZB01394.4	8	LODG	0.99	0.97	1.00	0.93	0.82
PZE-101239111	8	LODG	0.95	0.94	1.00	0.93	0.82
PZE-102000197	9	$EPHT_8$	1.00	0.64	0.09	1.00	0.89
PZE-102016498	10	CHLO	0.83	1.00	1.00	1.00	1.00
PZE-102039123	11	$EFMA_4$	0.02	0.04	0.00	0.60	0.52
PZE-102052836	12	THKW	0.92	0.90	0.61	0.97	0.56
PZE-102080077	13	KOIL	0.20	0.09	0.69	0.87	0.50
PZE-103029038	14	SFUS	0.91	0.95	1.00	0.90	0.82
PZE-103046561	15	GRYD	0.05	0.23	0.17	0.80	0.52
PZE-103053062	16	SFUS	0.54	1.00	0.88	0.93	0.98
PZE-103075286	17	SMUT	1.00	0.99	0.75	0.90	1.00
PZB01183.1	18	HUFL	1.00	1.00	0.97	0.97	0.79
SYN23238	19	$EVIG_8$	0.40	0.95	0.83	0.97	1.00
PZE-104017875	20	$REGR_{HOH}$	0.71	0.88	1.00	1.00	1.00
SYN4724	21	$REGR_{HOH}$	0.62	0.90	0.97	0.97	1.00
PZE-104070042	21	$REGR_{HOH}$	0.31	0.88	0.97	0.97	1.00
SYN2191	22	$REGR_{HOH}$	0.72	0.88	1.00	1.00	1.00
PZE-104110372	23	HUFL	1.00	1.00	0.81	0.76	1.00
PZE-104152153	24	LODG	1.00	1.00	1.00	1.00	0.84
PZE-104157783	25	LODG	0.91	0.83	0.81	0.23	0.45
PUT-163a-71762647-3467	26	HUCO	0.02	0.96	1.00	0.93	0.82
PZE-105019535	27	LODG	0.95	0.98	0.58	0.60	0.90

Marker	QTL	Trait	Frequency of the positive QTL allele				
			EU-D	**EU-F**	**DH-BU**	**DH-GB**	**DH-SC**
PZE-105024245	28	$EFMA_6$, $EFMA_8$, $TFMA_8$, $MAPL_8$, $EPHT_4$, $EPHT_6$, $EPHT_8$	0.01	0.04	0.00	0.30	0.34
PUT-163a-74237711-3635	29	ROWS	0.21	0.89	0.06	0.00	0.00
PZE-105136017	30	GERM	0.91	1.00	1.00	1.00	1.00
SYN12088	31	KOIL	0.54	0.73	0.31	0.83	0.37
PZE-106054182	31	KOIL	0.04	0.62	0.00	0.20	0.44
PZE-106054189	31	KOIL	0.04	0.62	0.00	0.20	0.44
PZE-106054245	31	KOIL	0.95	0.84	0.39	0.72	0.44
SYN35140	32	LODG	0.98	0.96	1.00	0.90	0.47
PZE-107000845	33	LODG	0.00	0.29	0.14	0.83	0.82
SYNGENTA6482	34	EMER, $REGR_{OLI}$	0.06	0.83	0.58	1.00	1.00
SYNGENTA6495	34	EMER, $REGR_{OLI}$	0.62	0.88	0.68	1.00	1.00
SYN13685	35	SFUS	0.64	0.98	1.00	0.87	0.90
PZE-107071640	36	HUCO	0.51	0.61	1.00	1.00	0.85
PZE-107089380	37	SMUT	0.90	0.68	0.72	0.48	0.56
PZE-107097762	38	KOIL	0.55	0.85	0.67	0.86	0.65
SYN13846	39	$REGR_{HOH}$	0.74	0.97	1.00	1.00	1.00
PZE-107109512	39	LODG, $REGR_{HOH}$	0.73	0.95	0.78	1.00	1.00
PZE-107109652	39	LODG	0.99	0.97	0.89	0.90	0.63
PZE-107113339	40	$REGR_{HOH}$	0.77	0.98	1.00	1.00	1.00
PZE-107113482	40	$REGR_{HOH}$	0.84	0.96	1.00	1.00	1.00
PZE-107113712	40	$REGR_{HOH}$	0.80	0.92	1.00	1.00	1.00
PZE-107113723	40	$REGR_{HOH}$	0.79	0.21	0.22	0.97	0.21
PZE-107130789	41	GERM, $EVIG_8$, $EFMA_8$, $EPHT_8$	0.38	0.68	0.72	0.57	0.03
PZE-108034742	42	MFLO	1.00	0.64	0.09	1.00	0.89
PZE-108036458	43	MFLO	0.64	0.95	0.97	1.00	1.00
SYN2640	44	ROWS	0.86	0.13	0.14	0.00	0.00
PZE-108072761	45	$REGR_{HOH}$	0.87	0.89	0.83	1.00	1.00
PZE-108072784	45	$REGR_{HOH}$	0.87	0.89	0.83	1.00	1.00
PZE-108072786	45	$REGR_{HOH}$	0.87	0.89	0.81	1.00	1.00
PZE-108072804	45	$REGR_{HOH}$	0.30	0.88	0.83	1.00	1.00
PZE-108072805	45	$REGR_{HOH}$	0.87	0.89	0.83	1.00	1.00
SYN17423	46	MFLO	0.61	0.90	1.00	1.00	1.00
PZE-108083889	47	$REGR_{HOH}$	0.84	0.88	1.00	1.00	1.00
SYN17921	48	$EVIG_4$	0.85	0.94	0.53	0.97	1.00
PZE-110054216	49	ROWS	0.42	0.02	0.00	0.00	0.00

† For traits description see table 2.