Technische Universität München
TUM School of Life Sciences

Genomic prediction in European flint maize landraces

Armin Christoph Hölker

Vollständiger Ausdruck der von der TUM School of Life Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation

Vorsitz:		Prof. Dr. Aurélien Tellier
Prüfende Dissertation:	1.	Prof. Dr. Chris-Carolin Schön
	2.	Prof. Dr. Tobias Würschum

Die Dissertation wurde am 03.01.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Life Sciences am 13.05.2024 angenommen.

Content

Summary

Global population growth is a major challenge to the agricultural sector. While plant breeding is a key contributor to food security, a fundamental prerequisite for germplasm improvement is the availability of sufficient genetic variation for target traits. Today's elite germplasm captures only a small proportion of the species-wide genetic diversity for most crops, thus broadening the genetic base of elite germplasm is an important task. A very rich source for additional diversity are landraces that are stored in seed banks all around the world. Their use is hampered by a lack of information (both, genetic and phenotypic), their heterogeneity and heterozygosity (in allogamous crops), and a performance gap to current elite germplasm. While the performance gap is only of minor importance in cases where the aim is to use single genes of qualitative traits (e.g. disease resistances) derived from landrace material, the utilization of landraces for improving quantitative traits in elite germplasm is more difficult. Efficient strategies for the successful utilization of landraces for this purpose are still in development.

In this thesis, three preselected European flint maize landraces were subjected to two different strategies of population development and the developed populations were characterized genotypically and phenotypically. The two contrasting strategies for obtaining populations of reproducible genetic units were the production of doubled-haploid lines (DH, as a pure approach without elite introgression) and crosses to an important European flint maize founder inbred line with subsequent selfing from the crosses (termed gamete capture (GC), as an example for an admixed approach). The DH and GC populations were genotyped at high density and phenotyped for over 25 traits in up to eleven environments to link molecular data to meaningful phenotypes. This unique resource allows to compare the different sampling strategies and study the prospects of genomic selection for landrace material in order to drive the development of landrace utilization strategies.

Both approaches (DH and GC) were able to capture the diversity present in the landrace in an unbiased way. With the theoretical framework developed as part of this thesis, the expectations for molecular and genetic variation in the derived populations can be formulated and were validated experimentally. Phenotypically, the landrace-derived populations did show the expected yield gap (15-20%). Genomic prediction accuracies reached promisingly high levels when calibration set size was sufficient (minimum $N > 200$, but no plateau reached up to $N = 450$). Prediction accuracies were highest within DH populations, followed by predictions within GC populations and predictions across the two

population types (from DH to GC and vice versa). Prediction accuracies in across landrace predictions were practically zero for DH populations, but higher for GC populations, which is to be expected because of the shared inbred line in the pedigree of all GC lines. In general, the prediction accuracies indicate that both pre-breeding strategies should include the use of genomic selection.

Although unbiased sampling from landraces was successful in both populations, the pure approach was shown to have several advantages over the admixed approach. The lines exhibit the full additive variance, show no masking of genetic effects by the crossing partner's alleles, no risk of reconstruction of the elite crossing partner genome in selection and a generally higher prediction accuracy. However, the efficiency of DH production varies greatly among different landraces. If the admixed approach is followed, the informed choice of the crossing partner is of crucial importance as it has large impact on the resulting populations. The choice between the pure and the admixed approach depends on if the production of DH lines is possible in the used genetic resource or crop (in extension to other allogamous crops beyond maize) and has to be evaluated together with the goal and the structure of the entire pre-breeding strategy. The results from this thesis might be complemented by additional research considering rapid genomic selection scenarios (for faster closing of the performance gap). The presented resource is an excellent starting point for such studies.

Zusammenfassung

Die Ernährung der wachsenden Weltbevölkerung ist eine große Herausforderung für die Landwirtschaft. Die Pflanzenzüchtung leistet einen wichtigen Beitrag zur Überwindung dieser Herausforderung. Eine Voraussetzung für den Selektionsfortschritt, und damit auch den Erfolg der Pflanzenzüchtung insgesamt, ist die Verfügbarkeit genetischer Variation. Für die meisten Kulturarten deckt das Elitematerial jedoch nur einen sehr kleinen Teil der speziesweiten Variation der Kulturart ab, weshalb eine Erweiterung der genetischen Basis des Elitematerials eine wichtige Aufgabe ist. Eine umfangreiche Quelle für momentan ungenutzte Diversität sind die in Genbanken auf der ganzen Welt aufbewahrten Landrassen. Deren Nutzung wird jedoch durch einen Mangel an Informationen über sie (phänotypisch sowie genotypisch), ihre Heterogenität und Heterozygotie sowie ihren Rückstand in der Ertragsleistung gegenüber Elitematerial gehemmt. Während für die Nutzung von Landrassen zur Verbesserung von qualitativen Merkmalen (z.B. Krankheitsresistenzen) das Ertragsdefizit ein eher geringes Problem darstellt, ist die Nutzung von Landrassen zur Verbesserung von quantitativen Merkmalen im Elitematerial eine große Herausforderung. Die dafür benötigten effizienten Nutzungsstrategien befinden sich in der Entwicklung.

In dieser Dissertation wurden drei vorselektierte europäische Hartmais-Landrassen zwei unterschiedlichen Populationsentwicklungsstrategien unterzogen und die entwickelten Populationen anschließend phänotypisch und genotypisch charakterisiert. Die beiden angewandten Strategien zur Erzeugung von Populationen aus reproduzierbaren genetischen Einheiten waren die Produktion von Doppelhaploiden (DH) Linien (als Beispiel für ein Vorgehen ohne Beimischung von Elitematerial) und Kreuzungen mit einer wichtigen Gründerlinie des europäischen Hartmais Elite-Genpools (als „gamete capture“ (GC) bezeichnet, als Beispiel für ein Vorgehen mit Beimischung von Elitematerial). Die DH und GC Populationen wurden hochauflösend genotypisiert und über 25 Merkmale in bis zu 11 Umwelten phänotypisch erfasst. Diese einzigartige Ressource erlaubt es, die beiden Strategien eingehend zu vergleichen und die Nutzung von genomischer Selektion im Landrassenmaterial zu untersuchen.

Die Populationen aus beiden Ansätzen repräsentierten die Diversität der Landrassen ohne nennenswerte Verluste. Die Erwartungswerte der molekularen und genetischen Variation in den jeweiligen Populationen wurden mit Hilfe der in dieser Arbeit vorgestellten Theorie vorhergesagt und experimentell bestätigt. Die Populationen aus Landrassen zeigten den erwarteten Ertragsrückstand (15-20%). Die Genauigkeit der genomischen Vorhersage war

hoch, wenn die Größe des Trainingsdatensatzes ausreichend groß war (mindestens N > 200, aber kein Plateau bis N = 450 erreicht). Die Vorhersagegenauigkeit war am höchsten bei Vorhersagen innerhalb der DH Populationen, gefolgt von Vorhersagen innerhalb der GC Populationen und Vorhersagen zwischen den beiden Populationsarten (von DH zu GC und *vice versa*). Die Vorhersagegenauigkeit zwischen Landrassen war praktisch Null bei DH und etwas höher für GC Populationen, was durch die von allen GC Individuen geteilte Inzuchtlinie im Stammbaum zu erklären ist. Insgesamt zeigen die Vorhersagegenauigkeiten, dass jede der beiden Züchtungsstrategien mit Landrassenmaterial genomische Vorhersagen nutzen sollte.

Obwohl die Landrassen durch beide Methoden der Populationsentwicklung gut repräsentiert werden können, konnte gezeigt werden, dass der Ansatz ohne Beimischung von Elitematerial mehrere Vorzüge mit sich bringt. Die Linien zeigen die volle additive Varianz, es gibt keine Maskierung von Landrassen-Allelen durch die Allele des Kreuzungspartners, kein Risiko einer Rekonstruktion des Genoms des Elite-Kreuzungspartners und eine insgesamt höhere Vorhersagegenauigkeit in der genomischen Selektion. Die Effizienz der DH Produktion variiert jedoch sehr stark zwischen verschiedenen Landrassen. Wenn der Ansatz mit Beimischung von Elitematerial verfolgt wird, ist eine fundierte Auswahl des Kreuzungspartners von sehr hoher Bedeutung, weil dieser einen sehr starken Einfluss auf die produzierten Populationen hat. Welche Methode zu bevorzugen ist hängt davon ab, ob eine DH Produktion in den ausgewählten Landrassen oder Kulturarten (bei anderen Kulturarten außer Mais) überhaupt möglich ist. Außerdem muss die Populationsentwicklung im Kontext der gesamten Züchtungsstrategie evaluiert werden. Die Ergebnisse dieser Dissertation können durch zusätzliche Forschung über die Verwendung schneller Zuchtschemata mit wiederholter genomischer Selektion erweitert werden. Die hier vorgestellten Ressourcen sind ein exzellenter Startpunkt für nachfolgende Forschung.

List of Figures

List of Tables

List of Abbreviations

AMOVA	analysis of molecular variance
ASI	anthesis-silking-interval
BLUP	best linear unbiased prediction
BLUE	best linear unbiased estimate
CT	cold tolerance
CV	cross-validation
DH	doubled-haploid
DMC	dry matter content
DT	drought tolerance
EH	ear height
EME	plant emergence
EV	early vigor
exPVP	expired Plant Variety Protection Act certificate
FF	female flowering
Fv/Fm	maximum photosystem II efficiency
GBLUP	genomic best linear unbiased prediction
GC	gamete capture
GD	genetic distance
GEBV	genomic estimated breeding value
GEM	Germplasm Enhancement of Maize
GP	genomic prediction
GS	genomic selection
HOPE	Hierarchical Open-Ended Corn Breeding System
KE	Kemater Landmais Gelb
LAMP	Latin American Maize Project
LD	linkage disequilibrium
LL	Lalin
LP	line per se
LPS	linkage phase similarity
LS	ancestral landrace sample
Mb	megabase pairs
MF	male flowering
MRD	modified Rogers' distance

NB	number of tassel branches
NIRS	near infrared spectroscopy
PE	Petkuser Ferdinand Rot
PH	plant height
QTL	quantitative trait locus
RL	root lodging
RR-BLUP	ridge regression best linear unbiased prediction
SL	spike length
SM	simple matching coefficient
SNP	single nucleotide polymorphism
SPAD	soil and plant analysis development chlorophyll meter
TA	tassel angle
TILL	tillering
TC	testcrosses
TDMY	total dry matter yield
TL	tassel length
UPGMA	unweighted pair group method with arithmetic mean
V3/V4/V6	vegetative growth stage with 3/4/6 leaves

Publications included in this thesis

Hölker et al. (2019)

Hölker AC, Mayer M, Presterl T, Bolduan T, Bauer E, Ordas B, Brauner PC, Ouzunova M, Melchinger AE, Schön C-C (2019) European maize landraces made accessible for plant breeding and genome-based studies. Theor Appl Genet 132:3333-3345, doi: 10.1007/s00122-019-03428-8.

Abstract

Maize landraces comprise large allelic diversity. We created doubled-haploid (DH) libraries from three European flint maize landraces and characterized them with respect to their molecular diversity, population structure, trait means, variances, and trait correlations. In total, 899 DH lines were evaluated using high-quality genotypic and multi-environment phenotypic data from up to 11 environments. The DH lines covered 95% of the molecular variation present in 35 landraces of an earlier study and represent the original three landrace populations in an unbiased manner. A comprehensive analysis of the target trait plant development at early growth stages as well as other important agronomic traits revealed large genetic variation for line per se and testcross performance. The majority of the 378 DH lines evaluated as testcrosses outperformed the commercial hybrids for early development. For total biomass yield, we observed a yield gap of 15% between mean testcross yield of the commercial hybrids and mean testcross yield of the DH lines. The DH lines also exhibited genetic variation for undesirable traits like root lodging and tillering, but correlations with target traits early development and yield were low or nonsignificant. The presented diversity atlas is a valuable, publicly available resource for genome-based studies to identify novel trait variation and evaluate the prospects of genomic prediction in landrace-derived material.

Candidate's contribution

The candidate made major contributions to planning of field trials and collection of phenotypic data, analyzed the data, created figures and tables, wrote the manuscript draft and made revisions.

Hölker et al. (2022)

Hölker AC, Mayer M, Presterl T, Bauer E, Ouzunova M, Melchinger AE, Schön C-C (2022) Theoretical and experimental assessment of genome-based prediction in landraces of allogamous crops. Proc Natl Acad Sci USA 119:e2121797119, doi: 10.1073/pnas.2121797119.

Abstract

Discovery and enrichment of favorable alleles in landraces is key to make them accessible for crop improvement. Here, we present two fundamentally different concepts for genome-based selection in landrace derived maize populations, one based on doubled-haploid lines derived directly from individual landrace plants, the other based on crossing landrace plants to a capture line. For both types of populations, we show theoretically how allele frequencies of the ancestral landrace and the capture line translate into expectations for molecular and genetic variances. We show that the DH approach has clear advantages over gamete capture with generally higher prediction accuracies and no risk of masking valuable variation of the landrace. Prediction accuracies as high as 0.58 for dry matter yield in the DH population indicate high potential of genome-based selection. Based on a comparison among traits, we show that the genetic make-up of the capture line has great influence on the success of genome-based selection and that confounding effects between the alleles of the landrace and the capture line are best controlled for traits for which the capture line does not outperform the ancestral population per se or in testcrosses. Our results will guide the optimization of genome-enabled pre-breeding schemes.

Candidate's contribution

The candidate made major contributions to conceiving the study, planning of field trials and phenotypic data collection, analyzed the phenotypic data, carried out genome-based prediction analysis and made major contributions to the first manuscript draft and manuscript revisions.

1 Introduction

1.1 Background

The global population is growing and forecasted to surpass nine billion within the next 20 years and reach ten billion before 2060 (United Nations, 2019). Additional pressure is applied to the food supply system by increasing wealth in many countries, resulting in greater demands for processed food, meat, dairy, and fish, which require more resources per calory than plant-based food (Godfray et al. 2010), as well as by the increasing demand for crops for fiber and fuel use (Edgerton 2009). Agricultural production is additionally challenged by the effects of climate change, land shortages due to urbanization, and other consequences of suboptimal land management like desertification, salinization, and soil erosion (Nellemann and MacDevette 2009). As a consequence, an immense increase in agricultural production has to be realized in a sustainable way that allows adaptation to the changing environmental conditions (Godfray et al. 2010; Lobell and Tebaldi 2014). A tremendous production gain of 125% was achieved during the so-called green revolution between 1960 and 2000 (Khush 2001). The main driver of production increase during the green revolution was genetic improvement and the widespread adoption of the improved varieties (Khush 2001). Plant breeding is of critical importance also for today's challenges, yet it relies on optimal exploitation of available genetic variation for ensuring the required genetic gains. However, current elite germplasm of many crop species represents only a small part of the total available diversity (McCouch et al. 2013) because it went through a series of diversity reducing bottlenecks in domestication and plant breeding (Tanksley and McCouch 1997; Mir et al. 2013; Russell et al. 2016). The evolution of diversity through those bottlenecks has been graphically described for three classes of genes by Yamasaki et al. (2005) (Figure 1). Neutral genes are expected to be only mildly affected by bottlenecks (genetic drift), while the diversity of domestication genes is heavily reduced already early, during the step from wild ancestor to landraces. The diversity in improvement genes is still large in landraces but greatly reduced in the step from landraces to modern inbred lines. The seed banks around the world harbor thousands of untapped landrace accessions with large remaining genetic diversity and the utilization of those resources is considered essential to leverage plant breeding for a sustainable intensification of agricultural production (Hoisington et al. 1999; Ortiz et al. 2010; McCouch et al. 2013).

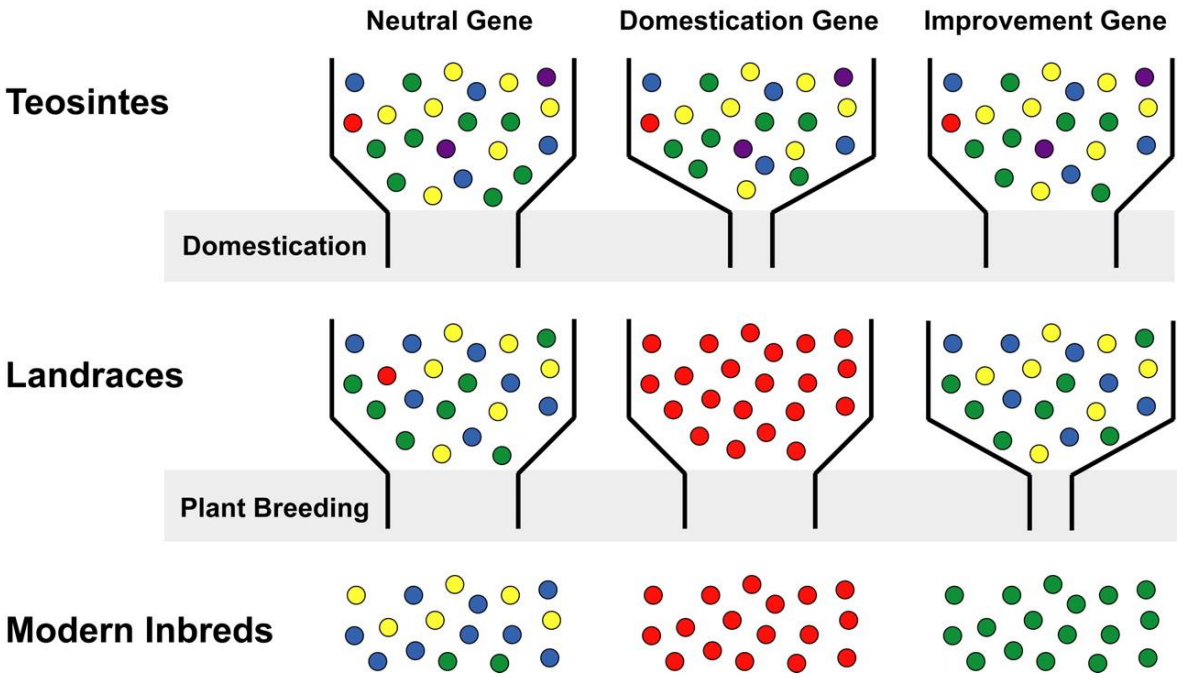


Figure 1: Effect of domestication and plant breeding on genetic diversity of maize genes (Yamasaki et al. 2005).

Maize (*Zea mays* L. ssp. *mays)* is the geographically most ubiquitous cereal (Leff et al. 2004), a critical staple crop worldwide (Nuss and Tanumihardjo 2010), and also an important model organism in biological research (Strable and Scanlon 2009). After the domestication from the wild relative teosinte (*Zea mays* L. ssp. *parviglumis)* approximately 10,000 years ago in today's Mexico (Piperno et al. 2009), maize has spread across the world, first within the Americas and after the discovery of the New World by Columbus also to Europe (Rebourg et al. 2003) and into Africa and Asia (Mir et al. 2013). The domestication and spread of maize across the world were accompanied by a series of genetic bottlenecks. During domestication, allelic diversity was reduced due to selection for domestication phenotypes such as increased ear size and no lateral branching (Piperno et al. 2009; Hufford et al. 2012). The following spread during the Americas took place over thousands of years (Mir et al. 2013), resulting in a high number of locally adapted landraces with a negative correlation between distance to the center of origin and diversity (Vigouroux et al. 2008). The introduction to Europe was very rapid with strong founder effects and followed by local adaptation through selection by farmers, leading to the creation of several hundred new landraces (Dubreuil et al. 2006). The thousands of landraces created during the global maize diffusion process maintained a high level of species-wide phenotypic and genetic variation across the globe (Buckler et al. 2006; Warburton et al. 2008; Sood et al. 2014). The change from open-pollinated to hybrid varieties more than 100 years ago was

a turning point for maize breeding, leading to tremendous yield increases on the one hand (Duvick 2005), but also strong genetic bottlenecks on the other hand. With the introduction of hybrid breeding, the heterotic pools of today were founded by intermating a small number of lines derived by selfing in landraces and subsequent breeding was characterized mainly by crossing within the established breeding pools (Messmer et al. 1992; White et al. 2020). As a result, the genetic basis of the current elite germplasm is comparatively narrow (Tanksley and McCouch 1997; Allier et al. 2019). To broaden the genetic diversity, new variation can be introduced by different approaches like crossing elite germplasm to non-adapted elite material or tapping the large diversity present in landraces and introduce it into the elite germplasm. The genetic material from which breeders can choose differs mainly in two important characteristics: level of adaptation to the target environments and performance gap to elite germplasm. Although landraces always exhibit a performance gap they are considered the prime resource for broadening the genetic basis of elite germplasm due to large genetic variation within and between them (Pollak 2003; Salhuana and Pollak 2006; Warburton et al. 2008; Strigens et al. 2013; McCouch et al. 2013; Navarro et al. 2017).

The use of genetic resources like landraces for improving mono- or oligogenic traits has been shown to be comparatively straightforward, as screening and mapping large numbers of accessions for large effect genes and their subsequent introgression into elite germplasm through backcrossing is feasible (Visscher et al. 1996; McCouch et al. 2012). Several examples for the application of this strategy exist for disease resistances (Marone et al. 2021), or abiotic stress tolerance like submergence (Bailey-Serres et al. 2010) or boron tolerance (Paull et al. 1992). However, most agronomically important traits are of a polygenic nature and examples for successful use of landraces for improving those traits are scarce (Sood et al. 2014).

Recently, it has been shown that landraces carry beneficial haplotypes that are not present in elite germplasm (Mayer et al. 2020; Würschum et al. 2022). When judging the breeding potential of landraces for specific quantitative traits the first challenge is the limited information available on seed bank accessions. A preselection solely relying on passport data might therefore be of limited value apart from the possibility to exclude e.g. non-adapted material based on its geographic origin. A possible source of information for characterizing germplasm in gene banks is genotypic data obtained through technologies such as array genotyping or whole-genome sequencing, which can be used for population genetic analyses of, e.g., molecular variation, linkage disequilibrium or population structure

(Mayer et al. 2017). Although cost-reduction in genotyping and sequencing technologies allows the generation of large-scale molecular inventories of plant genetic resources (Wang et al. 2017; Darrier et al. 2019), meaningful phenotypic data is usually lacking, constituting a major hindrance for the utilization of landraces in breeding (McCouch et al. 2012; Sood et al. 2014). In addition, if the crop is allogamous like maize, the landrace accessions themselves are collections of heterogeneous and heterozygous individuals, making phenotypic evaluation and maintenance difficult. For evaluation and breeding applications, reproducible genetic units are required. Different approaches for obtaining reproducible genetic units can be followed. They can mainly be divided into two fundamentally different concepts: landrace material can be used directly, keeping it “pure” without elite introgression, or crosses with elite material can be carried out, resulting in an admixed approach. Independent of the approach applied for obtaining reproducible genetic units, landrace-derived material usually shows a substantial performance gap for several traits compared to elite germplasm due to the presence of deleterious and unfavorable alleles, requiring efficient strategies for pre-breeding before introgression into current elite germplasm pools (Hallauer and Sears 1972; Holland et al. 1996; Böhm et al. 2017; Brauner et al. 2019).

For the pure pre-breeding approach, inbred lines can serve as reproducible genetic units. They can be produced directly from landraces by recurrent selfing, or, for accelerating the procedure, by in vivo haploid induction (Geiger 2009). It has been shown that the production of libraries of doubled-haploid (DH) lines from landraces is feasible (Hölker et al. 2019a) and that the libraries capture the allelic diversity of the original landraces in an unbiased way (Melchinger et al. 2017), but with some loss of diversity (Zeitler et al. 2020). In addition, DH lines have been proposed not only as a basis for breeding with landraces but also for maintaining the landraces themselves, as they are immortal resources that can be multiplied ad libitum (Melchinger et al. 2017).

An example for a coordinated project for utilizing landrace diversity in an admixed approach through crosses with elite germplasm is the Germplasm Enhancement of Maize (GEM) program (Pollak 2003), which followed the Latin American Maize Program (LAMP (Salhuana and Pollak 2006)). In LAMP, promising landrace accessions were identified and in GEM they were crossed to two proprietary elite inbred lines of industry cooperators in successive generations and subsequently inbred (Salhuana and Pollak 2006). After selection for adaptation to target environments in early generations, S_2-lines were selected based on their testcross performance (Salhuana and Pollak 2006), resulting in the

registration of the best performing lines from GEM. However, even after intense selection, the material showed rather poor agronomic properties and a substantial yield gap in comparison with elite materials (Salhuana and Pollak 2006).

While LAMP and GEM were solely based on phenotypic selection, modern pre-breeding strategies must consider the use of genomic prediction (GP). The possibility of assessing far more individuals in the lab compared to the field lead to increased selection intensity and higher genetic gain with the use of GP in elite germplasm (Crossa et al. 2017). The idea to predict the genetic value of selection candidates based on whole-genome molecular marker data was first suggested by Bernardo (1994) for the prediction of hybrid performance. The current form of genomic selection was proposed for cattle breeding in the landmark paper by Meuwissen et al. (2001), who used genome-wide dense marker maps and best linear unbiased prediction (BLUP) (Henderson 1975) as well as two Bayesian methods to establish the association between marker information and phenotypic performance. Genomic selection was adopted in animal (Hayes et al. 2009) and plant breeding (Crossa et al. 2014) and further development of additional methods, e.g. Bayesian methods (Gianola et al. 2009), and machine learning (Maenhout et al. 2010), followed. The statistical method for a particular application of GP needs to be chosen in accordance with the assumptions associated with the respective method and there is no best method for all traits and populations (Heslot et al. 2012). While for genomic predictions with the original BLUP approach, ridge regression is used (RR-BLUP), the genomic BLUP (GBLUP) uses a kinship matrix calculated from the marker data. Habier et al. (2007) have shown the equivalence of both approaches. GBLUP has the advantage of reduced complexity and therefore increased efficiency over RR-BLUP, if the number of markers is larger than the number of individuals in the prediction. GBLUP has developed into the benchmark method for genomic prediction due to its robustness in many different scenarios and, compared to Bayesian methods, a relatively low computational burden (Bernardo 2020).

In GP, the most important drivers of prediction accuracy are training set size, heritability of the phenotypic data, linkage disequilibrium (LD) and relatedness among genotypes in the training and prediction sets (Habier et al. 2007; Auinger et al. 2021). In autogamous crops, genetic variation within landraces is often low, so that prediction accuracy across landraces is of primary interest. High prediction accuracies for those types of predictions were observed in wheat (Crossa et al. 2016) and sorghum (Yu et al. 2016), making it promising to mine natural variation present in gene banks across a broad collection of

landraces. In contrast, in allogamous crops, a large proportion of the genetic variation is found within landraces (Böhm et al. 2017; Mayer et al. 2017), making the prediction of genetic values within landraces the prime target. Landraces of allogamous crops show much lower LD than elite germplasm (Mayer et al. 2017) and, if not artificially introduced during the multiplication of the landraces or the production of reproducible genetic units, progenies derived from landraces also have low relatedness because landraces have a history of random mating. The lower levels of LD and relatedness generally make employing GP in landraces more challenging compared to elite germplasm. Nevertheless, there are several cases for which the use of GP has been recommended for increased genetic gains or more efficient pre-breeding efforts. Studies with small DH populations derived from European flint maize landraces showed first promising indication of sufficient prediction accuracy for within-landrace prediction (Brauner et al. 2018). Genomic prediction models were suggested for identifying suitable genetic resources for enriching elite germplasm (Allier et al. 2020a), and simulation studies showed the usefulness of GP in pre-breeding strategies for improvement of landrace material prior to its introgression into elite germplasm (Gorjanc et al. 2016).

1.2 Outline

When working with genetic resources, the first decision to take is which of the vast amount of available genetic resource accessions to choose for the study. The available resources for genotyping and phenotyping can either be allocated to many accessions from various origins each with few individuals, thus maximizing overall diversity, or, in allogamous crops like maize, to large numbers of individuals from few landrace accessions for exploiting both, within- and between-landrace diversity. Sampling few individuals from a wide range of landraces can be advantageous for studying signals of adaptation (Navarro et al. 2017), crop evolution (Heerwaarden et al. 2011), or the effects of rare alleles (Kremling et al. 2018) but has limitations in breeding due to confounding effects of adaptation. This thesis follows the proposition of Mayer et al. (2017) to focus on a limited set of preselected landraces. In this study, large numbers of progenies from three different landraces were derived. The landraces were chosen according to their phenotypic and molecular variation (Hölker et al. 2019a) and for their adaptation to the European conditions to avoid undesired large effects of adaptative alleles, a confounding factor in the study of Navarro et al. (2017).

After the decision regarding the source material, reproducible genetic units are required for further study and for linking phenotypic and genotypic information. For obtaining reproducible genetic units for this thesis, the three selected landraces were subjected to

two different types of population development that represent the pure approach (without elite introgression) and the admixed approach (with elite introgression). In Hölker et al. (2019a), as an example of the pure approach, ~1,000 doubled-haploid (DH) lines were produced from landraces Kemater Landmais Gelb (KE), Petkuser Ferdinand Rot (PE) and Lalin (LL), and characterized for their phenotypic performance (as lines per se and testcrosses), as well as for their molecular properties and population structure, representing the first comprehensive study of such large DH populations from landraces. In Hölker et al. (2022), the same landraces were crossed to the line FV2, an elite founder inbred line of the European flint heterotic pool, derived from the landrace Lacaune, adapting a scheme suggested by Stadler (1944). Doing so, populations were produced to study an admixed approach of pre-breeding in comparison with the pure approach already introduced using DH lines. The population development schemes for the two approaches are displayed graphically in Figure 2.

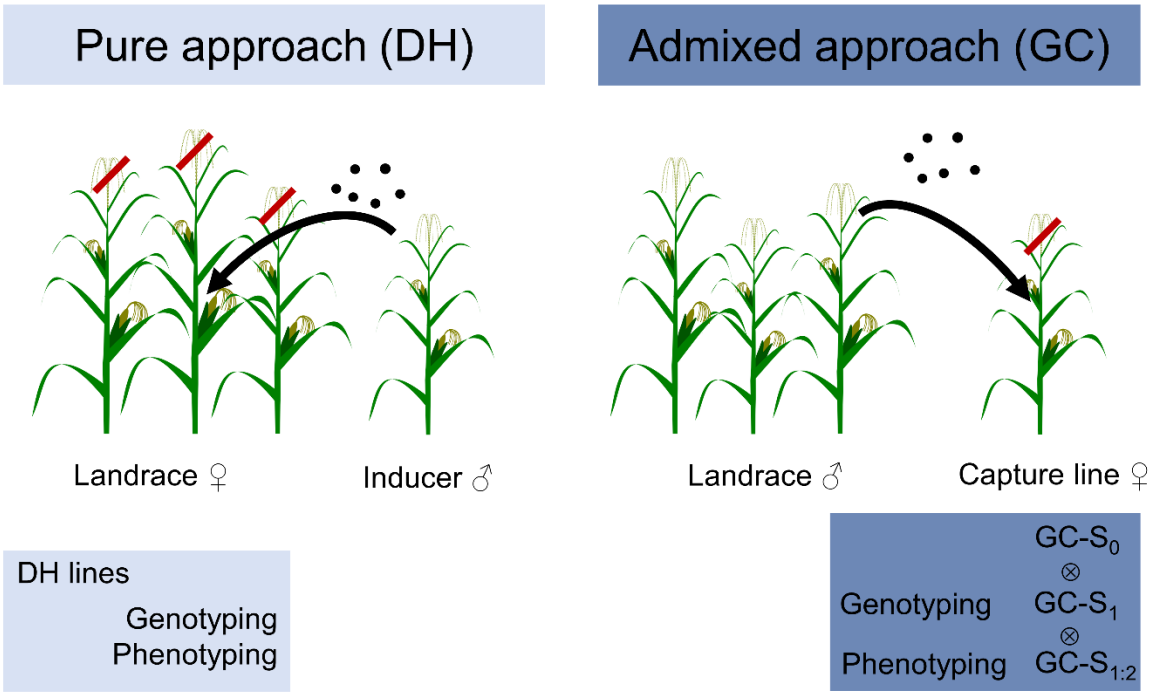


Figure 2: Population development schemes for the pure and admixed approaches. Modified after Hölker et al. (2022).

The choice of the sampling approach is of crucial importance and has wide-ranging implications for pre-breeding. Simulation studies have shown the merit of a fast reduction of the performance gap with the admixed approach, but also a high risk of reconstructing the elite parent genome in genomic selection during pre-breeding initiated from crosses of

genetic resources with elite germplasm (Gorjanc et al. 2016). In Hölker et al. (2022), both approaches were jointly investigated for their genomic prediction accuracy within as well as across populations and landraces. The method GBLUP was chosen as the statistical method for calculating genomic predictions in this thesis. The results in Hölker et al. (2022) were complemented with the development of a theoretical framework to link molecular inventories of genetic resources to phenotypic variation. The results are of crucial importance for making an informed choice of landraces and crossing partners in possible pre-breeding scenarios.

In this thesis, the findings of Hölker et al. (2019a) and Hölker et al. (2022) are combined with additional results relevant for evaluating the prospects of maize landraces for elite germplasm improvement with the use of genomic prediction in pre-breeding scenarios.

2 Material and methods

2.1 Genotypic and phenotypic data

2.1.1 Plant material

The three landraces Kemater Landmais Gelb (KE, Austria), Petkuser Ferdinand Rot (PE, Germany) and Lalin (LL, Spain) were selected from a set of 35 European maize landraces covering a broad area of Europe. The original analysis of the complete set of 35 landraces is described in Mayer et al. (2017). Criteria for selection of KE, PE and LL were that the landraces had low levels of linkage disequilibrium (LD) and population structure within populations and showed phenotypic variation for early development within landraces in preliminary field trials. Two different strategies for sampling gametes from the landraces were initiated for the selected landraces. For the first sampling strategy, the pure approach, doubled-haploid (DH) lines were directly derived from the landraces by in vivo haploid induction (Röber et al. 2005). For the second sampling strategy, the admixed approach, a modification of a scheme originally proposed by Stadler (1944) was used. For each of the landraces, 20 different and non-overlapping pollen mixtures from three landrace plants each were collected and used to pollinate the capture line FV2. FV2 is a founder line of the European flint heterotic group developed by INRAE from the French landrace Lacaune and was a very important parent line in commercial hybrids between the 1960s and 1990s. This procedure was termed gamete capture (GC). The plants resulting from this cross (GC-S_o) are all half-sibs and have one gamete from FV2 and one from the landrace. For obtaining populations for genotyping and phenotyping, the GC-S_0 plants were selfed to produce GC-S_1 ears. From each of the GC-S_1 ears, one plant was selfed again and genotyped. The field evaluation was performed on the resulting GC-S_2 lines, planted ear to row (subsequently referred to as GC-$S_{1:2}$). The populations for the pure and admixed approaches were derived from the same seed source of each landrace which is defined as the ancestral landrace (LS). An additional sample of the ancestral landrace (N = 48 per landrace) from this seed source was drawn for genotyping. In total, 1,015 DH lines were derived in the pure approach (516 DH_KE, 432 DH_PE, 67 DH_LL) and 957 GC-$S_{1:2}$ lines were derived in the admixed approach (288 GC_KE, 289 GC_PE, 380 GC_LL). Randomly chosen lines from the DH and GC populations of KE and PE (216 DH_KE, 203 DH_PE, 103 GC_KE, 54 GC_PE) were hand-crossed as pollinators onto the dent inbred line F353 (INRAE, France) for production of testcross seed.

2.1.2 Genotypic data

All material (inbred line FV2, DH and GC lines and ancestral landrace samples) was genotyped with 616,201 markers, using the 600k Affymetrix® Axiom® Maize Array (Unterseer et al. 2014). The datasets were filtered for markers assigned to the best quality class (Poly High Resolution, (Unterseer et al. 2014)), a call rate ≥ 0.9 and a known physical position on the AGPv4 (Jiao et al. 2017) B73 reference sequence. One sampled plant from the ancestral landrace sample of PE was excluded in this step due to an insufficient call rate. The remaining genotypes were subjected to following stringent quality filtering.

For the DH and GC lines used in Hölker et al. (2019a) and Hölker et al. (2022) an ancestry analysis was conducted with the software ADMIXTURE (Alexander et al. 2009) in supervised mode with four pre-defined groups (KE, PE, LL, FV2) determined from the sample of the ancestral landrace and the genotypic data of FV2. DH and GC lines with less than 75% concordance with the landrace assigned to by pedigree (or assigned landrace and FV2 in the case of GC) were excluded from further analysis. Markers and individuals with > 10% missing values were removed. Only in DH lines, markers and individuals with > 5% heterozygous calls were removed, and all remaining heterozygous calls were set to missing values. The missing values of the DH lines were imputed separately for each landrace using Beagle version 5.0 (Browning et al. 2018) and default settings. Pairwise modified Rogers' distances (MRD, (Wright 1978)) were calculated among DH lines and lines showing a MRD of < 0.05 were discarded as duplicates. For LS and GC, phasing of two gametes from each individual and imputation was also done with Beagle version 5.0, but with parameters iterations = 50, phase-segment = 10 and phase-states = 500 and the corresponding DH lines and FV2 were used as a reference panel during imputation.

After quality filtering and imputation 941 DH lines (501 DH_KE, 409 DH_PE, 31 DH_LL) and 286 gametes from the ancestral landrace (96 LS_KE, 94 LS_PE, 96 LS_LL) with 499,574 common markers remained for analysis in Hölker et al. (2019a). For the analyses in Hölker et al. (2022), LL was disregarded due to the small number of DH lines and only genotypes for which also phenotypic data was available were kept, resulting in a total of 1,512 genotypes (LS_KE = 48, LS_PE = 47, DH_KE = 471, DH_PE = 402, GC_KE = 274, GC_PE = 270), genotyped with 472,169 polymorphic single nucleotide polymorphisms (SNPs). The markers were coded as counts of the FV2 allele (0: homozygous for opposite allele of FV2, 1: heterozygous, 2: homozygous for FV2 allele).

2.1.2 Field experiments and phenotypic data analysis

Phenotyping for line per se performance of the DH lines was done in Germany at four locations in 2017 and three locations in 2018. The trials used ten separate 10 × 10 lattice designs in 2017 (1000 entries in total, of which 958 were DH lines and the rest checks), and eight separate 10 × 10 lattice designs in 2018 (800 entries in total, of which 756 were DH lines and the rest checks). In both years, a randomly chosen subset of 500 entries was also evaluated in five 10 × 10 lattice designs (500 entries, of which 458 and 468 where DH lines in 2017 and 2018, respectively, and the rest checks) at two locations in Spain. The trials were located in Einbeck (EIN, Germany, 2017 and 2018), Roggenstein (ROG, Germany, 2017 and 2018), Bernburg (BBG, Germany, 2017), Klein Wanzleben (KLW, Germany, 2018), Oberer Lindenhof (OLI, Germany 2017), Golada (GOL, Spain, 2017 and 2018) and Tomeza (TOM, Spain, 2017 and 2018). The reduction of the number of tested DH lines from 2017 to 2018 was due to seed shortages and the exclusion of lines that did not pass the above-described quality filtering of the genotypic data. The GC populations were evaluated in separate but adjacent field trials at locations EIN and ROG in 2017 and 2018. In both years, the GC trials were randomized in ten 10 x 10 lattice designs with 958 entries plus checks. The DH and GC trials were connected using common checks. In 2017, those checks were 14 flint (CH10 provided by Agroscope Changins-Wädenswil (Switzerland); D152, DK105, UH006, UH007, and UH009 provided by the University of Hohenheim (Germany); EP1 and EP44 provided by Misión Biológica de Galicia, Consejo Superior de Investigaciones Científicas, (CSIC, Spain); F03802, FV2, F283, F64, and F7 provided by Institut national de la recherche agronomique (INRAE, France); EC49A provided by Centro de Investigaciones Agrarias Mabegondo, Instituto Galego da Calidade Aumentaria (CIAM-INGACAL, Spain) and one dent line (F353, INRAE, also tester in testcross evaluations), all planted as duplicate entries. For the 2018 trials, the number of checks was reduced to four lines (DK105, EP1, FV2, F353) and each check was planted in each lattice design per location. In all trials, sampled plants from the ancestral landraces were included as quadruplicate entries. The DH and GC lines were evaluated with two replications per location. All line per se evaluation plots were single rows of 3 m length with 0.75 m distance between rows, resulting in a planting density of 8.8 plants m^{-2}.

Testcross performance was evaluated in four 10 × 10 lattice designs for the DH lines in 2018 (EIN, KLW, ROG, OLI) and 2019 (ROG and EIN); and in a generalized α-lattice design with 200 entries in ROG and EIN in 2019 for GC lines. Testcrosses of the ancestral landrace and either four (2018) or two (2019) inbred lines and either two (2018) or six (2019) commercial hybrids were included as checks in the trials and connected DH and GC trials.

Plots were double rows of 5 m or 6 m length with 0.75 m distance between rows and planting density varying between 9 and 11 plants m^{-2}. Sowing, fertilization, and plant protection in all trials (line per se and testcross) followed standard agricultural practice at the experimental stations.

In the DH line per se trials more than 25 traits were measured to cover as many different traits as possible for which the landraces showed variation. An emphasis was placed on traits related to early development. The early development related traits were early vigor (EV, at three growths stages V3, V4 and V6, 1-9 score, 1 = very poor, 9 = very vigorous), early plant height (PH, at V4 and V6 stage, cm), a cold tolerance score (CT, 1-9 score, 1 = low cold tolerance, 9 = high tolerance), and physiological traits measured in the Spanish environments only, like maximum efficiency of photosystem II (Fv/Fm, using a fluorometer (OS-30p, Opti-Sciences Inc., USA)) measured at V4 and V6, leaf greenness (SPAD, with chlorophyll content meter CCM-200, Opti-Sciences Inc., USA) measured at V3, V4 and V6. Final plant height (PH_final, cm), female flowering (FF, days after sowing until 50% of plants in a plot show silks) and male flowering (MF, days after sowing until 50% of the plants in a plot shed pollen) were measured as further important agronomic traits. At occurrence, additional traits like tillering (TILL, score 1-9, 1 = no tillers, 9 = many and long tillers), root lodging (RL, 1 = no lodging, 9 = all plants show severe lodging) and drought/heat tolerance (DT, 1-9 score, 1 = low drought/heat tolerance, 9 = high drought/heat tolerance) were measured at individual locations. At one location, also morphological traits of the tassel architecture were scored (tassel length, spike length, number of tassel branches, and tassel angle). In the testcross trials EV, PH, PH_final, FF, TILL and RL were scored as described for line per se testing. In GC trials only a small subset of traits was used for the comparative analyses and genomic prediction studies. The DH testcross trials were phenotyped for a subset of the already described traits and, in addition, testcross evaluation plots of GC and DH trials were harvested and total dry matter yield (TDMY, dt/ha) and dry matter content (DMC, %) were measured at forage harvest. See Table 1 for a detailed list of all trait and environment combinations used for this thesis.

Table 1: Table specifying which traits were measured in which environments for line per se (LP) and testcross (TC) trials. Traits are emergence (EME), early vigor and early plant height at stages V3, V4, and V6 (EV_V3, EV_V4, EV_V6, PH_V3, PH_V4, PH_V6), ear height (EH), final plant height (PH_final), male flowering (MF), female flowering (FF), anthesis-silking-interval (ASI), root lodging (RL), tillering (TILL), cold tolerance (CT), drought/heat tolerance (DT), tassel length (TL), spike length (SL), number of tassel branches (NB), tassel angle (TA), maximum photosystem II efficiency at stages V4 and V6 (Fv/Fm_V4, Fv/Fm_V6), leaf greenness at stages V3, V4, and V6 (SPAD_V3, SPAD_V4, SPAD_V6), dry matter content (DMC), and total dry matter yield (TDMY).

Trait	Environments per se performance DH	Environments per se performance GC	Environments testcross performance DH	Environments testcross performance GC
EME	All DH*	-	All 2018 DH environments	-
EV_V3	All DH*, except EIN 2018	-	KLW 2018, OLI 2018, ROG 2018	-
EV_V4	All DH*	-	All 2018 DH environments	-
EV_V6	All DH*	-	All 2018 DH environments	-
PH_V3	GOL and TOM 2017 and 2018	-	KLW 2018 and OLI 2018	-
PH_V4	All DH*	-	All 2018 DH environments	-
PH_V6	All DH*	EIN and ROG 2017 and 2018	All DH*	EIN and ROG 2019
EH	ROG 2017 and 2018	-	ROG 2018	
PH_final	All DH*	EIN and ROG 2017 and 2018	All DH*	EIN and ROG 2019
MF	EIN and TOM 2017 and 2018, GOL 2018	-	-	-
FF	All DH* except GOL 2017	EIN and ROG 2017 and 2018	All DH*, except KLW 2018	EIN and ROG 2019
ASI	EIN and TOM 2017 and 2018, GOL 2018	-	-	-
RL	BBG 2017, EIN 2017 and 2018,	-	EIN 2018, OLI 2018, ROG 2018	-

TILL	OLI 2017, ROG 2017 and 2018 EIN and ROG 2017 and 2018, KLW 2018	-	OLI 2018	-
CT	OLI 2017	-	-	-
DT	EIN 2018	-	-	-
TL	ROG 2018	-	-	-
SL	ROG 2018	-	-	-
NB	ROG 2018	-	-	-
TA	ROG 2018	-	-	-
Fv/Fm_V4	GOL and TOM; 2017 and 2018	-	-	-
Fv/Fm_V6	GOL and TOM 2017	-	-	-
SPAD_V3	GOL and TOM; 2017 and 2018	-	-	-
SPAD_V4	GOL and TOM; 2017 and 2018	-	-	-
SPAD_V6	GOL and TOM 2017	-	-	-
DMC	-	-	All DH*	EIN and ROG 2019
TDMY	-	-	All DH*	EIN and ROG 2019

* All DH environments for line per se performance are Einbeck (EIN, Germany, 2017 and 2018), Roggenstein (ROG, Germany, 2017 and 2018), Bernburg (BBG, Germany, 2017), Klein Wanzleben (KLW, Germany, 2018), Oberer Lindenhof (OLI, Germany 2017), Golada (GOL, Spain, 2017 and 2018) and Tomeza (TOM, Spain, 2017 and 2018); for testcross performance EIN, KLW, ROG, OLI in 2018 and ROG and EIN in 2019

The analysis of phenotypic data for Hölker et al. (2019a) was based on the following model:

$$y_{ijkopst} = \mu + m_i + \delta_{Checks} l_j + g_{k(ij)} + u_o + \delta_{Checks} lu_{jo} + gu_{ko(ij)} + k_{p(o)} + r_{s(op)} + b_{t(ops)} + \varepsilon_{ijkopst} \quad (1)$$

$y_{ijkopst}$ trait observation

μ overall mean

m_i effect of group *i*, with *i* = 1, 2, 3 (DH, LS, checks) for line per se and *i* = 1, 2 (DH, checks) for testcrosses

δ_{Checks} dummy variable with δ_{Checks} = 1 if the line belongs to DH and 0 otherwise

l_j effect of landrace *j* in group *i* = 1, with *j* = 1, 2, 3 (DH_KE, DH_PE, DH_LL) for line per se and *j* = 1, 2 (DH_KE, DH_PE) for testcrosses

$g_{k(ij)}$ effect of genotype *k* nested in group *i* and landrace *j*

u_o effect of environment *o*

lu_{jo} interaction effect of landrace *j* and environment *o*

$gu_{ko(ij)}$ interaction effect of genotype *k* and environment *o* nested in group *i* and landrace *j*

$k_{p(o)}$ effect of the lattice *p* nested in environment *o*

$r_{s(op)}$ effect of replicate *s* nested in lattice *p* and environment *o*

$b_{t(ops)}$ effect of block *t* nested in replicate *s*, lattice *p* and environment *o*

$\varepsilon_{ijkopst}$ residual error

All effects except m_i and l_j were treated as random for estimating genotype and genotype × environment variance components. Variance components for $g_{k(ij)}$ and $gu_{ko(ij)}$ were modeled individually for the three landraces, assuming that DH lines were unrelated. Residuals were assumed to be normally distributed with mean zero and different variances for DH lines (δ_{Checks} = 1) and checks/LS (δ_{Checks} = 0), but equal residual variances were assumed for all landraces in all environments. Raw data and outlier curation was done manually by inspection of residual plots. Measurements of lines that did not pass quality filtering of genotypic data but were still evaluated in the field in 2017 were set to missing values in the data analysis. Variance components and their standard errors were estimated using restricted maximum-likelihood estimation implemented in the ASReml-R package (Butler et al. 2009). Differences among means l_j were tested with pairwise *t*-tests with the R-package asremPlus. Heritabilities were calculated separately for each landrace on an entry-mean basis following Hallauer et al. (2010):

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{gu}^2}{n_u} + \frac{\sigma_\varepsilon^2}{n_u n_r}} \quad (2)$$

h^2 entry-mean heritability

σ_g^2 genotypic variance

σ_{gu}^2 genotype × environment variance

σ_ε^2 residual variance

n_u number of environments

n_r number of replications

Standard errors of heritability estimates were derived from standard errors of corresponding variance components with the delta method (Holland et al. 2010). Variance components and heritabilities exceeding twice their standard error were considered significant. Best linear unbiased estimates (BLUEs) of the genotype mean for each trait and DH line were obtained from a simplified model in Eq. (1) after dropping factors m_i, $\delta_{Checks} l_j$, and $\delta_{Checks} lu_{jo}$, and treating genotype as a fixed effect. The same model was also used for forming linear contrasts to test for significant differences (*t*-tests) between LS and the mean of the corresponding DH population (both for line per se and testcross performance) and between the mean of the commercial check hybrids and the mean of the DH population (only for testcross evaluation). The predicted response from selection within DH populations (line per se and testcrosses) was calculated according to Falconer and Mackay (1996) for a selection intensity of $\alpha = 10\%$ ($i_{(10\%)} \approx 1.76$):

$$\Delta G_{(\alpha)} = i_{(\alpha)} h \sigma_g \quad (3)$$

$\Delta G_{(\alpha)}$	predicted response from selection
$i_{(\alpha)}$	selection intensity
h	square root of the heritability
σ_g	genetic standard deviation

To account for mean differences and different selection responses, the usefulness criterion (Schnell 1983) was calculated as $U_{(10\%)} = \bar{x} \pm \Delta G_{(10\%)}$ where $\bar{x}$ = mean of the respective DH population. The BLUEs were used to calculate phenotypic correlations among traits as Pearson correlation coefficients within populations and in lines per se and testcrosses. For evaluating the prospects of selection on line per se performance, Spearman rank correlations were calculated across line per se and testcross performance for the same traits. Multiple-testing correction was done for all phenotypic correlations within DH populations by Bonferroni-Holm correction (Holm 1979). The model from Eq. (1) was expanded to a bivariate model with pairs of traits (or the same trait in lines per se and testcrosses) for estimating genetic correlations. The genetic correlations were considered significant if they exceeded twice their standard error.

This analysis of only DH lines was expanded for Hölker et al. (2022) to accommodate a joint analysis of the DH and GC experiments using the following model:

$$y_{ijkopst} = \mu + m_i + \delta_{Checks} l_j + g_{k(ij)} + u_o + gu_{ko(ij)} + \delta_{DH}\{lu_{jo} + k_{p(o)} + r_{s(op)} + b_{t(ops)}\} + \delta_{GC}\{lu_{jo} + k_{p(o)} + r_{s(op)} + b_{t(ops)}\} + \varepsilon_{ijkopst} \quad (4)$$

$y_{ijkopst}$	trait observation
μ	overall mean
m_i	effect of group *i*, with *i* = 1, 2, 3, 4 (GC, DH, LS, checks) for line per se and *i* = 1, 2, 3 (GC, DH, checks) for testcrosses
δ_{Checks}	dummy variable with δ_{Checks} = 1 if the line belongs to DH or GC and 0 otherwise
l_j	effect of landrace *j* in group *i* = 1, with *j* = 1, 2, 3, 4 (GC_KE, GC_PE, DH_KE, DH_PE) for lines per se and for testcrosses
$g_{k(ij)}$	effect of genotype *k* nested in group *i* and landrace *j*
u_o	effect of environment *o*
δ_{DH}	dummy variable with δ_{DH} = 1 if data belongs to the DH experiment and 0 otherwise
δ_{GC}	dummy variable with δ_{GC} = 1 if data belongs to the GC experiment and 0 otherwise
lu_{jo}	interaction effect for landrace *j* and environment *o*
$gu_{ko(ij)}$	interaction effect for genotype *k* and environment *o*
$k_{p(o)}$	effect of the lattice *p* nested in environment *o*
$r_{s(op)}$	effect of replicate *s* nested in lattice *p* and environment *o*
$b_{t(ops)}$	effect of block *t* nested in replicate *s*, lattice *p* and environment *o*
$\varepsilon_{ijkopst}$	residual error

The remaining analysis was done the same way as described above, e.g. same random and fixed effects, heritability calculation, calculation of BLUEs with the above-described simplified model that was also used to obtain tests for significant differences (*t*-tests) between LS, DH, GC and FV2 in linear contrasts, and estimation of genetic covariances and genetic correlations between line per se and testcross performance for a given trait in a bivariate extension of Eq. (4). In this case, significance of genetic covariances was tested in likelihood-ratio-tests comparing the model including the covariance with the reduced model without the covariance.

2.2 Genetic data analysis

2.2.1 Population structure and linkage disequilibrium

With DH and LS, a principial coordinate analysis (Gower 1966) was carried out with the R-package ape. The modified Rogers' distance (MRD, Wright 1978), a scaled Euclidean distance measure ranging between 0 and 1 was calculated for pairs of individual genotypes:

$$MRD = \frac{1}{\sqrt{2m}}\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{I_i}(p_{ij} - q_{ij})^2} \quad (5)$$

m	number of markers
I_i	number of alleles at marker *i* (for biallelic SNPs, $I_i = 2$)
p_{ij}, q_{ij}	allele frequencies of the *j*-th allele at the *i*-th locus in the two individuals under consideration ($p_{ij}, q_{ij} \epsilon \{0, 0.5, 1\}$)

The MRD matrices were hierarchically clustered using the unweighted pair group method with arithmetic mean (UPGMA) implemented in the hclust function in R. The results were displayed as 1-MRD. Linkage disequilibrium (LD) was measured using r^2 (Hill and Robertson 1968) for samples of 94 gametes within each population of landraces KE and PE. The value for r^2 was calculated for pairs of SNPs within a distance of 1 Mb. The r^2 decay with physical distance was investigated with nonlinear regression (Hill and Weir 1988). The physical distance δ for which the curve reaches r^2 = 0.2 is defined as the LD decay distance. LD across chromosomes was estimated by sampling 5,000 markers per chromosome with replacement for all 45 pairwise combinations of chromosomes and calculating r^2 for all pairs of markers across chromosomes. Following Schopp et al. (2017), linkage phase similarities (LPS) between populations were calculated and LPS according to physical distance was calculated by grouping pairs of markers into bins of between 10 kb and 1 Mb.

2.2.2 Molecular variance and genetic diversity

Different types of analysis of molecular variance (AMOVA) were carried out, all based on Euclidean distances (Excoffier et al. 1992). For Hölker et al. (2019a), the proportion of molecular variance explained by the three landraces under study was estimated by partitioning the molecular variation into within- and between-landrace components, using the DH and the original panel of 35 European landraces from Mayer et al. (2017) for comparison. A second AMOVA was carried out with DH and LS to decompose variation

within and between DH lines and LS gametes to investigate how much of the molecular variance lies within and between those groups. For Hölker et al. (2022), an AMOVA was done to estimate molecular variance within and between individuals of LS, DH, and GC separately for landraces KE and PE.

Genetic diversity was estimated by sampling 80 gametes from each population (LS, DH, and GC) and landraces KE and PE with 500 replicates for comparing the number and percentage of polymorphic markers across populations. The allele frequencies between DH and LS and between the experimental and expected GC were compared for KE and PE. The expected GC allele frequency was obtained by $(p+1)/2$, with *p* being the frequency of the FV2 allele in the respective LS. The simple matching coefficient (SM) was calculated across all SNP loci as described by Jacobson et al. (2015) and the genetic distance (GD) between two genotypes was measured as $GD = 1 - SM$.

2.3 Genome-based prediction

2.3.1 Genomic prediction model

Genome-based prediction in Hölker et al. (2022) for all scenarios and per se and testcross performance was done with genomic best linear unbiased prediction (GBLUP), always applying the following model

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Zu} + \mathbf{e} \quad (6)$$

$\mathbf{y}$	vector of BLUEs of the training set, from phenotypic analysis
$\mathbf{1}$	vector of 1s
μ	population mean
$\mathbf{Z}$	incidence matrix for $\mathbf{u}$
$\mathbf{u}$	vector of random genetic values with distribution $\mathbf{u} \sim N(0, \mathbf{U}\sigma_g^2)$; where $\mathbf{U}$ is the realized relationship matrix calculated after method 1 of VanRaden (2008) considering all genotypes (GC and DH of KE and PE) and σ_g^2 is the genetic variance pertaining to the GBLUP model
$\mathbf{e}$	vector of residuals, with distribution $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, where $\mathbf{I}$ is the identity matrix and σ_e^2 denotes the residual variance pertaining to the GBLUP model

Relationship matrices were calculated using R version 3.6.0 (R Core Team 2019) and R-package synbreed in version 0.12-9 (Wimmer et al. 2012). Estimation of the variance components pertaining to the GBLUP model was done using ASReml-R in version 3.0

(Butler et al. 2009). The genomic variance estimate was tested for significance with a likelihood-ratio-test by comparing the models with and without the model term. Predictive ability was determined by calculating the Pearson correlation between the predicted values with the observed phenotypic values in the prediction set. The results were reported as genomic prediction accuracy (ρ), which is the correlation between predicted and unobservable true genetic values. For estimating ρ, the calculated predictive abilities were divided by the square root of the heritability of the prediction set (Dekkers 2007).

2.3.2 Scenarios for genomic prediction

For obtaining a comprehensive evaluation of prediction accuracies in landrace material, several scenarios were considered for genomic prediction, which are depicted in detail in Figure 3.

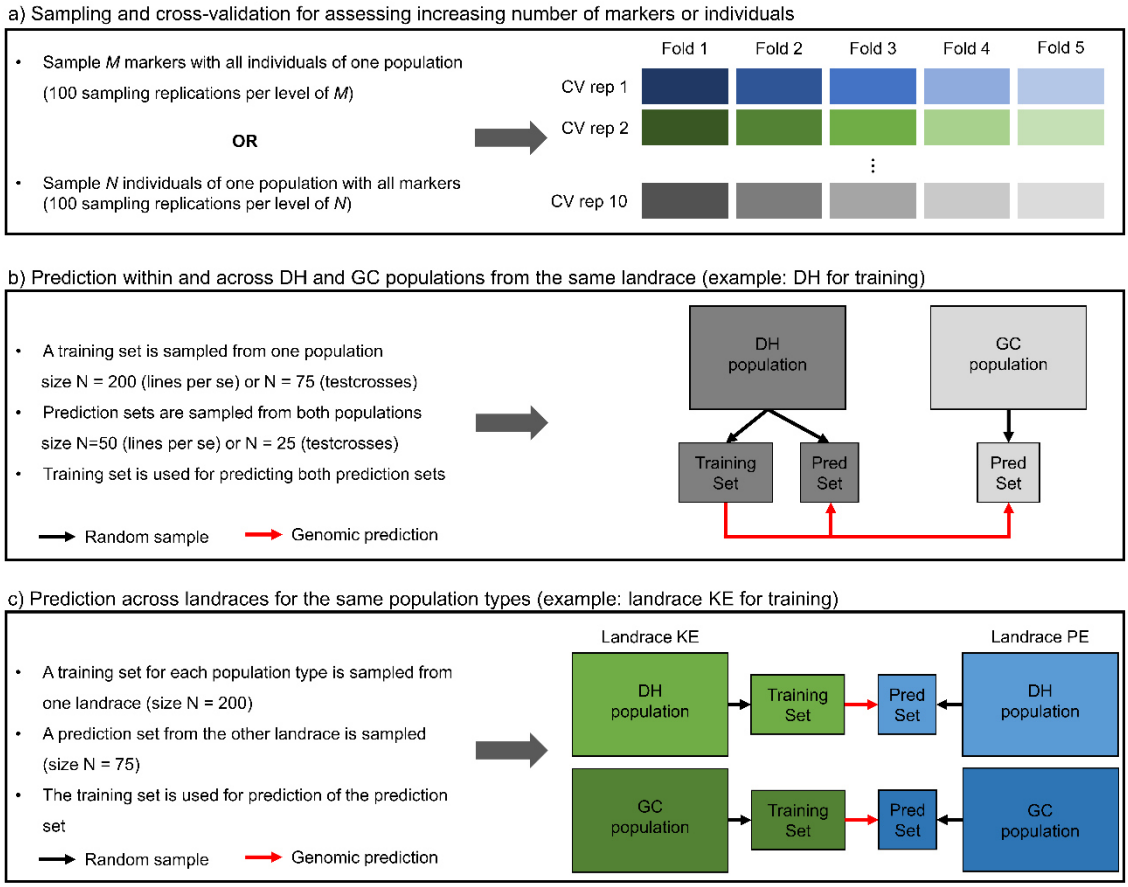


Figure 3: Genomic prediction scenarios in this thesis. For sampling the number of markers *M* and number of individuals *N*, 100 sampling replications were drawn for each level of *M* and *N*, respectively. On each of those samples, ten times fivefold cross-validation (CV) was applied (a). In each of the CV reps, the sample was split in five folds ten times and in each CV rep always four folds are used for predicting the remaining fold. The prediction and training folds are rotated until

all folds are predicted once. For predictions across and within DH and GC populations (b), one population (either DH or GC, in the example it is DH) is used for sampling both, a training and a prediction set. In addition, a prediction set is also sampled from the other population of the same landrace. The training set is used to predict both prediction sets (b). For predictions across landraces (c), again a training set for each population type and landrace was sampled and used to predict the same population type of the other landrace. For b) and c) 100 samples were drawn.

The scenarios can be divided into studying the influence of technical parameters like number of markers and number of individuals in the training set on the prediction accuracy on the one hand and in investigating the prediction accuracy in prediction scenarios within and across populations and landraces on the other hand. For studying the influence of number of markers *M* on prediction accuracy, 1,000, 2,000, 5,000, 10,000, 15,000, 20,000, 50,000, 100,000, 150,000, 200,000, and 250,000 markers were sampled. Sample size *N* was increased from 50 to the maximum possible number for the respective population in increments of 50 (except for the last increment up to the maximum). The sampling of *M* and *N* without replacement was repeated 100 times (except for the maximum value of *N*) yielding 100 sampling replications. Within each sample, 10 times fivefold cross-validation (CV) was carried out with line per se data. Prediction accuracies ρ were averaged across the 50 CV replications and the 95% quantile was calculated from the 100 sampling replications. With small sample sizes *N*, the mixed model algorithm did not converge for some training sets. In those cases, ρ was set to missing value.

For comparing ρ within and between the DH and GC populations, either 200 (per se performance) or 75 (testcross performance) genotypes were randomly sampled from one of the two population types (either DH or GC) as the training set. The prediction set always comprised a disjoint set of 50 (per se performance) or 25 (testcross performance) genotypes either from the same or a different population. Sampling without replacement was repeated 100 times. The DH and GC populations of KE and PE were used as training population with line per se data, with testcross data only GC_KE was analyzed and GC_PE was excluded as training population due to its small sample size. The same sampling procedure was applied for across landrace predictions, but in this case only predictions with line per se data were studied.

3 Discussion

Continuous genetic improvement of crops is necessary to secure agricultural productivity and meet the challenges current production systems are facing. Increased awareness of the environmental impact of management practices and the changing climate will be a key requirement in future breeding programs. However, in elite germplasm relevant genetic variation for genetic improvement towards stress resilience and resource use efficiency in crops is limited. In this thesis, the objective was to develop whole-genome based strategies for the utilization of genetic resources, such as landraces, for elite germplasm improvement for quantitative traits, using European flint maize landraces as an example. In the following chapter, the results from the two publications included in this thesis (Hölker et al. 2019a, 2022) are discussed and complemented by additional results.

3.1 Choice of landraces and sampling method

3.1.1 Choice of source material

The global germplasm of maize genetic resources, currently stored in gene banks across the world, contains extensive genetic variation (Lu et al. 2009), resulting from a long history of open pollination, mutation, drift, and selection over a broad range of environments (Mercer and Perales 2010; Hufford et al. 2012). Due to local adaptation to specific environments, it is likely that landraces contain favorable alleles that could prove useful for the improvement of current elite germplasm (Hellin et al. 2014; Mayer et al. 2020). However, the total number of accessions stored in gene banks is too large to be able to study all of them in detail (Hoisington et al. 1999), thus a preselection of the source material is a prerequisite for working with this source of diversity. The first question therefore is how to decide which landrace or landraces should be evaluated in more detail. For an informed decision regarding how to sample (many accessions with few individuals vs. many individuals from few accessions), knowledge concerning the distribution of genetic diversity within and across landraces is important. The molecular diversity of landraces of allogamous crops has been studied in many ways, but a common result was that the majority of molecular variation can be found within instead of across landraces (Sood et al. 2014). For European flint maize, chosen as an example for this thesis, Mayer et al. (2017) showed that sampling from five of the 35 studied landraces was, on average, enough to capture 95% of the total molecular variation of the whole set. Therefore, sampling of many individuals of fewer landraces seems a suitable strategy for the improvement of elite germplasm for quantitative traits. On the other hand, sampling few individuals from many

landraces has been shown to be appropriate for studying broad signals of adaptation, crop evolution or the effects of rare alleles (Heerwaarden et al. 2011; Navarro et al. 2017; Kremling et al. 2018). Such a diverse set of landraces is expected to harbor a wide range of adaptive alleles, which might hamper the landraces' utilization in a specific target population of environments. Thus, the use of more diverse material (compared to material already adapted to the target region) offers more novel genetic variation, but comes at a much higher investment for the successful incorporation into elite germplasm of a given target environment (Sood et al. 2014).

In the Latin American Maize Project (LAMP, late 1980s and 1990s), 12 countries of North and South America evaluated their native accessions of genetic resources together. In total, 12,000 accessions were phenotypically evaluated (per se and in elite testcrosses) in a wide range of environments and several testing stages (Pollak 2003). The best LAMP accessions were identified, and pre-breeding was initiated during the follow-up project Germplasm Enhancement of Maize (GEM). These efforts were based on phenotypic selection only and consumed a lot of field-testing capacities and time. More recent studies indicated that genomic prediction can be employed also across a broad range of diversity (Crossa et al. 2016; Yu et al. 2016) and could therefore be used to identify genetic resources that are promising based on their predicted performance. Allier et al. (2020b) developed a strategy for a targeted identification of donors from a panel of genetic resource inbred lines (important founder lines, lines derived from landraces, elite material in the public domain and breeding material from public institutes) that have been characterized phenotypically and genotyped with molecular markers. The donors were selected to complement the elite germplasm, balancing their own genetic value as an indicator of short-term genetic gain, and their allelic originality at quantitative trait loci (QTLs), influencing the expectation for long-term genetic gain (Allier et al. 2020b). However, this concept has not been proven in practice. Other studies in maize evaluating the use of landrace material for elite germplasm improvement sampled only one or very few individuals per landrace (Takuno et al. 2015; Navarro et al. 2017). Only few reports sampled larger numbers of progenies per landrace with 21 to 113 DH lines per landrace from up to six different landraces (Böhm et al. 2017; Melchinger et al. 2017; Brauner et al. 2019; Würschum et al. 2022).

As a conclusion, sampling from landrace source material should be guided by the respective research aim. When investigating genome-based prediction in landrace-derived material for improving elite germplasm, as in this thesis, the strategy of sampling intensively

from fewer landraces that are already adapted to the target environment should be the method of choice. Even after making general decisions on the strategy and reducing the plethora of possible choices of source material to the ones originating from the same geographic region as the breeding target (e.g. European flint for the improvement of flint elite material, as for this thesis), the number of possible accessions is still prohibitively large. The main options for further material selection are the use of small-scale geographic origin, molecular data, and/or predicted or observed phenotypic performance.

During the preselection carried out on the material for this thesis, all of these information sources were used: geographic origin, phenotypic performance and variability for the traits of interest as well as molecular data. The search was restricted to material adapted to the target region so that only European flint landraces were considered. A phenotypic pre-screening for variation in early development within landraces was done and high molecular diversity within the chosen landraces was ensured. The landraces KE, PE, and LL selected for this study together represent 95% of the molecular variance of the full set of 35 landraces analyzed by Mayer et al. (2017). Thus, the applied preselection resulted in a higher representation of the full allelic spectrum in these three landraces than expected from random sampling (Hölker et al. 2019a). The three landraces constitute a promising starting point for improving European flint maize elite breeding pools, as the selected landraces cover much of the relevant diversity, show no indication of pronounced hidden population structure (Mayer et al. 2017) and the landraces are expected to be adapted to the Central European target region (Mayer et al. 2020).

The next crucial step for initiating a pre-breeding effort is to develop populations of reproducible genetic units that capture the genetic variance of the source material in the best possible way for further breeding activities.

3.1.2 Population development in pre-breeding

Reproducible genetic units can be obtained in different ways that come with different advantages and disadvantages. The most basic categorization of strategies is into “pure” approaches, keeping landrace material without elite introgression, or “admixed” approaches, where crosses between landrace and elite material are made before the evaluation process. Eventually, all pre-breeding efforts will lead to crosses between landrace and elite germplasm, but the timing of these crosses has large impact, e.g. on the opportunities for selection along the way and the final output. The main criteria for comparing different strategies of population development are the required time and resources, the expected performance level of the resulting population and the unbiased

representation of the original source by the developed population. The general expectation for all landrace material is that a substantial gap of performance to current elite germplasm exists and needs to be bridged (Sood et al. 2014; Böhm et al. 2017).

When landraces are evaluated choosing the “pure” approach, reproducible genetic units can be obtained either by recurrent selfing or by direct production of DH lines. Recurrent selfing requires only a low to medium effort per generation, but several subsequent generations (> 5) need to be produced before sufficient inbreeding levels are reached for the material to be stable and reproducible. Selection can be applied during the inbreeding process, which would theoretically increase the performance level, but only for highly heritable traits that can be evaluated on the lines per se. The recurrent inbreeding additionally harbors the risk of losing many lines in late generations of inbreeding, as the percentage of homozygous loci increases, which also increases the risk of unmasking deleterious alleles in a homozygous state.

The production of DH lines yields fully inbred lines much faster and promises to purge the population of deleterious alleles immediately instead of after several generations of selfing, which should generally result in a performance improvement compared to the original landrace (Melchinger et al. 2017). However, the initial effort for producing DHs is often high due to the low efficiency of DH production in genetic resource material (Melchinger et al. 2017). Some landraces carry such a high genetic load of deleterious alleles that efforts for production of DH lines are prohibitive for these landraces. The low efficiency in DH production for LL observed in this thesis compared to the other two landraces may indicate a stronger load of deleterious alleles in LL, which could possibly also result in a stronger effect on phenotypic performance.

All pure methods of population development have in common that the performance level of the resulting populations is mainly determined by the choice of the source material, as the possibilities for reducing the performance gap to elite material during population development (selection during selfing, purging of deleterious alleles in DH process) are minor compared to the effect a direct cross with more elite material has. The pure populations are expected to result in mostly unbiased representations of the source landraces, as no introgression of other germplasm happens and the applied selection, if any, cannot be focused on the most important traits, like yield. Any deviation in allelic frequencies or phenotypic performance from the original source population would originate from purging or, in the case of DH production, from unintentional selection through the DH production procedure. A previous study by Melchinger et al. (2017) showed that the

production of DH lines from maize landraces results in populations that represent the ancestral landrace in an unbiased way, but the study indicated strongly varying efficiency in the DH production across different source landraces, which is most likely a result of the substantial number of deleterious alleles the landraces harbor (Yang et al. 2017) that are uncovered instantaneously during DH production. An analysis of the same data focusing on those deleterious, often rare (Mezmouk and Ross-Ibarra 2014), alleles by Zeitler et al. (2020) indicated that the DH production indeed resulted in limited loss of diversity in some regions of the genome. In summary, the pure approach in general, and the DH approach in particular, is suitable for successful sampling and production of reproducible genetic units from landraces for subsequent pre-breeding efforts. The performance gap to elite germplasm will more or less persist during population development with a pure approach.

The admixed approaches have in common that the source material is crossed to a crossing partner, which could, for example, be an inbred line. A process like this was termed gamete selection by Stadler et al. (1944). The initial crossing can be considered low effort, but the resulting F1-hybrids are not yet reproducible genetic units. They need at least a few generations of selfing or could also be followed by an additional DH production step to obtain fully homozygous inbred lines. Independent of the exact procedure, the performance, molecular variation and representation of the source material by the resulting population is strongly influenced by the properties of the crossing partner and the selection pressure applied during population development. The higher the crossing partner's performance level, the smaller will be the overall performance gap of the resulting population to elite germplasm. However, simulation studies indicate that a strong crossing partner and immediate selection carry the risk of reconstructing the elite crossing partner's genome in subsequent selection cycles (Gorjanc et al. 2016). A crossing partner of lower performance level is expected to result in less selection pressure against the source material but will on the other hand result in a larger remaining performance gap (Hölker et al. 2022). In Hölker et al. (2022) theoretical derivations were developed to derive expectations for the molecular and genetic variation in the admixed approach.

For this thesis the goal was to empirically compare the pure and the admixed approach. The aim was to show how molecular variances of the resulting populations can be predicted and to investigate if the molecular variance also predicts genetic variation measured in multi-environment field trials. Therefore, the three preselected landraces were intensively sampled by direct DH production from the landraces (pure approach) and by crossing to a capture line (admixed approach, following a modification of Stadler (1944)).

The required efforts for DH production confirmed the varying efficiencies in DH production observed by Melchinger et al. (2017). The landrace LL resulted in a much smaller number of DH lines compared to the other two landraces, although the input during DH production for LL was even larger than for the two other landraces. The analysis of population structure and molecular variation showed that a low level of population structure was already present in all of the ancestral landrace populations, but no further structuring was introduced through the DH production process (Hölker et al. 2019a). On the phenotypic level, the direct comparison of performance of the ancestral landrace and the derived DH population is not possible due to the different levels of inbreeding in the two populations. However, a comparison can be made on the level of testcrosses, as, without directed selection, the mean testcross performance of the DH population is expected to be equal to the testcross performance of the ancestral landrace. As shown in Hölker et al. (2019a) and Hölker et al. (2022), those differences were almost always nonsignificant for the DH populations. The result shows that no inadvertent selection pressure was applied. For the DH, it indicates that the possible loss of diversity during DH production indicated by Zeitler et al. (2020) did not affect the phenotypic performance, at least for the landraces KE and PE. All landraces were sampled from gene bank accessions and their history before storage in the gene bank and also the propagation history since storage is not fully documented. The landraces KE and PE might have gone through stronger bottlenecks than LL before storage in a gene bank or inadvertently during propagation cycles, increasing the efficiency of DH production by a reduction of genetic load (Hölker et al. 2019a). However, no testcrosses for LL were produced for this research so it remains an open question if the low production efficiency is an indication of selection pressure that changed the phenotypic performance in this case.

The generic theoretical framework developed as part of this thesis allows to predict the molecular variances in the GC populations and highlights the large influence of the choice of the capture line (Hölker et al. 2022). Depending on which alleles are present in the capture line and landrace, the genetic variance of the resulting population can be larger, equal or smaller than in a corresponding pure population of DH lines. In the GC populations evaluated in this study the proportion of the capture line genome was on average about 50%, thus meeting expectations. In addition, a significant correlation between per se performance and capture line genome proportion (ranging between 21.5 and 75.6% among the GC lines) could not be observed except for flowering time in testcrosses of KE. Although significant, the correlation was still only weak (r = -0.25).

The genetic makeup and performance of the capture line can have a large influence on the derived populations (mean, genetic variance, risk of masking (favorable) landrace alleles). In the case of the studied populations, the difference between phenotypic per se performance of the capture line and the mean of the ancestral landrace provided an indication for which traits the capture line might carry alleles absent in the landrace. The capture line used in this study is an important founder line of Flint elite germplasm. It needs to be shown that this pattern also persists for other combinations of landraces and capture lines, especially when the capture line is a very recent high performing elite line. Locus-specific allele frequencies can be predicted from the molecular data. Hölker et al. (2022) showed that those parameters do not provide sufficient guidance for the choice of the capture line in a quantitative trait setting, where many loci contribute to the genetic variance. In summary, the DH and GC approach both capture the diversity present in the landrace.

3.1.3 Phenotypic performance of landrace material

The possibility to capture molecular variation in populations derived from landraces has been described intensively for qualitative traits like disease resistance. In this thesis, the hypothesis that the molecular variance of the landraces translates to genetic variation in quantitative traits could be validated for both, DH and GC populations (Hölker et al. 2019a, 2022). The landrace-derived populations were evaluated for line per se and testcross performance and revealed large genetic variances within landraces but only small phenotypic differences among landrace means. Due to the availability of data for only three landraces, a full decomposition of genetic variance into within and across landrace variance was not meaningful, but the results are in agreement with a previous study of European flint maize landraces that sampled more landraces with less individuals and decomposed genetic variances with the result of the majority of variation being found within landraces (Böhm et al. 2017). However, the large number of over 25 traits studied in this thesis revealed that the variation of individual traits must still be evaluated specifically as genetic variance differed substantially among traits and landraces (Hölker et al. 2019a).

Several studies point towards a yield performance gap between landrace and elite material of between 15 and 40%. These results are in agreement with the 15% yield gap found in Hölker et al. (2019a) for the DH populations in comparison with a small number of elite check hybrids evaluated together in 2018. In Hölker et al. (2022) another year of testcross evaluation with more commercial check hybrids and the joint evaluation of DH and GC populations confirmed a yield gap between 13% (DH_KE) and 16% (GC_PE). For

appropriate assessment of the yield differences, it has to be considered that the tester used in this study (F353) was released in 2001 and therefore is expected to have better performance than genetic resources but also a yield gap in comparison to recent commercial check hybrids that are highly selected on general and specific combining ability. A small set of DH lines from landraces that have been preselected, but not further bred or recombined, and tested with a modern tester in a public breeding program in direct comparison to current elite material of the same breeding program showed that some individual lines reached competitive yield (Hölker et al. 2019b). However, on average the DH lines showed a substantial yield gap of about 17% compared to the recent elite lines (Hölker et al. 2019b). For this thesis, six of the best performing (total dry matter yield) landrace DH lines from PE were used for testcross seed production with a modern elite dent tester. The testcrosses were evaluated in trials together with new flint elite breeding lines crossed to the same tester by the breeding company KWS SAAT SE & Co. KGaA. The results for total dry matter content (DMC) and total dry matter yield (TDMY) are shown in Figure 4. The six preselected DH lines yielded on average about 9% less than the mean of all elite entries. The best performing DH_PE line yielded about 5% less than the elite entries. The performance gap to the best 5% of elite entries amounts to about 17% on average and 13% for the best performing DH_PE line, respectively.

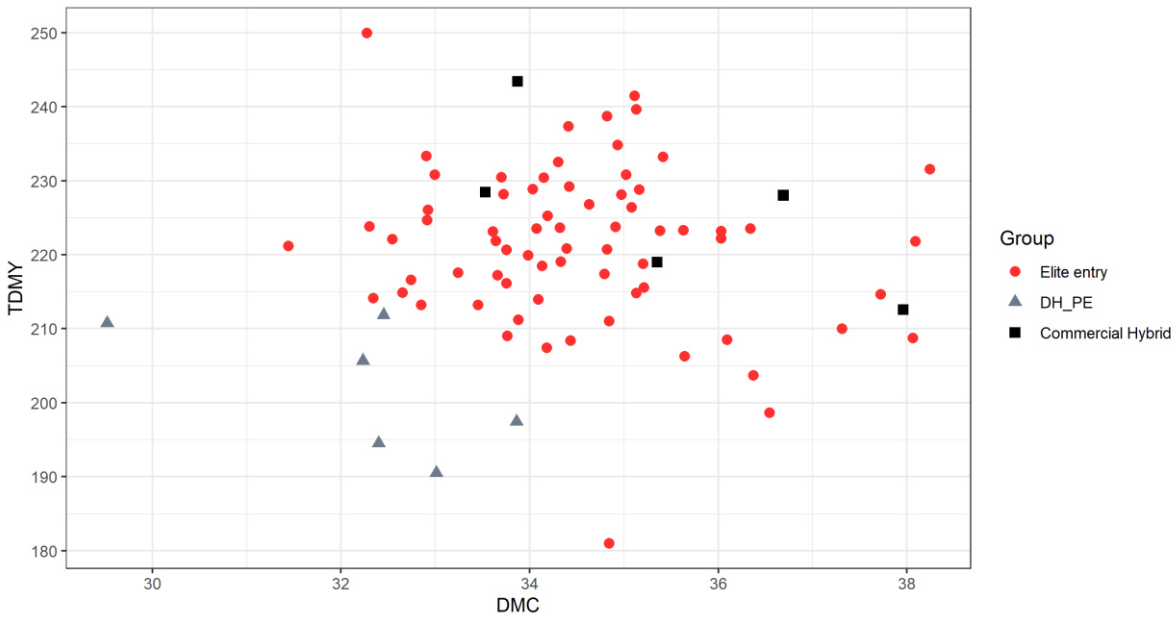


Figure 4: Scatter plot of total dry matter content (DMC) and total dry matter yield (TDMY) at silage maturity for testcrosses of elite inbred lines from a commercial breeding program (Elite entry), six preselected DH lines from landrace Petkuser Ferdinand Rot (DH_PE) crossed to the same elite tester, and five commercial hybrids. Presented results are adjusted means across up to six locations in Germany in 2021.

Even more interesting than the yield gap is the difference in total dry matter content. The landrace PE is expected to be early in maturity, with its origin in the northeast of Germany (Brandenburg). As previously established, the DH process did not introduce a selection. The total dry matter content of the six DHs selected for elite testcrosses was not lower than the mean of the total DH_PE population in previous trials with F353 as a tester. Thus, the difference in total dry matter content was also not introduced by the selection after the DH process. Nevertheless, the DH_PE lines have, on average, 2.3% lower total dry matter content than the mean of the elite entries. This indicates that either the dry-down characteristics of the landrace PE differs strongly or that it actually belongs into a later maturity range than the early flint material from the elite breeding program it was compared with.

Besides the yield, there are many other traits relevant in the changing environment we are currently facing. Landraces are expected to carry beneficial variation for traits they have either been selected for in the past or also due to a reduced loss of genetic variation through selection compared to elite material. As shown by Mayer et al. (2020), the studied landraces actually carried beneficial haplotypes for cold tolerance and early plant development. For this thesis, early development under cold conditions was chosen as an example trait. The landraces are derived from regions with naturally occurring cold conditions during early development and, indeed, many testcrosses of landrace-derived lines (both DH and GC and both landraces) outperformed a set of recent commercial hybrids for example for early plant height at V6 stage.

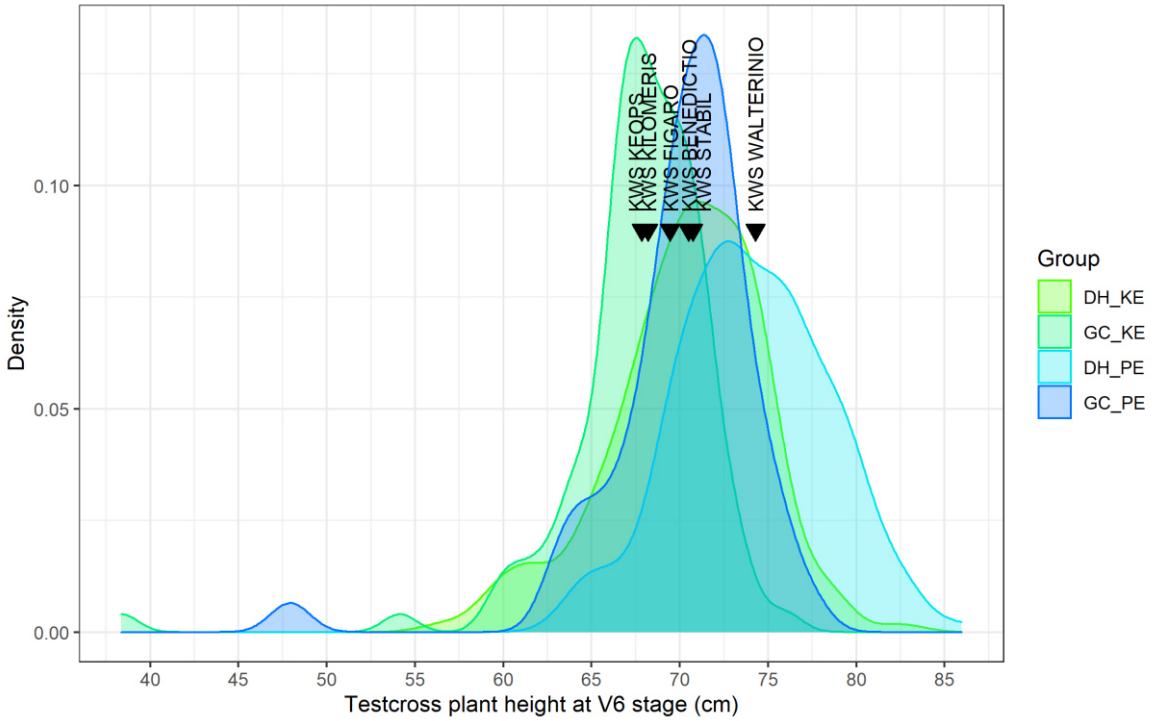


Figure 5: Density plot for testcross plant height at V6 stage (cm) for testcrosses of lines from DH_KE (N = 183), GC_KE (N = 103), DH_PE (N = 173), and GC_PE (N = 54) along with the performance of six commercial check hybrids (indicated as arrows). Phenotypic evaluation was done in in two environments in 2019.

Similar advantages for early development performance have been shown also for other European maize landraces (Weiß et al. 2022a). The early development trait is only one example showing that a proportion of the variation present in landraces is beneficial and useful for elite germplasm improvement. This example also highlights the necessity of developing suitable strategies for the utilization of the diversity, as in all landrace material favorable genetic variation can be linked to unfavorable alleles for yield and possibly also other agronomically relevant traits. The trait networks presented in Hölker et al. (2019a), indicate that the genetic correlations between early development traits and the most important agronomically unfavorable traits (e.g. root lodging, tillering) are only low to intermediate, making an improvement of agronomic performance while keeping the good performance in early development traits feasible.

3.2 Genome-based prediction accuracy in landrace populations

Genome-based prediction has quickly developed into a standard method in plant breeding. Genomic prediction in elite germplasm is strongly driven by relatedness in the breeding populations. In populations derived from panmictic landraces relatedness

between individuals is low (Hölker et al. 2019a), thus raising the question whether genome-based prediction will result in equally high predictive abilities as observed in elite germplasm. The prediction accuracies presented as part of this thesis demonstrate great potential for genome-based selection also in landrace-derived material (Hölker et al. 2022). For this thesis a variety of important factors that are known to influence the prediction accuracy were studied in detail.

3.2.1 Factors influencing genome-based prediction accuracy

Many parameters are known to influence genome-based prediction accuracy, for example relatedness, linkage phase similarity, and number of polymorphic SNPs shared between training and validation set (Auinger et al. 2021). Those parameters in return influence the maximum achievable prediction accuracy and the required training set size and number of markers to reach said maximum prediction accuracy. Therefore, especially the required number of markers to reach a plateau in the prediction accuracy can be indicative for population structure and LD in the population. With large LD blocks and very strong population structure, the required number of markers for the maximum prediction accuracy is expected to be low. In elite germplasm, < 1,000 markers can already be sufficient to describe the relationship well enough to reach sufficiently high prediction accuracies (Crossa et al. 2014), but many studies in maize used around 10,000 to 20,000 markers (Windhausen et al. 2012; Albrecht et al. 2014; Auinger et al. 2021). In research on genetic resources, studies often use much larger numbers of markers in order to capture all present diversity (Yu et al. 2016; Navarro et al. 2017). For landraces KE and PE, the number of markers used in the prediction did have a strong influence on the prediction accuracy, as expected (Hölker et al. 2022). The plateau of prediction accuracy was reached at around 15,000 markers with the studied traits early and final plant height and flowering time. This is only a small proportion of the total number of genotyped markers used on the landrace-derived populations, but quite high compared to elite germplasm, where prediction accuracies close to the maximum could sometimes be reached with less than 1000 markers (Crossa et al. 2014; Albrecht 2015). More markers are required for successful genomic predictions in landrace material, as the level of LD and the amount of population structure are much lower compared to common population types in elite breeding (e.g. biparental populations). The plateau of prediction accuracy at 15,000 markers indicates that the remaining polymorphisms of the 600k chip seem to be well-represented already through this limited set of markers (e.g. because of LD). However, the landraces were not part of the discovery set used for array design of the chip, which might result in missing loci that are relevant in the landrace material but were not polymorphic in the discovery set

(Nielsen 2004). An inclusion of such loci could theoretically lead to further increase of the prediction accuracy and would have to be investigated with sequence-based genotyping approaches instead of pre-defined chips, requiring further research.

A second well-known factor influencing the prediction accuracy is the size of the training set (N). Increasing training set size has been shown to generally increase prediction accuracy (Auinger et al. 2021). So far, most results available in literature were not obtained in the context of genetic resources, though. The size of the training set could be even more important in this case, as the LD and population structure in the landrace-derived populations is lower than in elite germplasm. For within landrace predictions, Brauner et al. (2018) showed strong influence of the training set size on the prediction accuracy and, with the limited available maximum size, they did not reach any plateau in prediction accuracy. The results of this thesis (Hölker et al. 2022) highlighted that even with large samples derived from a single landrace (up to $N > 450$), no clear plateau of prediction accuracy was reached yet. Therefore, for successfully applying genomic selection in landrace-derived populations large training sets are required and the maximum sample size studied in this thesis can be recommended as a lower boundary.

3.2.2 Prediction accuracy within and across populations and landraces

Sampling of admixed and pure populations from different landraces allows to study prediction accuracies within and across the sampled populations and landraces. The main questions to be answered are (i) if there are systematic differences between pure and admixed populations with respect to within population prediction accuracies, (ii) how high prediction accuracies are in across population predictions for the same landrace (pure to admixed or vice versa), and (iii) which prediction accuracies can be achieved in across landrace predictions, dependent on the used population type. The results of this thesis give insights into how to set up efficient GP schemes for landrace improvement in pre-breeding. Obviously, pre-breeding activities from landrace material aim at closing the performance gap to elite germplasm as fast as possible. To achieve this, a high prediction accuracy is key. However, the achieved prediction accuracy must not be evaluated isolated from other aspects that influence the maximum possible genetic gain like time required for population development and recycling. It could for example be a potential strategy to set up high-quality training sets using the pure DH approach and afterwards use these calibrations for predicting admixed populations from the same landrace, which can be produced at much lower costs than DH populations. This would require the high effort for DH production only once, but is only feasible, if the high-quality DH calibration

set has sufficiently high prediction accuracy when used for predicting other admixed populations from the same landrace.

The prediction results in this study show differences in prediction accuracy among traits, but the overall tendency is that the population type determines the prediction accuracy. DH lines had consistently higher prediction accuracies than GC lines at comparable sample sizes in calibration and prediction sets (Hölker et al. 2022). Prediction accuracies across the different types of populations were lower than within populations and no advantage of the direction of prediction (from DH to GC or vice versa) could be observed. Thus, the use of one high-quality DH population in training for prediction of large GC populations in pre-breeding is not very promising.

In sorghum it has been shown that predictions across a broad set of landrace accessions had high prediction accuracy, but this result was influenced strongly by mean differences driving prediction accuracies (Yu et al. 2016). The prediction of mean differences is not useful for ranking the individuals within a landrace in across landrace predictions. In this thesis, prediction accuracies across landraces differed for the two population types. In DH populations, prediction accuracy was close to zero. The GC populations had higher accuracies across landraces (Hölker et al. 2022). The generally high prediction accuracy within DH populations can be attributed to accurate phenotyping (no segregation in lines) and the full exploitation of the genetic variance in the DH lines. For a given trait, nonsignificant estimates of the genetic variance in the training set occurred in more samples in GC than in DH, indicating the lower genetic variance present. Training and prediction within GC populations resulted in higher prediction accuracies than the across population prediction. In the GC population development process, a relatedness among the produced populations was introduced with the use of the common capture line FV2, resulting in higher prediction accuracies. This increase was mainly driven by increased linkage phase similarities in GC lines due to shared haplotypes from FV2 (Hölker et al. 2022). Accordingly, the, at a first glance, advantage of higher across landrace prediction accuracies in GC populations is not useful in pre-breeding as it relies mainly on the common capture line and the situation in the DH populations represents the true expected across-landrace prediction accuracies.

3.2.3 Implications for the use of genomic prediction in pre-breeding

This thesis delivers important insights on the use of genomic prediction in pre-breeding with landrace material. Achieving high genomic prediction accuracies in landrace-derived populations is possible and prediction accuracies > 0.5 with training set sizes > 200 can

be considered realistic in DH populations from landraces. With sufficiently sized high-quality calibration data, despite much lower LD and relatedness, prediction accuracies within landrace-derived populations can be close to values observed in elite germplasm (Auinger et al. 2021). Therefore, any strategy for utilizing landrace material should make use of genomic prediction.

The method used for developing populations of reproducible genetic units from landraces (pure vs. admixed) has strong influence on prediction accuracy. Prediction accuracies in this thesis were consistently higher for the fully inbred DH populations. However, genomic prediction in GC populations was also feasible, just at a lower prediction accuracy. The overall impact of this result must be considered as part of a comprehensive pre-breeding strategy. Achieving the highest possible prediction accuracy will not be the goal of a pre-breeding strategy. It is the overall selection gain per year of pre-breeding that should be maximized and in which the prediction accuracy is only one component. Here, the admixed approach can still have advantages, e.g. if the capture line is a high-performing elite line the performance gap can be reduced in the GC populations or if DH line production is not possible with reasonable resource input. If the prediction accuracy in the GC lines is, at least partially, driven by additive effects of the capture line haplotypes might be undesirable if the goal is to select the best landrace-derived haplotypes. However, it is not proven that this will necessarily be the case.

As an alternative to genomic prediction, the use of phenomic selection may be a promising approach to improve prediction accuracy in across-population and across-landrace scenarios, as constructing the relationship matrix from spectral data and using this matrix for prediction of genotypes has been shown to be less influenced by population structure and therefore may exceed the accuracy of marker-based predictions (Weiß et al. 2022b). The two relationship matrices can also be combined in multi-kernel models that have been shown to be superior to any of the two single-kernel prediction models (Galán et al. 2020; Robert et al. 2022). The implementation of phenomic selection therefore could be a comparatively cheap and promising extension to further enhance prediction accuracies, due to the cost-efficient options to screen samples with near-infrared spectroscopy or gather hyperspectral imaging data through drones. This topic requires further research.

3.3 Prospects of improving elite germplasm through landraces

Maize is known for its very large species-wide variation, of which only a part is currently captured in current elite breeding germplasm (Buckler et al. 2006; Sood et al. 2014). This

is not inherently a problematic situation, but in order to make selection gain, genetic variation needs to be present and intense selection, as done in the breeding process, leads to a reduction in genetic variation. This could eventually lead to a situation where further gain is not possible without the introduction of new genetic variation. While the variance in breeding programs will be monitored by breeders to avoid the depletion of variation for traits that are currently in the focus of the breeding, this is not the case for many traits currently not in focus of the breeding activities. However, the future changes in agriculture and the climatic conditions in which maize is grown in may require a focus on traits other than yield. For new target traits elite germplasm might then still exhibit genetic variation by chance, but this is not guaranteed. Genetic drift introduced through the history of strong bottlenecks in domestication and breeding may have resulted in fixation of alleles for traits relevant for future crop cultivation. Another possibility is that landraces that harbored variation for the desired traits might not have contributed as founders to the current elite germplasm at all. If this was the case, there is the need to look for new sources of variation. The obvious choice before turning to landrace material as a source would of course be competitor material under the breeder's exemption, where possible, or US material with expired Plant Variety Protection Act certificates (exPVP). Although this material group has either no (competitor material) or only a small to intermediate performance gap (exPVP), it still requires caution and a strategy for integrating it successfully into the breeding germplasm (Allier et al. 2020a). If there is no beneficial variation present in these sources, it will be necessary to turn to landrace material with the desired variation. As shown in this thesis and by Mayer et al. (2020), preselected landraces can be competitive with elite material in some specific traits and do indeed carry beneficial haplotypes. For having a chance to successfully utilize landraces for elite germplasm improvement two main aspects need to be fulfilled: First, the pressure on the elite germplasm breeding program to search for new variation and accept the possibility of a yield drawback through landrace introgression needs to be high enough to invest extra efforts into breeding strategies using landrace-derived material. Secondly, the drawbacks of landrace utilization need to be minimized by having efficient strategies in place. The amount of pressure exerted onto breeding programs to start venturing into genetic resources is driven by characteristics of their germplasm pools (e.g. lack of variation for specific traits) and external factors that might come through demand changes based on climate or also political changes (e.g. Green New Deal and European Green Deal). The required efficient strategies for the incorporation of landraces into elite germplasm are under development and there are several conclusions that can be drawn from this thesis. The choice of source material is

very important and should be as informed as possible (Mayer et al. 2017). The production of reproducible genetic units from landraces can be done in different ways, following either admixed or pure approaches. Both approaches were shown to be feasible, but with many advantages of the pure approach, as presented in this thesis. Nevertheless, if following a pure approach efficiently is not possible (e.g. allogamous crops without DH system), making crosses for an admixed approach is suitable, too.

The private/public partnership projects LAMP and GEM were already introduced earlier. They are examples for large pre-breeding efforts from landraces and follow the admixed approach. The industry partners used in total 51 preselected LAMP accessions and crossed them to each industry partner's proprietary inbred lines to obtain progeny containing 50% (crosses to one elite inbred line) or 25% (use 50% inbred/elite crosses and cross to another, but different, inbred line) exotic material (Pollak 2003). Later, populations were also produced using recycled GEM and exPVP lines and genomic selection (GS) was evaluated for selections in GEM (Rogers et al. 2022). As can be seen by the fact that the GEM program has been running since in 1993, it was initiated as a long-term pre-breeding project. Up to now, 334 inbred lines were released (Rogers et al. 2022), most of them coming from the conventional breeding and bridging process described above. Only recently, the first evaluation of GS in GEM was considered successful so that it was recommended to switch from the original phenotyping-only protocol to the use of GS in GEM (Rogers et al. 2022) for the future. Another project carrying the acronym HOPE (Hierarchical Open-Ended Corn Breeding System) was designed to increase genetic diversity and long-term gain by a continuous introgression of variability-adding germplasm into breeding populations (Cramer and Kannenberg 1992). The introgression followed a pyramid-like scheme with increasingly stringent selection criteria while the material moved from a low via an intermediate and a high to a final elite performance level. The HOPE results reported in 1992 reflect efforts from 5 years during the 1980s and thus were based on phenotypic selections only (using mass selection in the lower, ear-to-row methods in the intermediate and reciprocal recurrent selection in the highest performance levels) and without any consideration of molecular information. The authors could show improved performance, but the project cannot be viewed as a suitable strategy under today's conditions.

Beyond the described example projects of practical pre-breeding efforts, studies on the theory and simulations of pre-breeding schemes have been done in many ways. Breeding programs often use truncation selection for highest yield as an intuitive way to ensure high

short-term genetic gain. This method of selection is associated with loss of genetic variation (Jannink 2010) and may lead to a reduction in long-term genetic gain (Vanavermaete et al. 2020). In the context of admixed populations in pre-breeding it additionally carries the risk of reconstructing the elite crossing partner's genotype during selection after crosses between genetic resource and elite material (Gorjanc et al. 2016). Any pre-breeding strategy for the utilization of genetic resources has large long-term genetic gain as the goal and will not be able to outperform a conventional breeding program using simple truncation selection from elite material in the short term. Therefore, for the pre-breeding strategies maintaining diversity besides selecting the best genomic estimated breeding values (GEBVs) is necessary. For elite breeding populations, several ways to maximize long-term genetic gain were suggested, like applying the expected maximum haploid breeding value (Müller et al. 2018), optimum contribution selection (Gorjanc et al. 2018) or the scoping method which combines a performance-based truncation selection (GEBV-based) with a mating design that maximizes genetic variation of the offspring (marker-based coupling of diverse crossing partners) to maximize long-term genetic gain (Vanavermaete et al. 2020). These methods were tested on populations exhibiting broad genetic variation. However, when using landrace material for elite germplasm improvement, the breeding population might rather be a variation depleted population with little room for genetic gain for certain traits. Vanavermaete et al. (2021) developed a strategy for this specific case and termed it "deep scoping". Briefly, the method involves using a selected fraction of a breeding population for pre-breeding and a combination of short-term gains with purely elite individuals with highest GEBVs and a layered pre-breeding that introduces favorable alleles gradually from the source to the elite population. The simulation results indicate that the deep scoping method increases the achievable long-term genetic value of the breeding program while not increasing its costs (Vanavermaete et al. 2021). It would be very interesting to implement deep scoping in pre-breeding and evaluate the simulation results empirically.

Genomic selection is working very well and should be part of any strategy for the utilization of landraces in elite germplasm improvement. GS is suitable to speed up the breeding process and therefore increase genetic gains, which is why it also is standard practice in elite breeding. Recently, it has been suggested for elite breeding to increase speed of breeding (and therefore selection gain) with stable budget even further by inserting selection and recombination cycles that are only based on GS while making use of contra season nurseries that can handle three or four generations per year (Bernardo 2021). Moving from adding one additional recombination as suggested by Bernardo (2021) to

several recombinations without retraining of the prediction model may increase selection gain even further and has been proposed as rapid cycling genomic selection (Zhang et al. 2017). Encouraging results for such schemes have been shown for wheat (Dreisigacker et al. 2023) and it can be expected that those rapid GS breeding schemes are also adopted in a broad range of commercial breeding schemes, including in pre-breeding. The expected increase in selection gain by adoption of those new fast schemes in elite breeding further enlarges the challenge of closing the performance gap in pre-breeding with landrace material. Further research is warranted if and how rapid GS schemes in landrace-derived populations can be employed. If the prediction accuracy remains sufficiently high over several cycles without retraining, two different applications can be imagined for pre-breeding with landraces: On one hand, a small number of rapid GS cycles over 1-2 years could be used for fast extraction of the best alleles per landrace as a short-termed approach. The rapid cycling would then be a way to recombine the genetic material as often as possible in a very quick way in order to enrich as many beneficial alleles of the landrace as possible in just a few individuals to be integrated into the elite germplasm. On the other hand, a recurrent improvement of a specific landrace population over a longer term could also be possible with rapid GS. In both cases, a very big challenge is the huge variability in landrace-derived populations. The selection procedure needs to be well designed and sophisticated to produce useful material in the end. Besides the required reduction of the yield gap, the landrace material may show large variation for many other unwanted agronomic traits like tillering, root lodging, disease susceptibility and many more. In conventional breeding schemes more time is available to also make selections for these traits while in rapid GS every relevant trait needs to be covered through predictions and the different traits need to be considered in selection simultaneously, e.g. by an index or separate truncation selections, which is also a challenge. The unique data and plant material generated for this thesis are an ideal starting point for further research on the use of rapid-GS schemes in genetic resources.

3.4 Conclusion

In this thesis, the prospects of using maize landraces for elite germplasm improvement with the use of genomic prediction in pre-breeding scenarios were evaluated. The main conclusions can be summarized as follows:

- Sampling from landraces following a pure or admixed approach is equally possible and both result in an unbiased representation of the original landrace. The pure approach has several advantages (full additive variance, no masking by crossing partner alleles, no risk of reconstruction of elite crossing partner genome, higher prediction accuracy).
- As expected from literature results, the landrace-derived populations exhibit a performance gap to elite germplasm of about 15-20% in yield. Besides the yield performance gap, landraces exhibit variation for unwanted traits against which selection pressure needs to be applied. However, landraces also show large variation and superior performance compared to elite hybrids for desirable traits, for which early development is one example in this thesis.
- If an admixed approach is followed, intense consideration related to the choice of the crossing partner is strongly advised, as its genetic makeup and performance has very large influence on the populations that are produced and the risk of a reconstruction of the elite crossing partner in subsequent selections.
- The genomic prediction accuracies are strongly influenced by factors already known from research in elite germplasm. They are training set size, linkage disequilibrium and population structure, linkage phase similarity between training and prediction set and number of markers required for maximum prediction accuracy.
- If training set sizes and number of markers are sufficiently high, prediction accuracies close to those in elite germplasm can be achieved.
- Using genomic prediction while working with landrace material is feasible. It is not restricted to a specific utilization strategy and can be used no matter which type of populations are chosen. Genomic prediction is already a standard tool in elite breeding and should also be part of any strategy for utilizing genetic resources.
- In particular, fast recombinations in rapid cycles with heavy use of genomic selection (without retraining before each selection) might be an interesting option for fast closure of the performance gap when working with landrace material, but require further research before application in pre-breeding.

4 References

Albrecht MT (2015) Genome-based prediction of testcross performance in maize (Zea mays L.). Technische Universität München

Albrecht T, Auinger H-J, Wimmer V, et al (2014) Genome-based prediction of maize hybrid performance across genetic groups, testers, locations, and years. Theoretical and Applied Genetics 127:1375–1386. https://doi.org/10.1007/s00122-014-2305-z

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655–1664. https://doi.org/10.1101/gr.094052.109

Allier A, Teyssèdre S, Lehermeier C, et al (2019) Assessment of breeding programs sustainability: application of phenotypic and genomic indicators to a North European grain maize program. Theor Appl Genet 132:1321–1334. https://doi.org/10.1007/s00122-019-03280-w

Allier A, Teyssèdre S, Lehermeier C, et al (2020a) Optimized breeding strategies to harness genetic resources with different performance levels. BMC Genomics. https://doi.org/10.1186/s12864-020-6756-0

Allier A, Teyssèdre S, Lehermeier C, et al (2020b) Genomic prediction with a maize collaborative panel: identification of genetic resources to enrich elite breeding programs. Theor Appl Genet 133:201–215. https://doi.org/10.1007/s00122-019-03451-9

Auinger H-J, Lehermeier C, Gianola D, et al (2021) Calibration and validation of predicted genomic breeding values in an advanced cycle maize population. Theor Appl Genet 134:3069–3081. https://doi.org/10.1007/s00122-021-03880-5

Bailey-Serres J, Fukao T, Ronald P, et al (2010) Submergence Tolerant Rice: SUB1's Journey from Landrace to Modern Cultivar. Rice 3:138–147. https://doi.org/10.1007/s12284-010-9048-5

Bernardo R (1994) Prediction of Maize Single-Cross Performance Using RFLPs and Information from Related Hybrids. Crop Science 34:20–25. https://doi.org/10.2135/cropsci1994.0011183X003400010003x

Bernardo R (2020) Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE. Heredity 125:375–385

Bernardo R (2021) Upgrading a maize breeding program via two-cycle genomewide selection: Same cost, same or less time, and larger gains. Crop Science 61:2444–2455. https://doi.org/10.1002/csc2.20516

Böhm J, Schipprack W, Utz HF, Melchinger AE (2017) Tapping the genetic diversity of landraces in allogamous crops with doubled haploid lines: a case study from European flint maize. Theoretical and Applied Genetics 130:861–873. https://doi.org/10.1007/s00122-017-2856-x

Brauner PC, Müller D, Schopp P, et al (2018) Genomic Prediction Within and Among Doubled-Haploid Libraries from Maize Landraces. Genetics. https://doi.org/10.1534/genetics.118.301286

Brauner PC, Schipprack W, Utz HF, et al (2019) Testcross performance of doubled haploid lines from European flint maize landraces is promising for broadening the genetic base of elite germplasm. Theor Appl Genet 132:1897–1908. https://doi.org/10.1007/s00122-019-03325-0

Browning BL, Zhou Y, Browning SR (2018) A One-Penny Imputed Genome from Next-Generation Reference Panels. The American Journal of Human Genetics 103:338–348. https://doi.org/10.1016/j.ajhg.2018.07.015

Buckler ES, Gaut BS, McMullen MD (2006) Molecular and functional diversity of maize. Current Opinion in Plant Biology 9:172–176. https://doi.org/10.1016/j.pbi.2006.01.013

Butler D, Cullis BR, Gilmour A, Gogel B (2009) ASReml-R reference manual. The State of Queensland, Department of Primary Industries and Fisheries, Brisbane

Cramer MM, Kannenberg LW (1992) Five Years of HOPE: The Hierarchical Open-Ended Corn Breeding System. Crop Science 32:1163–1171. https://doi.org/10.2135/cropsci1992.0011183X003200050021x

Crossa J, Jarquín D, Franco J, et al (2016) Genomic Prediction of Gene Bank Wheat Landraces. G3 (Bethesda) 6:1819–1834. https://doi.org/10.1534/g3.116.029637

Crossa J, Pérez P, Hickey J, et al (2014) Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity 112:48–60. https://doi.org/10.1038/hdy.2013.16

Crossa J, Pérez-Rodríguez P, Cuevas J, et al (2017) Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. Trends in Plant Science 22:961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Darrier B, Russell J, Milner SG, et al (2019) A Comparison of Mainstream Genotyping Platforms for the Evaluation and Use of Barley Genetic Resources. Frontiers in Plant Science. https://doi.org/10.3389/fpls.2019.00544

Dekkers JCM (2007) Marker-assisted selection for commercial crossbred performance. J Anim Sci 85:2104–2114. https://doi.org/10.2527/jas.2006-683

Dreisigacker S, Pérez-Rodríguez P, Crespo-Herrera L, et al (2023) Results from rapid-cycle recurrent genomic selection in spring bread wheat. G3 Genes|Genomes|Genetics 13:jkad025. https://doi.org/10.1093/g3journal/jkad025

Dubreuil P, Warburton ML, Chastanet M, et al (2006) More on the introduction of temperate maize into Europe: large-scale bulk SSR genotyping and new historical elements. Maydica 51:281–291

Duvick DN (2005) The Contribution of breeding to yield advances in maize (Zea mays L.). In: Advances in Agronomy. Academic Press, pp 83–145

Edgerton MD (2009) Increasing Crop Productivity to Meet Global Needs for Feed, Food, and Fuel. Plant Physiology 149:7–13. https://doi.org/10.1104/pp.108.130195

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Falconer D, Mackay T (1996) Introduction to quantitative genetics, 4th edn. Longmann Scientific & Technical, New York

Galán RJ, Bernal-Vasquez A-M, Jebsen C, et al (2020) Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye. Theor Appl Genet 133:3001–3015. https://doi.org/10.1007/s00122-020-03651-8

Geiger HH (2009) Doubled Haploids. In: Handbook of Maize. Springer, New York, NY, pp 641–657

Gianola D, de Los Campos G, Hill WG, et al (2009) Additive genetic variability and the Bayesian alphabet. Genetics 183:347–363

Godfray HCJ, Beddington JR, Crute IR, et al (2010) Food Security: The Challenge of Feeding 9 Billion People. Science 327:812–818. https://doi.org/10.1126/science.1185383

Gorjanc G, Gaynor RC, Hickey JM (2018) Optimal cross selection for long-term genetic gain in two-part programs with rapid recurrent genomic selection. Theor Appl Genet 131:1953–1966. https://doi.org/10.1007/s00122-018-3125-3

Gorjanc G, Jenko J, Hearne SJ, Hickey JM (2016) Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. BMC Genomics 17:30. https://doi.org/10.1186/s12864-015-2345-z

Gower JC (1966) Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis. Biometrika 53:325–338. https://doi.org/10.2307/2333639

Habier D, Fernando RL, Dekkers JCM (2007) The Impact of Genetic Relationship Information on Genome-Assisted Breeding Values. Genetics 177:2389–2397. https://doi.org/10.1534/genetics.107.081190

Hallauer AR, Carena MJ, Miranda Filho J de (2010) Quantitative Genetics in Maize Breeding, 6th edn. Springer Science & Business Media, New York

Hallauer AR, Sears JH (1972) Integrating Exotic Germplasm into Corn Belt Maize Breeding Programs 1. Crop Science 12:203–206. https://doi.org/10.2135/cropsci1972.0011183X001200020016x

Hayes BJ, Visscher PM, Goddard ME (2009) Increased accuracy of artificial selection by using the realized relationship matrix. Genetics research 91:47–60

Heerwaarden J van, Doebley J, Briggs WH, et al (2011) Genetic signals of origin, spread, and introgression in a large sample of maize landraces. Proceedings of the National Academy of Sciences 108:1088–1092. https://doi.org/10.1073/pnas.1013011108

Hellin J, Bellon MR, Hearne SJ (2014) Maize Landraces and Adaptation to Climate Change in Mexico. Journal of Crop Improvement 28:484–501. https://doi.org/10.1080/15427528.2014.921800

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423–447

Heslot N, Yang H-P, Sorrells ME, Jannink J-L (2012) Genomic selection in plant breeding: a comparison of models. Crop science 52:146–160

Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. Theoret Appl Genetics 38:226–231. https://doi.org/10.1007/BF01245622

Hill WG, Weir BS (1988) Variances and covariances of squared linkage disequilibria in finite populations. Theoretical Population Biology 33:54–78. https://doi.org/10.1016/0040-5809(88)90004-4

Hoisington D, Khairallah M, Reeves T, et al (1999) Plant genetic resources: What can they contribute toward increased crop productivity? Proceedings of the National Academy of Sciences 96:5937–5943

Hölker AC, Mayer M, Presterl T, et al (2019a) European maize landraces made accessible for plant breeding and genome-based studies. Theor Appl Genet 132:3333–3345. https://doi.org/10.1007/s00122-019-03428-8

Hölker AC, Mayer M, Presterl T, et al (2022) Theoretical and experimental assessment of genome-based prediction in landraces of allogamous crops. Proceedings of the National Academy of Sciences 119:e2121797119. https://doi.org/10.1073/pnas.2121797119

Hölker AC, Schipprack W, Utz HF, et al (2019b) Progress for testcross performance within the flint heterotic pool of a public maize breeding program since the onset of hybrid breeding. Euphytica 215:50. https://doi.org/10.1007/s10681-019-2370-0

Holland JB, Goodman MM, Castillo-Gonzalez F (1996) Identification of Agronomicaily Superior Latin American Maize Accessions via Multi-Stage Evaluations. Crop Science 36:778–784. https://doi.org/10.2135/cropsci1996.0011183X003600030041x

Holland JB, Nyquist WE, Cervantes-Martínez CT (2010) Estimating and Interpreting Heritability for Plant Breeding: An Update. In: Plant Breeding Reviews. Wiley-Blackwell, pp 9–112

Holm S (1979) A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics 6:65–70

Hufford MB, Xu X, van Heerwaarden J, et al (2012) Comparative population genomics of maize domestication and improvement. Nat Genet 44:808–811. https://doi.org/10.1038/ng.2309

Jacobson A, Lian L, Zhong S, Bernardo R (2015) Minimal Loss of Genetic Diversity after Genomewide Selection within Biparental Maize Populations. Crop Science 55:783–789. https://doi.org/10.2135/cropsci2014.10.0744

Jannink J-L (2010) Dynamics of long-term genomic selection. Genetics Selection Evolution 42:35. https://doi.org/10.1186/1297-9686-42-35

Jiao Y, Peluso P, Shi J, et al (2017) Improved maize reference genome with single-molecule technologies. Nature 546:524–527. https://doi.org/10.1038/nature22971

Khush GS (2001) Green revolution: the way forward. Nat Rev Genet 2:815–822. https://doi.org/10.1038/35093585

Kremling KAG, Chen S-Y, Su M-H, et al (2018) Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. Nature 555:520–523. https://doi.org/10.1038/nature25966

Leff B, Ramankutty N, Foley JA (2004) Geographic distribution of major crops across the world. Global Biogeochemical Cycles 18:. https://doi.org/10.1029/2003GB002108

Lobell DB, Tebaldi C (2014) Getting caught with our plants down: the risks of a global crop yield slowdown from climate trends in the next two decades. Environ Res Lett 9:074003. https://doi.org/10.1088/1748-9326/9/7/074003

Lu Y, Yan J, Guimarães CT, et al (2009) Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. Theor Appl Genet 120:93–115. https://doi.org/10.1007/s00122-009-1162-7

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

Marone D, Russo MA, Mores A, et al (2021) Importance of Landraces in Cereal Breeding for Stress Tolerance. Plants 10:1267. https://doi.org/10.3390/plants10071267

Mayer M, Hölker AC, González-Segovia E, et al (2020) Discovery of beneficial haplotypes for complex traits in maize landraces. Nature Communications 11:4954. https://doi.org/10.1038/s41467-020-18683-3

Mayer M, Unterseer S, Bauer E, et al (2017) Is there an optimum level of diversity in utilization of genetic resources? Theor Appl Genet 130:2283–2295. https://doi.org/10.1007/s00122-017-2959-4

McCouch S, Baute GJ, Bradeen J, et al (2013) Agriculture: Feeding the future. Nature 499:23–44. https://doi.org/10.1038/499023a

McCouch SR, McNally KL, Wang W, Sackville Hamilton R (2012) Genomics of gene banks: A case study in rice. American Journal of Botany 99:407–423. https://doi.org/10.3732/ajb.1100385

Melchinger AE, Schopp P, Müller D, et al (2017) Safeguarding Our Genetic Resources with Libraries of Doubled-Haploid Lines. Genetics 206:1611–1619. https://doi.org/10.1534/genetics.115.186205

Mercer KL, Perales HR (2010) Evolutionary response of landraces to climate change in centers of crop diversity. Evolutionary applications 3:480–493

Messmer MM, Melchinger AE, Boppenmaier J, et al (1992) RFLP analyses of early-maturing European maize germ plasm. Theor Appl Genet 83:1003–1012. https://doi.org/10.1007/bf00232964

Meuwissen T, Hayes B, Goddard M (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Mezmouk S, Ross-Ibarra J (2014) The Pattern and Distribution of Deleterious Mutations in Maize. G3: Genes, Genomes, Genetics 4:163–171. https://doi.org/10.1534/g3.113.008870

Mir C, Zerjal T, Combes V, et al (2013) Out of America: tracing the genetic footprints of the global diffusion of maize. Theor Appl Genet 126:2671–2682. https://doi.org/10.1007/s00122-013-2164-z

Müller D, Schopp P, Melchinger AE (2018) Selection on Expected Maximum Haploid Breeding Values Can Increase Genetic Gain in Recurrent Genomic Selection. G3: Genes, Genomes, Genetics. https://doi.org/10.1534/g3.118.200091

Navarro JAR, Willcox M, Burgueño J, et al (2017) A study of allelic diversity underlying flowering-time adaptation in maize landraces. Nature Genetics 49:476–480. https://doi.org/10.1038/ng.3784

Nellemann C, MacDevette M (2009) The environmental food crisis: the environment's role in averting future food crises: a UNEP rapid response assessment. UNEP/Earthprint

Nielsen R (2004) Population genetic analysis of ascertained SNP data. Hum Genomics 1:218. https://doi.org/10.1186/1479-7364-1-3-218

Nuss ET, Tanumihardjo SA (2010) Maize: A Paramount Staple Crop in the Context of Global Nutrition. Comprehensive Reviews in Food Science and Food Safety 9:417–436. https://doi.org/10.1111/j.1541-4337.2010.00117.x

Ortiz R, Taba S, Tovar VHC, et al (2010) Conserving and Enhancing Maize Genetic Resources as Global Public Goods – A Perspective from CIMMYT. Crop Science 50:13–28. https://doi.org/10.2135/cropsci2009.06.0297

Paull JG, Nable RO, Rathjen AJ (1992) Physiological and genetic control of the tolerance of wheat to high concentrations of boron and implications for plant breeding. Plant Soil 146:251–260. https://doi.org/10.1007/BF00012019

Piperno DR, Ranere AJ, Holst I, et al (2009) Starch grain and phytolith evidence for early ninth millennium BP maize from the Central Balsas River Valley, Mexico. Proceedings of the National Academy of Sciences 106:5019–5024

Pollak LM (2003) The history and success of the public-private project on germplasm enhancement of maize (GEM). Advances in Agronomy 78:46–89

R Core Team (2019) R: A language and environment for statistical computing. Vienna

Rebourg C, Chastanet M, Gouesnard B, et al (2003) Maize introduction into Europe: the history reviewed in the light of molecular data. Theoretical and Applied Genetics 106:895–903

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

Robert P, Brault C, Rincent R, Segura V (2022) Phenomic Selection: A New and Efficient Alternative to Genomic SelectionGenomic selection (GS). In: Ahmadi N, Bartholomé J (eds) Genomic Prediction of Complex Traits: Methods and Protocols. Springer US, New York, NY, pp 397–420

Rogers AR, Bian Y, Krakowsky M, et al (2022) Genomic prediction for the Germplasm Enhancement of Maize project. The Plant Genome 15:e20267. https://doi.org/10.1002/tpg2.20267

Russell J, Mascher M, Dawson IK, et al (2016) Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. Nature Genetics 48:1024–1030. https://doi.org/10.1038/ng.3612

Salhuana W, Pollak L (2006) Latin American Maize Project (LAMP) and Germplasm Enhancement of Maize (GEM) project: generating useful breeding germplasm. Maydica 51:339–355

Schnell F (1983) Probleme der Elternwahl - Ein Überblick. Bericht-Arbeitstagung der Arbeitsgemeinschaft der Saatzuchtleiter

Schopp P, Müller D, Technow F, Melchinger AE (2017) Accuracy of Genomic Prediction in Synthetic Populations Depending on the Number of Parents, Relatedness, and Ancestral Linkage Disequilibrium. Genetics 205:441–454. https://doi.org/10.1534/genetics.116.193243

Sood S, Flint-Garcia S, Willcox MC, Holland JB (2014) Mining Natural Variation for Maize Improvement: Selection on Phenotypes and Genes. In: Genomics of Plant Genetic Resources. Springer, Dordrecht, pp 615–649

Stadler LJ (1944) Gamete selection in corn breeding. J Am Soc Agron 36:988–989

Strable J, Scanlon MJ (2009) Maize (Zea mays): a model organism for basic and applied research in plant biology. Cold spring harbor protocols 2009:pdb. emo132

Strigens A, Schipprack W, Reif JC, Melchinger AE (2013) Unlocking the Genetic Diversity of Maize Landraces with Doubled Haploids Opens New Avenues for Breeding. PLoS ONE. https://doi.org/10.1371/journal.pone.0057234

Takuno S, Ralph P, Swarts K, et al (2015) Independent Molecular Basis of Convergent Highland Adaptation in Maize. Genetics 200:1297–1312. https://doi.org/10.1534/genetics.115.178327

Tanksley SD, McCouch SR (1997) Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. Science 277:1063–1066. https://doi.org/10.1126/science.277.5329.1063

United Nations, (2019) World population prospects 2019: department of economic and social Affairs. World Population Prospects 2019

Unterseer S, Bauer E, Haberer G, et al (2014) A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. BMC Genomics 15:823. https://doi.org/10.1186/1471-2164-15-823

Vanavermaete D, Fostier J, Maenhout S, De Baets B (2020) Preservation of Genetic Variation in a Breeding Population for Long-Term Genetic Gain. G3 Genes|Genomes|Genetics 10:2753–2762. https://doi.org/10.1534/g3.120.401354

Vanavermaete D, Fostier J, Maenhout S, De Baets B (2021) Deep scoping: a breeding strategy to preserve, reintroduce and exploit genetic variation. Theor Appl Genet 134:3845–3861. https://doi.org/10.1007/s00122-021-03932-w

VanRaden PM (2008) Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science 91:4414–4423. https://doi.org/10.3168/jds.2007-0980

Vigouroux Y, Glaubitz JC, Matsuoka Y, et al (2008) Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. American Journal of Botany 95:1240–1253. https://doi.org/10.3732/ajb.0800097

Visscher PM, Haley CS, Thompson R (1996) Marker-Assisted Introgression in Backcross Breeding Programs. Genetics 144:1923–1932. https://doi.org/10.1093/genetics/144.4.1923

Wang C, Hu S, Gardner C, Lübberstedt T (2017) Emerging Avenues for Utilization of Exotic Germplasm. Trends in Plant Science 22:624–637. https://doi.org/10.1016/j.tplants.2017.04.002

Warburton M, Reif J, Frisch M, et al (2008) Genetic diversity in CIMMYT nontemperate maize germplasm: landraces, open pollinated varieties, and inbred lines. Crop Science 48:617–624

Weiß TM, Li D, Roller S, et al (2022a) How can we breed for phosphate efficiency in maize (Zea mays)? Plant Breeding 141:733–744. https://doi.org/10.1111/pbr.13054

Weiß TM, Zhu X, Leiser WL, et al (2022b) Unraveling the potential of phenomic selection within and among diverse breeding material of maize (*Zea mays* L.). G3 Genes|Genomes|Genetics 12:jkab445. https://doi.org/10.1093/g3journal/jkab445

White MR, Mikel MA, de Leon N, Kaeppler SM (2020) Diversity and heterotic patterns in North American proprietary dent maize germplasm. Crop Science 60:100–114. https://doi.org/10.1002/csc2.20050

Wilde K, Burger H, Prigge V, et al (2010) Testcross performance of doubled-haploid lines developed from European flint maize landraces. Plant Breeding 129:181–185. https://doi.org/10.1111/j.1439-0523.2009.01677.x

Wimmer V, Albrecht T, Auinger H-J, Schön C-C (2012) synbreed: a framework for the analysis of genomic prediction data using R. Bioinformatics 28:2086–2087. https://doi.org/10.1093/bioinformatics/bts335

Windhausen VS, Atlin GN, Hickey JM, et al (2012) Effectiveness of Genomic Prediction of Maize Hybrid Performance in Different Breeding Populations and Environments.

G3 Genes|Genomes|Genetics 2:1427–1436. https://doi.org/10.1534/g3.112.003699

Wright S (1978) Evolution and the Genetics of Populations, Volume 4. University of Chicago Press

Würschum T, Weiß TM, Renner J, et al (2022) High-resolution association mapping with libraries of immortalized lines from ancestral landraces. Theor Appl Genet 135:243–256. https://doi.org/10.1007/s00122-021-03963-3

Yamasaki M, Tenaillon MI, Vroh Bi I, et al (2005) A Large-Scale Screen for Artificial Selection in Maize Identifies Candidate Agronomic Loci for Domestication and Crop Improvement. The Plant Cell 17:2859–2872. https://doi.org/10.1105/tpc.105.037242

Yang J, Mezmouk S, Baumgarten A, et al (2017) Incomplete dominance of deleterious alleles contributes substantially to trait variation and heterosis in maize. PLOS Genetics 13:e1007019. https://doi.org/10.1371/journal.pgen.1007019

Yu X, Li X, Guo T, et al (2016) Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nature Plants 2:16150. https://doi.org/10.1038/nplants.2016.150

Zeitler L, Ross-Ibarra J, Stetter MG (2020) Selective Loss of Diversity in Doubled-Haploid Lines from European Maize Landraces. G3: Genes, Genomes, Genetics 10:2497–2506. https://doi.org/10.1534/g3.120.401196

Zhang X, Pérez-Rodríguez P, Burgueño J, et al (2017) Rapid Cycling Genomic Selection in a Multiparental Tropical Maize Population. G3 Genes|Genomes|Genetics 7:2315–2326. https://doi.org/10.1534/g3.117.043141

5 Publications

This section includes reprints of the two publications underlying this thesis. The two publications, including supplementary material, can be accessed via the following links:

Hölker et al. (2019)

https://link.springer.com/article/10.1007/s00122-019-03428-8

Hölker et al. (2022)

https://www.pnas.org/doi/full/10.1073/pnas.2121797119

Theoretical and Applied Genetics (2019) 132:3333–3345
https://doi.org/10.1007/s00122-019-03428-8

ORIGINAL ARTICLE

European maize landraces made accessible for plant breeding and genome-based studies

Armin C. Hölker[1] · Manfred Mayer[1] · Thomas Presterl[2] · Therese Bolduan[2] · Eva Bauer[1] · Bernardo Ordas[3] · Pedro C. Brauner[4,5] · Milena Ouzunova[2] · Albrecht E. Melchinger[4] · Chris-Carolin Schön[1]

Received: 3 July 2019 / Accepted: 17 September 2019 / Published online: 26 September 2019

Key message **Doubled-haploid libraries from landraces capture native genetic diversity for a multitude of quantitative traits and make it accessible for breeding and genome-based studies.**

Abstract Maize landraces comprise large allelic diversity. We created doubled-haploid (DH) libraries from three European flint maize landraces and characterized them with respect to their molecular diversity, population structure, trait means, variances, and trait correlations. In total, 899 DH lines were evaluated using high-quality genotypic and multi-environment phenotypic data from up to 11 environments. The DH lines covered 95% of the molecular variation present in 35 landraces of an earlier study and represent the original three landrace populations in an unbiased manner. A comprehensive analysis of the target trait plant development at early growth stages as well as other important agronomic traits revealed large genetic variation for line per se and testcross performance. The majority of the 378 DH lines evaluated as testcrosses outperformed the commercial hybrids for early development. For total biomass yield, we observed a yield gap of 15% between mean testcross yield of the commercial hybrids and mean testcross yield of the DH lines. The DH lines also exhibited genetic variation for undesirable traits like root lodging and tillering, but correlations with target traits early development and yield were low or nonsignificant. The presented diversity atlas is a valuable, publicly available resource for genome-based studies to identify novel trait variation and evaluate the prospects of genomic prediction in landrace-derived material.

Introduction

Maize (*Zea mays* L. ssp. *mays*) seed banks around the world harbor thousands of landrace accessions, representing a rich resource of currently untapped native diversity that could be harnessed for plant improvement and adaptation to environmental changes (Hoisington et al. 1999; Ortiz et al. 2010; McCouch et al. 2013; Hellin et al. 2014; Wang et al. 2017). European flint maize went through several bottlenecks, the first of which occurred in the Americas (Doebley et al. 1986), followed by the introduction to Europe (Rebourg et al. 2003). In the course of maize breeding, landraces were replaced by hybrids. For the establishment of hybrid breeding, only a limited set of founder landraces was sampled, and the inbred lines produced were subjected to second cycle breeding (Messmer et al. 1992; Barrière et al. 2006). Subsequent selection at high intensities has led to an additional decline in genetic diversity of elite germplasm, especially within the flint heterotic pool important for European maize breeding (Messmer et al. 1992; Reif et al. 2005a, b; Lu et al. 2009). Revisiting the vast diversity of landraces stored in seed banks is considered a promising approach

Communicated by Jianbing Yan.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-019-03428-8) contains supplementary material, which is available to authorized users.

✉ Chris-Carolin Schön
chris.schoen@tum.de

[1] Plant Breeding, TUM School of Life Sciences Weihenstephan, Technical University of Munich, 85354 Freising, Germany

[2] Maize Breeding, KWS SAAT SE, 37574 Einbeck, Germany

[3] Misión Biológica de Galicia, Spanish National Research Council (CSIC), 36080 Pontevedra, Spain

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

[5] Present Address: Maize Breeding, KWS SAAT SE, 37574 Einbeck, Germany

for broadening the genetic base of current germplasm pools (Pollak 2003; Salhuana and Pollak 2006; Warburton et al. 2008; Strigens et al. 2013; McCouch et al. 2013; Navarro et al. 2017). However, opening this avenue for quantitative traits entails considerable challenges, and efficient strategies are still lacking.

In a first step, the most promising landraces have to be identified from several thousand stored in seed banks, even if only the flint pool is of interest. Information on stored landraces is limited, and the choice has to be based either on passport data from seed banks, or the per se and/or testcross performance of the landraces has to be evaluated in field trials (Pollak 2003; Salhuana and Pollak 2006; Böhm et al. 2014). In allogamous species like maize, landrace collections represent populations of heterogeneous and heterozygous individuals. Thus, the evaluation of populations either per se or in testcrosses would disregard the large genetic variation found within landraces, and without prior self or cross, it is not possible to evaluate the breeding potential of individual genotypes. In order to harness the genetic diversity within landraces, reproducible genetic units such as libraries of doubled-haploid (DH) lines from landraces have been suggested to overcome some of the aforementioned drawbacks since they are suitable for genotyping and high-precision phenotyping (Wilde et al. 2010; Strigens et al. 2013; Melchinger et al. 2017). Diversity from landraces captured in such DH libraries could help in improving traits such as plant development at early growth stages, for which genetic variation is small in breeding material. However, improving quantitative traits by utilizing lines derived from landraces is complex because the targeted introgression of favorable alleles at major genes is not possible (Bernardo 2002). Any introgression of landrace material therefore carries the risk of an undesired correlated response in traits other than the trait under selection due to the overall poor agronomic performance of the landrace material. To achieve a targeted utilization of natural diversity, an exhaustive characterization of line per se and testcross performance for the trait of interest and as many other agronomic and morphological traits as possible has to be carried out in order to develop a pre-breeding strategy that allows introgression of favorable diversity into elite germplasm without introducing major disadvantages in other traits (Sood et al. 2014).

In the research at hand, we employed large-scale production of DH lines to make native diversity for quantitative traits in maize landraces accessible for the purpose of germplasm improvement and genome-based studies. Our objectives were (i) to create a publicly available diversity atlas of European flint maize by characterizing landrace-derived DH libraries genotypically and phenotypically for line per se and testcross performance, (ii) to provide a comprehensive analysis of the DH libraries in terms of population structure, performance level, trait correlations, and genetic variances for a broad range of traits, and (iii) to gain insights into potential strategies for capturing native diversity for use in germplasm improvement.

Materials and methods

Plant material

The three landraces Kemater Landmais Gelb (KE, Austria), Petkuser Ferdinand Rot (PE, Germany), and Lalin (LL, Spain) were chosen for the production of DH lines because they showed phenotypic variation for early development as well as low levels of linkage disequilibrium (LD) and population structure within populations. They were selected from a set of 35 European maize landraces covering a broad geographical region of Europe that was described in detail by Mayer et al. (2017). Together, they represented 95.0% of the molecular variance of the full set of 35 landraces. From the selected landraces, 1015 DH lines (516 KE, 432 PE, 67 LL) were produced and multiplied using the in vivo haploid induction method (Röber et al. 2005). Phenotyping of lines per se (LP) was conducted in 2017 and 2018. Testcrosses (TC) of a subset of 378 DH lines from landraces KE and PE were evaluated in 2018. To warrant successful TC evaluation, the shortest, earliest, and late maturing lines as well as lines with a high score for lodging were not included in the TC production. The dent line F353 (Institut national de la recherche agronomique, INRA) was used as the female parent in TC production to ensure uniform seed quality across DH lines and because variation in tassel architecture of DH lines hampered detasseling.

Analysis of genotypic data and population structure

The 1015 DH lines and 144 S_0 plants (48 per landrace) from the landraces KE, PE, and LL were genotyped using the 600 k Affymetrix® Axiom® Maize Array (Unterseer et al. 2014). Only markers assigned to the best quality class (Unterseer et al. 2014), with a call rate of ≥ 0.9 and with a known physical position in the B73 reference sequence [AGPv4, (Jiao et al. 2017)], were used for the analyses. One S_0 plant from landrace PE was excluded due to an insufficient call rate (≤ 0.9). Assignment of lines to their respective landrace was performed using the ADMIXTURE software tool (Alexander et al. 2009) in supervised mode with three pre-defined groups (KE, PE, and LL) that were determined from S_0 plants. DH lines with less than 75% concordance with the landrace to which they were assigned by pedigree records were excluded from further analysis. Markers and individuals with > 10% missing values were removed. In DH lines, markers and individuals with > 5% heterozygous genotype calls were discarded, and all remaining heterozygous

calls were set to missing values. Missing values in the DH lines were imputed separately for each landrace using BEAGLE 5.0 (default parameters) (Browning et al. 2018). Missing values in the S_0 plants were imputed, and two gametes were phased from each S_0 plant separately in each landrace using BEAGLE 5.0 (iterations = 50, phase-segment = 10, phase-states = 500) and a reference panel consisting of the corresponding DH lines. Pairwise modified Rogers' distances [MRD; (Wright 1978)] were calculated, and DH lines showing a pairwise MRD of < 0.05 were assumed to be duplicates and excluded from further analyses. Markers were identified which overlapped between DH lines and S_0 gametes. Quality filtering and imputation resulted in 941 DH lines (501 KE, 409 PE, and 31 LL) and 286 S_0 gametes (96 KE, 94 PE, and 96 LL) genotyped with 499,574 common markers.

We performed a principal coordinate analysis [Gower (1966), R-package ape] based on MRD for DH lines and S_0 plants. The MRD matrices of DH lines and S_0 plants were hierarchically clustered using the unweighted pair group method with arithmetic mean (UPGMA) implemented in the hclust function in R and are displayed as 1-MRD. In order to estimate the proportion of molecular variance explained by the three landraces under study, an analysis of molecular variance [AMOVA; Excoffier et al. (1992)] was performed to partition the molecular variation into within- and between-landrace components. This analysis used the panel of 35 European landraces described by Mayer et al. (2017) for comparison. In addition, a second AMOVA decomposing the variance within and between DH lines and S_0 gametes was performed to investigate how much of the molecular variance lies within and between those two groups.

Field experiments and phenotypic analysis

Line per se (LP) performance was evaluated in Germany during 2017 using ten separate 10 × 10 lattice designs in four locations (1000 entries: 958 DH lines plus checks) and during 2018 using eight 10 × 10 lattice designs in three locations (800 entries: 756 DH lines plus checks). A randomly chosen subset (five 10 × 10 lattice designs, 458 and 468 DH lines plus checks in 2017 and 2018, respectively) was evaluated in two locations in Spain in both years. The trial locations were Einbeck (EIN, Germany, 2017 + 2018), Roggenstein (ROG, Germany, 2017 + 2018), Bernburg (BBG, Germany, 2017), Klein Wanzleben (KLW, Germany, 2018), Oberer Lindenhof (OLI, Germany, 2017), Golada (GOL, Spain, 2017 + 2018), and Tomeza (TOM, Spain, 2017 + 2018). See Table S1 for a detailed description of the test locations [geographical coordinates, elevation, precipitation, temperature; the climate data was obtained from the Bavarian State Research Center for Agriculture, Landwirtschaftliches Technologiezentrum Augustenberg, and Menne et al. (2012)]. Each combination of year and location was considered to be one environment in later analyses. The number of lines tested had to be reduced between 2017 and 2018 due to seed shortage and the exclusion of lines that did not pass the quality control described above for the genotypic data analysis. In 2017, 14 flint (CH10 provided by Agroscope Changins-Wädenswil (Switzerland); D152, DK105, UH006, UH007, and UH009 provided by the University of Hohenheim (Germany); EP1 and EP44 provided by Misión Biológica de Galicia, Consejo Superior de Investigaciones Científicas, (CSIC, Spain); F03802, F2, F283, F64, and F7 provided by Institut national de la recherche agronomique (INRA, France); EC49A provided by Centro de Investigaciones Agrarias Mabegondo, Instituto Galego da Calidade Aumentaria (CIAM-INGACAL, Spain) and one dent (F353, INRA, tester in testcross evaluation) inbred line served as checks and were included as duplicate entries. The checks were chosen in order to exhibit variation in plant development at early growth stages and flowering time. In 2018, the number of checks was reduced to four lines (DK105, EP1, F2, and F353) included in each lattice design per location (eight in Germany, five in Spain). In both years, the three landraces were included as quadruplicate entries. Plots were single rows 3 m in length with a distance of 0.75 m between rows and twenty plants per plot, corresponding to a sowing density of about 9 plants m^{-2}.

The testcrosses (TC) were evaluated in four 10 × 10 lattice designs in four locations in Germany in 2018 (EIN, KLW, ROG, OLI). In the TC trials, testcrosses of lines DK105, EP1, and F2 as well as testcrosses of the two landraces KE and PE and two commercial hybrid varieties (CH1 = KWS Stabil, CH2 = KWS Figaro) were planted as checks. The testcrosses of landraces KE and PE were planted in one lattice only, while all other checks were planted in every lattice. In TC, plots were double rows 5 m in length at locations ROG and OLI and 6 m in length at locations KLW and EIN, in both cases with 0.75 m distance between rows. Sowing density followed local practice at the experimental stations and varied between 9 and 11 plants m^{-2}. Fertilization and plant protection were carried out according to standard agricultural practices in both the LP and the TC trials.

In the LP trial, a total of 25 morphological, agronomic, and early-development-related traits were measured (Table S2 provides detailed information on trait × environment combinations). The traits that were scored in ≥ 10 environments included emergence (EME, ratio of emerged plants to sown seeds, %), early vigor (EV, at three different growth stages V3, V4, and V6, 1–9 score, 1 = very poor vigor, 9 = very vigorous), early plant height (PH, at V4 and V6, average over three measured plants per plot, cm), final plant height (PH_final, cm), and female flowering (FF, d). Root lodging at the R6 stage (RL, 1 = no lodging, 9 = all plants showing severe lodging) was scored in six environments;

tillering (TILL, 1 = no tillers, 9 = all plants showing many and long tillers) and male flowering (MF, d) were scored in five environments. The anthesis-silking interval (ASI, d) was calculated for the environments in which both MF and FF were scored. Ear height (EH, cm) was measured in four environments. In the Spanish environments, physiological traits like the maximum efficiency of photosystem II [Fv/Fm, using a fluorometer (OS-30p, Opti-Sciences Inc., USA)] were measured at stages V4 (2017 + 2018) and V6 (only 2017), and leaf greenness (SPAD) was measured by a chlorophyll content meter (CCM-200, Opti-Sciences Inc., USA; V3, V4 in both years, V6 only 2017). Reaction to stress was scored as cold tolerance (CT, 1–9 score, 1 = low cold tolerance, 9 = high cold tolerance; symptoms were chlorosis and necrosis on the leaves) after a very cold night with a slight frost at OLI 2017, drought/heat tolerance (DT, 1–9 score, 1 = low drought/heat tolerance, 9 = high drought/heat tolerance; symptoms were dry leaves and tassels) at EIN 2018, and rust susceptibility (binary) at TOM 2018. Traits related to tassel architecture were measured in ROG 2018. Tassel length was measured from the lowest tassel branch to the tassel tip (TL, cm), spike length was measured as the length of the top spike (SL, cm), the number of branches was counted (NB), and the tassel angle was scored on a 1–9 scale (TA, 1 = completely upright, 9 = branches horizontal). In the TC trial, EME, EV, PH, EH, PH_final, FF, TILL, and RL were scored as was described for LP. In addition, TC plots were harvested with a forage harvester to measure total dry matter yield (TDMY, dt/ha) and dry matter content (DMC, through near infrared spectroscopy or drying, in %).

The statistical model for estimating genotype and genotype × environment interaction variance components for lines derived from the same landrace was

$$y_{ijkopst} = \mu + m_i + \delta_{ij} l_j + g_{k(ij)} + u_o + \delta_{ij} lu_{jo} + gu_{ko(ij)} + k_{p(o)} + r_{s(op)} + b_{t(ops)} + \varepsilon_{ijkopst} \quad (1)$$

where $i = 1, 2, 3$ denotes three groups, i.e., DH lines from landraces (DHL), checks (CH), and landrace populations (LR_S$_0$); $j = 1, 2, 3$ denotes the three landraces KE, PE, and LL; μ is the overall mean; m_i is the effect of group i; l_j is the effect of landrace j in group $i = 1$; δ_{ij} is a dummy variable with $\delta_{ij} = 1$ for $i = 1$ and $j = 1, 2, 3$ and $\delta_{ij} = 0$ otherwise; $g_{k(ij)}$ is the effect of line k nested in group i and landrace j; u_o is the effect of environment o; lu_{jo} is the interaction of landrace j and environment o; $gu_{ko(ij)}$ is the interaction effect for genotype k and environment o. The effects $k_{p(o)}$, $r_{s(op)}$, $b_{t(ops)}$, and $\varepsilon_{ijkopst}$ refer to the effect of the lattice (nested in environments), replicate (nested in lattices in environments), incomplete block (nested in replicates in lattices in environments), and the residual error, respectively. All effects except m_i and l_j were treated as random. Genotype and genotype × environment ($gu_{ko(ij)}$) variance components were modeled individually for the three landraces ($j = 1, 2, 3$), assuming that DH lines across and within landraces were unrelated. Residuals were assumed to be normally distributed with mean zero and two heterogeneous variances, one for $\delta_{ij} = 1$ and one for $\delta_{ij} = 0$ assigning the same residual variance to all three landraces in all environments. Raw data and outliers were manually curated by inspection of residual plots. Since genotyping and the first year of phenotyping were carried out in parallel, some lines were evaluated in the field during 2017 that did not pass quality control in the genotypic data analysis. Measurements for those entries were treated as missing values in the data analysis. The same model was used for the analysis of TC experiments, except that $i = 1, 2$ referred to DHL and CH and $j = 1, 2$ referred to landraces KE and PE. Restricted maximum-likelihood estimation implemented in the ASReml-R package (Butler et al. 2009) was used for estimating variance components and their standard errors. Differences among means l_j were tested with pairwise t-tests using the R-package asremlPlus. Trait heritabilities were calculated on an entry-mean basis within landraces (Hallauer et al. 2010), and standard errors of heritability estimates were derived from standard errors of corresponding variance components using the delta method (Holland et al. 2010). Heritabilities and variance component estimates exceeding twice their standard errors were considered significant. Best linear unbiased estimates (BLUEs) of the genotype mean for each trait and DH line were obtained from a simplified version of the model in Eq. (1), dropping factors m_i, $\delta_{ij} l_j$ and $\delta_{ij} lu_{jo}$ and treating genotype (g_k) as a fixed effect. This model was also used to form linear contrasts used to test for significant differences (t-tests) between original landraces and the mean of the corresponding DH library (LP and TC) and between the mean of the two check hybrids and the mean of the DH library (TC only). We calculated the predicted response from selection within DH libraries (LP and TC) according to Falconer and Mackay (1996) as $\Delta G_{(\alpha)} = i_{(\alpha)} h \sigma_G$, where $i_{(\alpha)}$ = selection intensity for selection with $\alpha = 10\% \left(i_{(10\%)} \approx 1.76\right)$, h = square root of heritability, and σ_G = genetic standard deviation. To account for mean differences and different selection responses, we calculated the usefulness criterion (Schnell 1983) as $U_{(10\%)} = \bar{x} \pm \Delta G_{(10\%)}$ where $\bar{x}$ = mean of the respective DH library. Phenotypic correlations among traits were calculated from BLUEs as Pearson correlation coefficients within libraries in LP and TC, respectively. For evaluating the prospects of selection on LP performance in this material, we calculated Spearman rank correlations for same traits across LP and TC. To adjust for multiple testing, Bonferroni–Holm correction was applied for significance tests of phenotypic correlations in each DH library (Holm 1979). For estimating genetic covariances and genetic correlations between traits, the model in Eq. (1) was expanded to a bivariate model with pairs of traits. Genetic correlations were considered

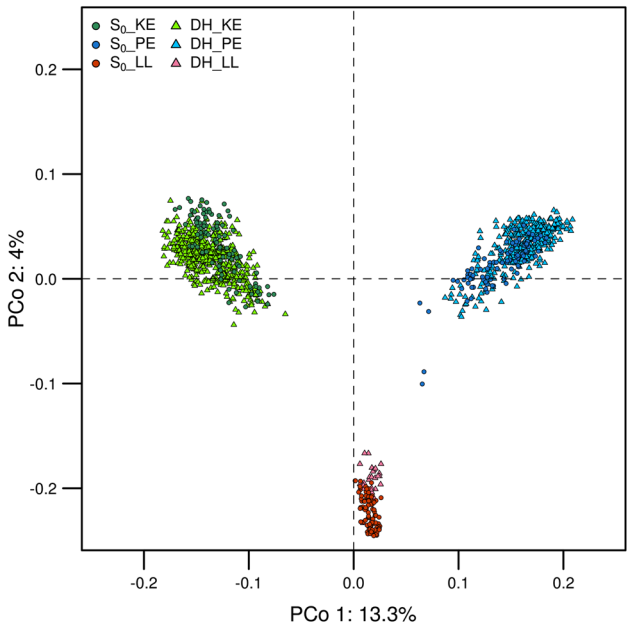


Fig. 1 Principal coordinate analysis (PCoA) of DH libraries and S_0 gametes based on modified Rogers' distances between individuals. Landrace KE is colored in *green*, PE in *blue*, and LL in *red*. Darker colors were used for S_0 gametes and brighter ones for DH. S_0 gametes were plotted as *filled circles* and DH lines as *filled triangles*. Axis labels show the percentage of explained variance per principal coordinate (PCo)

significant if they exceeded twice their standard error. The same method was applied for estimating genetic correlations between LP and TC performance.

In summary, high-quality phenotypic line per se data are available from up to 11 environments for 899 DH lines (471 KE, 402 PE, and 26 LL) and for a subset of 378 lines (190 KE, 188 PE) that were evaluated as testcrosses in four environments. For all lines, data on almost 500,000 SNP markers are available.

Results

Population structure and molecular variation

The principal coordinate analysis clearly separated the three landraces, with the first two coordinates explaining 13.3% and 4% of the total molecular variance, respectively (Fig. 1). DH lines and S_0 gametes derived from the same landrace clustered together except for four gametes from S_0_PE, which fell outside the PE-cluster. Complementing our data with those from Mayer et al. (2017) revealed that S_0 gametes sampled from landraces KE, PE, and LL individually accounted for 77, 75, and 89% of the total molecular variance captured in the collection of 35 European landraces used in their study. The AMOVA on S_0 gametes and DH lines from the same landrace confirmed the results from the PCoA. While 95.3, 96.6, and 96.7% of the molecular variance were found within S_0 and DH of KE, PE, and LL, respectively, less than 5% of the molecular variance was explained by differences between S_0 gametes and DH lines of different landraces. Matrices of 1-MRD

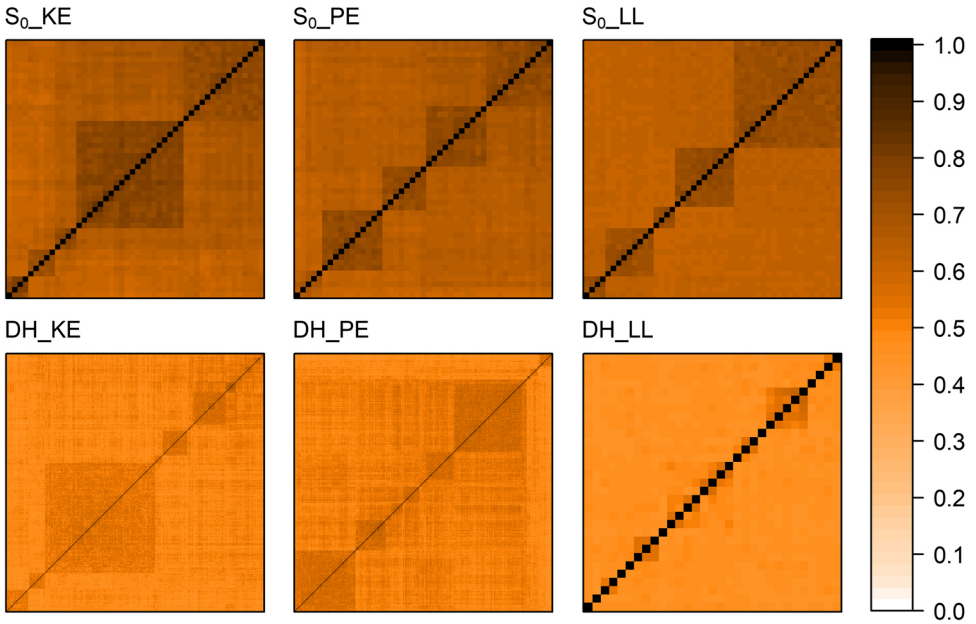


Fig. 2 Heatmaps of 1-MRD matrices S_0_KE ($N=48$ individuals), DH_KE ($N=471$ lines), S_0_PE ($N=47$ individuals), DH_PE ($N=409$ lines), S_0_LL ($N=48$ individuals), and DH_LL ($N=31$ lines). Matrices were ordered according to hierarchical clustering with UPGMA

(Fig. 2) gave no indication of pronounced population structure for either DH or S_0 plants. As expected, the similarity between S_0 plants within landraces was on average higher than in DH lines due to the higher level of heterozygosity in the former.

Phenotypic variation within and across landraces

In the following, we will refer to a subset of traits as "core traits" since they are considered most important for improvement of early plant development in elite germplasm. These traits were EV_V4 and PH_V4 as representatives for early development, RL and TILL as representatives for traits for which genetic variation is not acceptable in elite germplasm, PH_final and FF as important agronomic traits, and DMC and TDMY for evaluating yield performance. Phenotypic variation for core traits within and across landraces is shown in Fig. 3 (LP) and Fig. 4 (TC) and for all other traits in Fig. S1 (LP) and Fig. S2 (TC). Phenotypic means, variance components, and heritabilities for all traits are provided in Table S3 and Table S4 for LP and TC performance, respectively. The DH libraries exhibited considerable phenotypic variation for all traits. In LP and TC, a similar range of trait values was observed for all DH libraries. Probably due to the small sample size, distribution of phenotypes in LL deviated slightly from the other two landraces, e.g., for traits EV and TILL. Mean performance differed significantly ($P < 0.05$) across landraces for 20 out of 25 traits in LP and for 5 out of 14 traits in TC, which was a result of the high-quality phenotypic data and large sample sizes of KE and PE. As expected, mean LP performance of the DH libraries was significantly ($P < 0.05$) lower than the respective landraces for almost all traits. The reduction was most pronounced for early development traits, final plant height, and photosynthetic efficiency (Fig. 3, Fig. S1). Flowering time of the DH library was delayed by 10 (LL) and 6 (KE, PE) days compared to the non-inbred material. While the LL DH library had consistently lower mean performance in early development traits, ear height, and final plant height compared to KE and PE, this was not true for the original landraces.

When choosing DH lines to be evaluated as TC, we had applied mild selection for flowering time, plant height, and lodging (see "Materials and methods" for details). Mean TC performance of the DH libraries KE and PE did not differ significantly from the TC mean of their respective landrace populations for all traits except for TDMY in PE, indicating that DH lines evaluated as TC represented a random sample

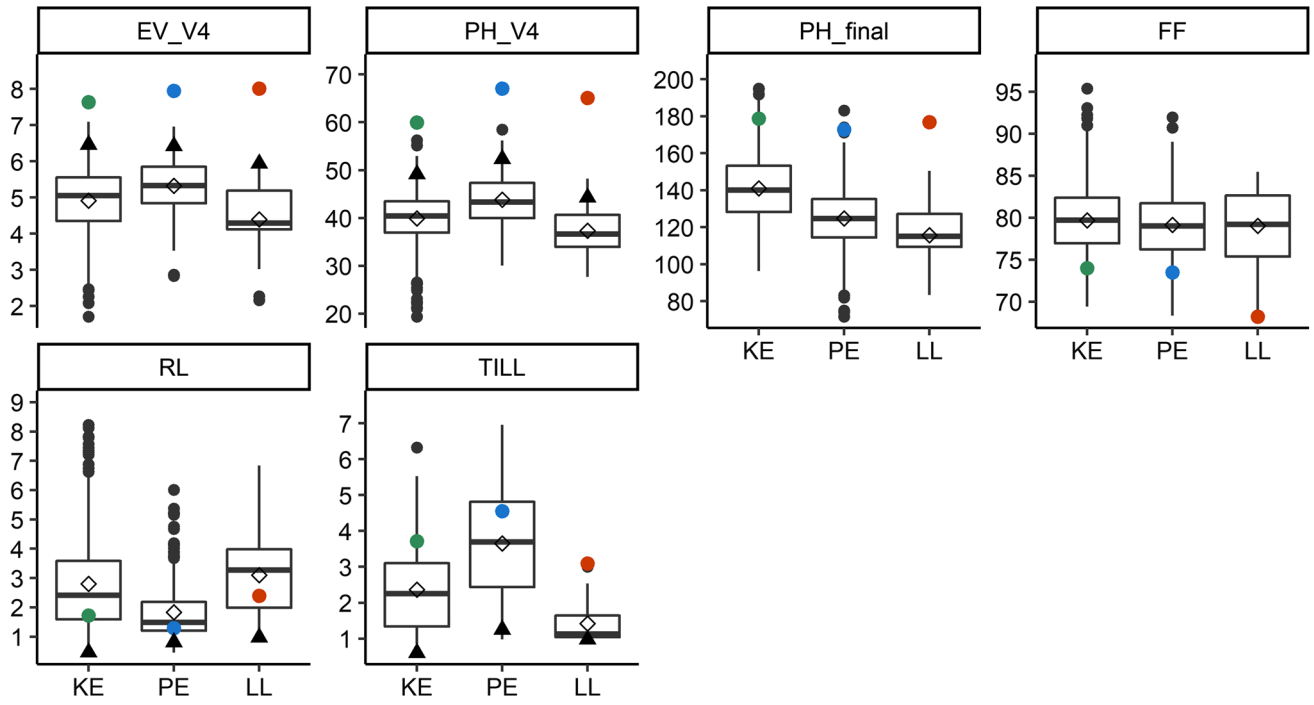


Fig. 3 Boxplots of phenotypic data for line per se (LP) performance for the DH libraries from landraces KE, PE, and LL. Boxplots show the upper and lower quartiles, median (*horizontal bar*), mean (*open diamond*), whiskers (*vertical bars*), and the performance of the respective landrace (*filled circle* in *green*, *blue*, and *red* for KE, PE, and LL, respectively). Points above and below the whiskers indicate values ± 1.5 times the interquartile range. Usefulness for a selection intensity of 10% ($U_{10\,\%}$) is indicated with *black filled triangles*. Traits are early vigor and early plant height at stage V4 (EV_V4, PH_V4), final plant height (PH_final), female flowering (FF), root lodging (RL), and tillering (TILL)

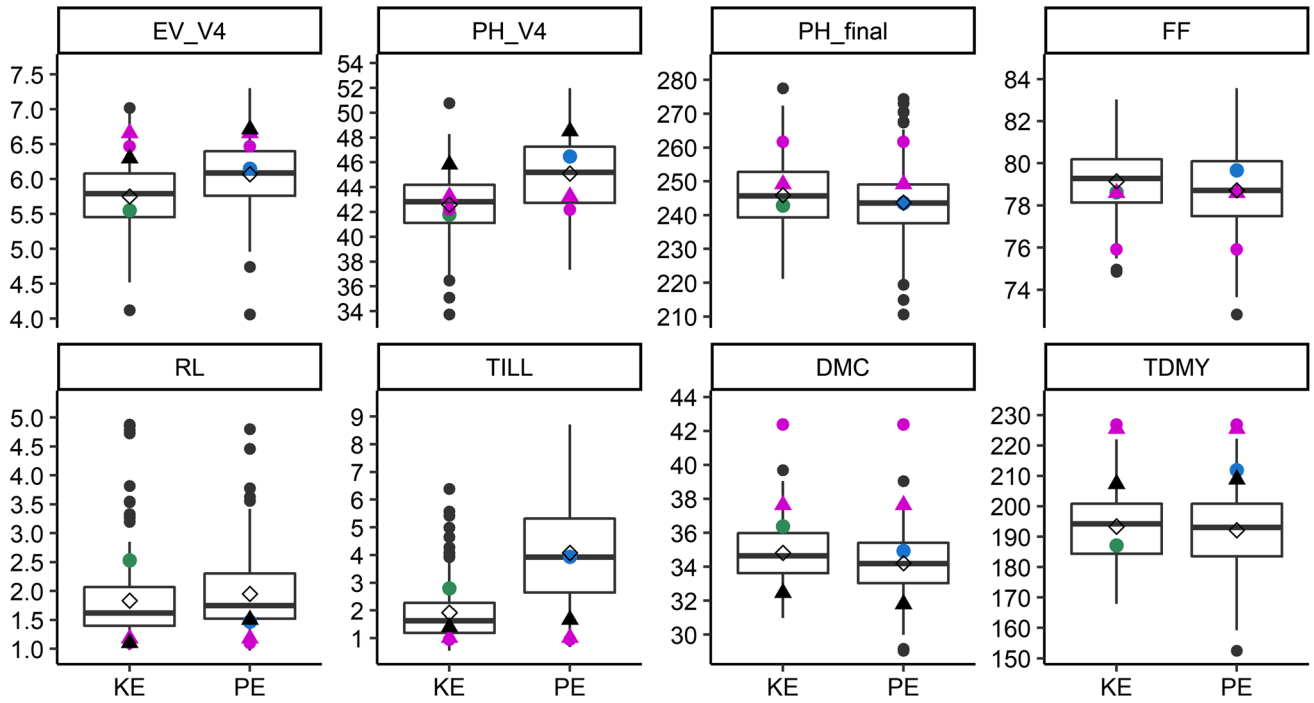


Fig. 4 Boxplots of phenotypic data for testcross (TC) performance for DH libraries from landraces KE and PE. Boxplots show the upper and lower quartiles, median (*horizontal bar*), mean (*open diamond*), whiskers (*vertical bars*) and the performance of the respective landrace (*filled circle* in *green* and *blue* for KE and PE, respectively). Points above and below the whiskers indicate values ± 1.5 times the interquartile range. Performance of the two commercial check hybrids is indicated with a *filled circle* and *filled triangle* in *magenta* for CH1 and CH2, respectively. Usefulness for a selection intensity of 10% ($U_{10\ \%}$) is indicated with *black filled triangles*. Traits are early vigor and early plant height at V4 stage (EV_V4, PH_V4), final plant height (PH_final), female flowering (FF), root lodging (RL), tillering (TILL), dry matter content (DMC), and total dry matter yield (TDMY)

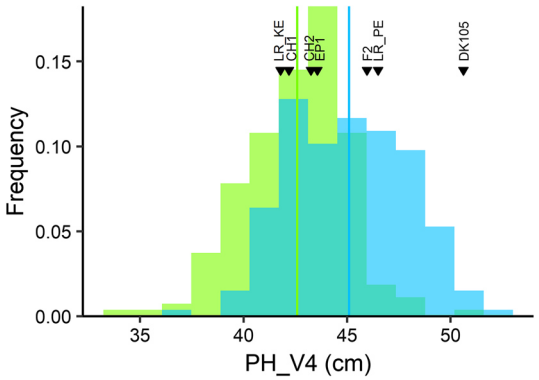


Fig. 5 Histogram of testcross (TC) performance of DH lines from landraces KE ($N = 190$, colored in *green*) and PE ($N = 188$, colored in *blue*) for trait early plant height at V4 stage (PH_V4), including the mean of the DH lines per population (*green* and *blue horizontal bar*) and the performance of testcrosses of lines EP1, F2, DK105, the landrace populations (LR_KE, LR_PE), as well as two commercial check hybrids (CH1, CH2) indicated by labeled *black triangles*

of the entire DH library. The TC of many DH lines outperformed the commercial hybrids as well as the TC of founder lines and landraces for the target trait early development, as is shown for PH_V4 in Fig. 5. Only the testcross of inbred DK105 fell into the upper 10% of the distribution of PH_V4. As expected, the commercial hybrids significantly ($P < 0.05$) outperformed the TC mean of DH lines for TDMY by about 15% and, in contrast to the DH lines, showed no TILL or RL (Fig. 4).

Genetic variances were highly significant in LP and TC for most traits under investigation (Table S3, Table S4). Variance component estimates for LL were similar to the other two libraries, but, due to the small sample size, they were estimated with considerably larger error, resulting in nonsignificant genetic variances for PH_V3, TILL, ASI, photosynthesis-related traits, and SPAD. As expected from quantitative genetic theory, genetic variance component estimates were smaller in TC than in LP. In the statistical model, we allowed for heterogeneity of genetic variances estimated within landraces, but only a few traits (e.g., DT, RL) showed strong differences (> twofold) in genetic variance estimates between KE and PE in LP, which were even alleviated in TC.

In LP, trait heritabilities were generally high and similar across landraces, ranging from 0.35 to 0.96. Except for PH_V3, TILL, ASI, Fv/Fm, and SPAD in LL, the

heritability estimate always exceeded twice its standard error (Table S3). In TC, heritabilities were slightly lower overall than in LP (Table S4), ranging from 0.31 to 0.92, which was expected from the lower number of testing environments and the lower genetic variance compared to LP.

Variation across environments

DH libraries were evaluated in a total of 11 environments comprising seven different locations and two years. Locations covered a geographical region spanning from northern Germany to northwestern Spain at altitudes ranging from 29 to 706 m above sea level (Table S1). Average temperatures differed by 5 °C between the coldest (OLI 2017, 14.0 °C) and the warmest (TOM 2018, 19.0 °C) environments, and precipitation varied from 159 (KLW 2018) to 548 mm (ROG 2018) during the vegetation period. The ratio of genotype by environment and genotype variance components depended on the trait under study. In LP, values ranged from 0.11 (EH in KE) to 1.22 (ASI in PE), but varied between 0.2 and 0.7 for most traits with a mean of 0.51 (Table S3). Similar ratios were observed in TC (Table S4).

Correlations between locations for traits measured in at least five environments ranged from 0.40 to 0.87 in 2017 and from 0.19 to 0.86 in 2018 (Table S6). Correlations between years of a given trait and location ranged from 0.31 to 0.83 (Table S6).

Trait correlations

In LP and TC, phenotypic correlations among early development traits measured at different growth stages were high and stable across DH libraries, ranging from 0.58 to 0.95 (Fig. 6). The corresponding genetic correlations were slightly higher, ranging from 0.65 to 1 (Fig. S3). For LL in LP, only phenotypic correlations among early development traits (ranging from 0.82 to 0.93, data not shown), PH_final and EH (0.75), and FF and MF (0.69) were significant.

In LP, the early development traits showed intermediate to high positive phenotypic and genetic correlations with final plant height (phenotypic 0.4 to 0.6, genetic 0.4 to 0.7). In TC, only the phenotypic correlation between PH_V6 and

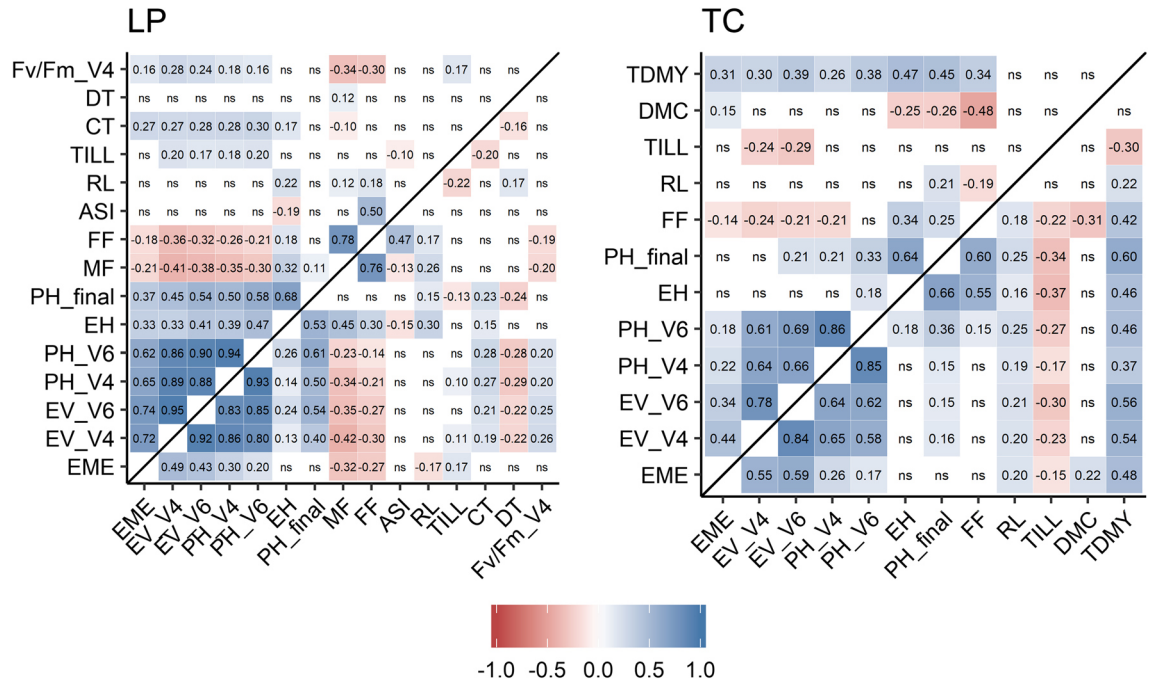


Fig. 6 Phenotypic Pearson correlation coefficients for line per se [LP, *left, N* = 471 (KE) and 402 (PE)] and testcross [TC, *right*, *N* = 190 (KE) and 188 (PE)] data within DH libraries KE (above diagonal) and PE (below diagonal) for the traits emergence (EME), early vigor, and early plant height at stages V4 and V6 (EV_V4, EV_V6, PH_V4, PH_V6), ear height (EH), final plant height (PH_final), male flowering and female flowering (MF, FF), anthesis-silking interval (ASI), root lodging (RL), tillering (TILL), cold tolerance (CT), drought/heat tolerance (DT), maximum photosynthetic efficiency at V4 stage (Fv/Fm_V4), dry matter content (DMC), and total dry matter yield (TDMY). *P*-values were adjusted using Bonferroni–Holm correction for multiple testing. Nonsignificant correlations are labeled with ns

final plant height was significant, but it was lower than in LP. Genetic correlations for EV_V4, EV_V6, PH_V4, and PH_V6 with PH_final ranged between 0.2 and 0.5. Intermediate positive correlations were found between early development traits and TDMY in TC (phenotypic 0.3 to 0.6, genetic 0.4 to 0.8) as well as negative correlations of early development with flowering time in LP (FF and MF, phenotypic −0.2 to −0.4, genetic −0.2 to −0.5) in KE and PE libraries. Phenotypic and genetic correlations of RL with all other traits were nonsignificant or small in LP and TC. The same was true for TILL except for TC of PE, where intermediate correlations with early and late plant height, ear height, and TDMY were observed.

Phenotypic correlations between LP and TC performance were significant for all traits except EME. Genetic correlations between LP and TC were intermediate (early development traits, 0.35 to 0.68) to high (PH_final and FF > 0.78) (Table S5).

Discussion

Our study is part of a long-term research project which aims to make maize landrace diversity amenable to plant breeding (www.europeanmaize.net). We produced DH libraries from three landraces for obtaining reproducible genetic units for phenotyping and genotyping and characterized them comprehensively to build a publicly available, immortal genetic resource that is ready to use for pre-breeding and for investigations on functional diversity and the prospects of genomic prediction.

DH libraries capture native diversity for germplasm improvement

The three landraces were chosen to represent the molecular variance of the European landraces characterized by Mayer et al. (2017). Individually, they accounted for more than 75% of the molecular variance in this collection, together for 95%. These findings corroborate results from the literature where it has been shown for several outcrossing species, including maize, that a large proportion of the molecular variation can be found within landraces, while differences between landraces account only for a small proportion (Böhm et al. 2014; Greene et al. 2014; Monteiro et al. 2016). Genotyping with the SNP array technology might have led to an overestimation of the captured molecular variance due to an enrichment of markers with intermediate allele frequencies. For truly quantitative traits, however, the contribution of rare alleles to the additive genetic variance is small and the molecular variance assessed with array data should translate directly into genetic variation observable in phenotypes. With only three (LP) or two (TC) landraces in the statistical model, decomposition of the genetic variance within and across landraces is not meaningful, but from Figs. 3, 4 and Figs. S1, S2 it becomes obvious that differences in trait means across landraces were small compared to the range of values within landraces. Although each landrace accounted for a large proportion of molecular variance individually, we still advise to analyze progenies from several landraces for capturing the genetic variance segregating in a germplasm pool. Molecular variance might be a good indicator for genetic variance averaged across traits, but variation for individual traits must be evaluated for each landrace specifically, as was shown here for TILL, RL, DT, and CT. Different landraces may also differ with respect to their success rates in DH production (Melchinger et al. 2017), pointing to different multiplication histories. While KE and PE may have encountered bottlenecks or inbreeding in the past, LL seems to carry a much higher genetic load that limited the production of fully homozygous DH lines for this landrace. This assumption is also supported by the significantly lower LP mean performance of the LL DH library for early development, ear height, and final plant height compared to KE and PE that was not observed for the original landraces.

The DH libraries generated in this study represented their respective landraces accurately in terms of molecular variance. DH lines and S_0 gametes from the same landrace overlapped nicely in the PCoA (Fig. 1) and the AMOVA showed that almost all molecular variation was found within S_0 gametes and DH lines (> 95%) and not between them. Individuals sampled from a maize landrace are assumed to be unrelated, but pairwise comparisons share different numbers of alleles alike in state, leading to variation in similarity between them. Patterns of variation in similarity were comparable for S_0 plants and DH lines (Fig. 2), corroborating that the two types of progeny represent their original landraces in a similar way. We thus conclude that the three DH libraries derived from KE, PE, and LL represent a valuable resource for genetic improvement of elite flint germplasm, since they cover a large proportion of the genomic and genetic variance of the landrace collection described in Mayer et al. (2017).

Improving early plant development

In many growing regions worldwide, maize encounters low to moderate temperatures during the early vegetative phase. Under these conditions, accelerated early development can increase final biomass yield. Genetic enhancement of early growth can also improve resource efficiency, preserve soil fertility, and reduce the need for herbicide treatment. European flint maize germplasm has been adapted to the temperate climate conditions of Northern and Central Europe through breeding, but genetic variation for early

development under cool temperatures has been depleted simultaneously (Greaves 1996; Rodríguez et al. 2010).

In LP and TC of the DH libraries, the target trait early development assessed through early vigor scores and early plant height measurements showed ample genetic variation (Figs. 3, 4, Figs. S1, S2). In TC, the majority of DH lines outperformed the commercial hybrids for PH_V4, and only one check (F353 × DK105) ranged among the best 10% DH lines, suggesting that the DH libraries can serve as a valuable source of alleles for improving early development traits of the elite European flint pool (Fig. 5).

Identifying maize flint germplasm with superior early growth has been the objective of several studies in both, field and controlled environments (Peter et al. 2009a, b; Rodríguez et al. 2010; Revilla et al. 2016). In most studies, early development was assessed as a visual score, which delivers ordinal endpoints and can be rather subjective. On the other hand, early plant height measurements consume considerable resources. Early vigor scores showed a substantially higher correlation with plant emergence compared to early plant height in this research (Fig. 6, Fig. S3). Even though all TC seed was produced on inbred line F353, the higher phenotypic and genetic correlation of early vigor and EME was maintained. For PH_V4, the commercial hybrids were on average not different from the TC mean of the DH libraries, but they scored better for EV. Thus, the early plant height measurement neglects information that can be accounted for by EV scores, such as differences in leaf coloration or the overall lower EME of the DH library testcrosses. In addition, genetic correlations between TDMY and EV were substantially higher compared to between TDMY and PH_V4 supporting the hypothesis that, although highly correlated, the two types of measurements target different components of early development. For a comprehensive characterization of early growth development, it seems advisable to assess both, EV and early plant height. To allow dissection of early growth development into its genetic components and consequently provide a better understanding of the underlying genetic mechanisms, we propose establishing growth models by monitoring early development at high resolution in time using remote sensing in the field (Huang et al. 2013; Bendig et al. 2015) or in controlled conditions (Gioia et al. 2017). The three DH libraries KE, PE, and LL are most suitable for further investigation on this topic as they exhibit more pronounced genetic variation in early growth traits than can be expected from elite material (Revilla et al. 1999; Peter et al. 2009a).

Comprehensive phenotypic characterization of DH libraries

The prospects for the genetic improvement of elite germplasm for early growth development through the use of landrace-derived material have to be evaluated in a multi-trait context. Comprehensive data on trait correlations are crucial in order to avoid undesired selection response in traits of agronomic importance.

In LP, EV_V4 and PH_V4 showed intermediate negative genetic correlations with flowering time and positive genetic correlations with PH_final, corroborating results of Böhm et al. (2017) on DH lines derived from landraces. Thus, selection for accelerated early development will lead to increased plant height and early flowering which, depending on the target environment, might not be desirable. The DH libraries also showed variation for RL and TILL. Given the low levels of genetic correlations with early development traits and the usefulness of the best 10% of DH lines being close to zero, a simultaneous reduction or removal of lodging and tillering should be possible in a recurrent selection program devoted to the improvement of early development traits. In TC, correlations between early development traits and TDMY were positive. However, the commercial hybrids significantly outperformed the DH lines for TDMY, while testcrosses of founder lines (F2, EP1, DK105) lay well within the range of the DH libraries for both traits (Fig. 4, Fig. S4). The yield gap between the mean testcross yield of the DH lines and the mean testcross yield of two commercial hybrids amounted to about 15% and was comparable to what was reported in the literature for other European landraces (Wilde et al. 2010; Brauner et al. 2019). The usefulness of the best 10% DH lines in KE and PE, respectively, remained 8% below the performance level of the commercial hybrids for TDMY (Fig. 4). Given that the inbred line F353 used as tester for the DH libraries was developed about 20 years ago (year of release 2001, C. Bauland, personal communication) and that the parental components of the commercial check hybrids were highly selected based on their general and specific combining ability, the difference in TDMY between commercial hybrids and the top 10% DH lines seems small and could likely be reduced by the use of modern testers (Hölker et al. 2019). In many material groups, a negative correlation between DMC and TDMY is expected. In our research, phenotypic correlations between TDMY and DMC were nonsignificant when averaged across environments (Fig. 6, Fig. S4) as well as in all four individual environments where TC performance was evaluated (data not shown). This outcome can most likely be attributed to the exceptionally hot and dry conditions during the 2018 growing season (Table S1), the genetic material under study, or an interaction of both. Thus, an additional year of TC evaluation, including more and also later maturing commercial check hybrids, will be conducted for investigating the DMC/TDMY relationship in material derived from genetic resources more closely and for evaluating the overall yield potential of the DH libraries.

Multi-environment testing

One of the aims of this study was to assess trait differentiation in diverse environments and to estimate the magnitude of genotype × environment interactions of landrace-derived material. Thus, the chosen environments covered a broad spectrum of target regions for European flint material (Table S1). Despite locations with very different climatic conditions (e.g., OLI and TOM) and large differences in temperature and precipitation in 2017 and 2018, the ratio of genotype × environment and genetic variance (σ^2_{gu} : σ^2_g, Tables S3, S4) was moderate for most traits. If landraces from which DH libraries are derived are adapted to similar environmental conditions as the target elite breeding germplasm, the confounding effects of adaptive alleles and strong genotype × environment interactions can be avoided and meaningful phenotypes obtained. Thus, our results are encouraging with respect to the prospects of incorporating environmentally stable alleles from pre-selected DH libraries into elite germplasm.

Evaluating landrace-derived material in 11 environments might not be practicable for applied pre-breeding programs. In this study, the large number of test environments was highly useful because we detected the segregation of unfavorable alleles in specific environments such as segregation for rust in TOM (Fig. S5) and drought susceptibility in EIN (Fig. S6), both in 2018. Although infections with rust or severe drought may not occur frequently, it would be devastating if these susceptibilities were transferred inadvertently to elite germplasm through the introgression of landrace-derived material. If evaluating the landrace-derived material in a large number of environments is not possible, prioritized testing in environments known for high disease pressure, abiotic stress, or frequent occurrence of undesirable traits like RL is highly advisable.

DH libraries from landraces make native diversity accessible

The DH libraries presented in this study link the large molecular diversity present in landraces to meaningful phenotypes. DH lines from landraces outperformed flint founder lines and commercial hybrids in early development, and as immortal genetic units they are directly accessible for plant breeding. Improving one or several target traits and simultaneously closing the performance gap between elite and landrace-derived genetic material for multiple traits of agronomic importance requires efficient recurrent population improvement. In this context, knowledge of trait correlations is crucial in order to broaden the narrow genetic base of the elite flint germplasm pool without introducing undesired traits from landraces into elite breeding populations. To obtain maximum selection gain per unit time, theory offers different strategies, such as multi-stage or index selection (Bernardo 2002), which need to be evaluated in the framework of the respective breeding programs. Optimal strategies may vary conditional on species, budget, and short-term or long-term perspectives. Böhm et al. (2017) suggested multi-stage phenotypic selection of landrace-derived DH libraries. In a simulation study, Gorjanc et al. (2016) compared different scenarios for initiating pre-breeding for maize landraces using genomic prediction (GP) and suggested starting directly from landraces (e.g., without crossing to elite lines).

The implementation of GP in pre-breeding of landrace-derived material is still underexploited. The comprehensive phenotypic data and derived quantitative genetic parameters presented for the three DH libraries in this study provide an excellent basis for optimizing genome-based pre-breeding schemes. Multi-environment phenotypic data are available for model training in LP and TC. Sample sizes and marker densities are large, allowing to investigate the effects of population size and required marker densities in populations with relatively low linkage disequilibrium compared to elite germplasm. In addition to investigating the prospects of genome-based prediction, our data provide a comprehensive framework for the discovery of genes controlling favorable and unfavorable traits as well as for the genetic analysis of additional relevant traits such as nutrient efficiency, photosynthesis-related traits, and additional biotic and abiotic stress tolerances.

Acknowledgements We are grateful to Barbara Eder and Joachim Eder (Bavarian State Research Center, Institute for Crop Science and Plant Breeding, Freising, Germany), Alain Charcosset and Cyril Bauland (INRA, Gif-Sur-Yvette, France), Anne Zanetto (INRA, Mauguio, France), the INRA station of Saint Martin de Hinx, Pedro Revilla and Rosa Ana Malvar (CSIC, Pontevedra, Spain), Angel Alvarez (CSIC, Zaragoza, Spain), and Jesús Moreno González (CIAM, A Coruña, Spain) for providing landrace and/or elite line seeds. We are indebted to the technical staff at KWS SAAT SE, Misión Biológica de Galicia, Spanish National Research Council (CSIC), Technical University of Munich, and University of Hohenheim for their skilled support in carrying out the extensive phenotypic evaluation required for this study. This study was funded by the Federal Ministry of Education and Research (BMBF, Germany) within the scope of the funding initiative "Plant Breeding Research for the Bioeconomy" (Funding ID: 031B0195, project "MAZE") and by the KWS SAAT SE under Ph.D. fellowships for Manfred Mayer and Armin C. Hölker.

Author contribution statement CCS, EB, MO, TP, and AEM conceived the study; CCS, MO, and AEM acquired funding for the study; MO, TP, and TB developed the plant material; MM investigated genotypic data and performed analyses; ACH, MM, EB, TB, TP, BO, and PCB collected the phenotypic data; ACH performed phenotypic data analysis and drafted the manuscript; CCS edited the manuscript; all authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical standards The authors declare that this study complies with the current laws of the countries in which the experiments were performed.

Availability of data and materials All data and material are available through material transfer agreements upon request.

References

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655–1664. https://doi.org/10.1101/gr.094052.109

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

Bendig J, Yu K, Aasen H, Bolten A, Bennertz S, Broscheit J, Gnyp ML, Bareth G (2015) Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. Int J Appl Earth Obs Geoinf 39:79–87. https://doi.org/10.1016/j.jag.2015.02.012

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

Böhm J, Schipprack W, Mirdita V, Utz HF, Melchinger AE (2014) Breeding potential of european flint maize landraces evaluated by their testcross performance. Crop Sci 54:1665–1672. https://doi.org/10.2135/cropsci2013.12.0837

Böhm J, Schipprack W, Utz HF, Melchinger AE (2017) Tapping the genetic diversity of landraces in allogamous crops with doubled haploid lines: a case study from European flint maize. Theor Appl Genet 130:861–873. https://doi.org/10.1007/s00122-017-2856-x

Brauner PC, Schipprack W, Utz HF, Bauer E, Mayer M, Schön CC, Melchinger AE (2019) Testcross performance of doubled haploid lines from European flint maize landraces is promising for broadening the genetic base of elite germplasm. Theor Appl Genet. https://doi.org/10.1007/s00122-019-03325-0

Browning BL, Zhou Y, Browning SR (2018) A one-penny imputed genome from next-generation reference panels. Am J Hum Genet 103:338–348. https://doi.org/10.1016/j.ajhg.2018.07.015

Butler D, Cullis BR, Gilmour A, Gogel B (2009) ASReml-R reference manual. The State of Queensland, Department of Primary Industries and Fisheries, Brisbane

Doebley JF, Goodman M, Stuber CW (1986) Exceptional genetic divergence of northern flint corn. Am J Bot 73:64–69. https://doi.org/10.1002/j.1537-2197.1986.tb09681.x

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Falconer D, Mackay T (1996) Introduction to quantitative genetics, 4th edn. Longmann Scientific & Technical, New York

Gioia T, Galinski A, Lenz H, Müller C, Lentz J, Heinz K, Briese C, Putz A, Fiorani F, Watt M, Schurr U, Nagel KA (2017) GrowScreen-PaGe, a non-invasive, high-throughput phenotyping system based on germination paper to quantify crop phenotypic diversity and plasticity of root traits under varying nutrient supply. Funct Plant Biol 44:76–93. https://doi.org/10.1071/FP16128

Gorjanc G, Jenko J, Hearne SJ, Hickey JM (2016) Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. BMC Genom 17:30. https://doi.org/10.1186/s12864-015-2345-z

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338. https://doi.org/10.2307/2333639

Greaves JA (1996) Improving suboptimal temperature tolerance in maize—the search for variation. J Exp Bot 47:307–323. https://doi.org/10.1093/jxb/47.3.307

Greene SL, Kisha TJ, Yu L-X, Parra-Quijano M (2014) Conserving plants in gene banks and nature: investigating complementarity with *Trifolium thompsonii* morton. PLoS ONE. https://doi.org/10.1371/journal.pone.0105145

Hallauer AR, Carena MJ, de Miranda Filho J (2010) Quantitative genetics in maize breeding. Springer, Berlin

Hellin J, Bellon MR, Hearne SJ (2014) Maize landraces and adaptation to climate change in Mexico. J Crop Improv 28:484–501. https://doi.org/10.1080/15427528.2014.921800

Hoisington D, Khairallah M, Reeves T, Ribaut JM, Skovmand B, Taba S, Warburton M (1999) Plant genetic resources: what can they contribute toward increased crop productivity? Proc Natl Acad Sci 96:5937–5943

Hölker AC, Schipprack W, Utz HF, Molenaar WS, Melchinger AE (2019) Progress for testcross performance within the flint heterotic pool of a public maize breeding program since the onset of hybrid breeding. Euphytica 215:50. https://doi.org/10.1007/s10681-019-2370-0

Holland JB, Nyquist WE, Cervantes-Martínez CT (2010) Estimating and interpreting heritability for plant breeding: an update. In: Janick J (ed) Plant breeding reviews. Wiley-Blackwell, Hoboken, pp 9–112

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70

Huang Y, Thomson SJ, Hoffmann WC, Lan Y, Fritz BK (2013) Development and prospect of unmanned aerial vehicle technologies for agricultural production management. Int J Agric Biol Eng 6:1–10. https://doi.org/10.25165/ijabe.v6i3.900

Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, Campbell MS, Stein JC, Wei X, Chin CS, Guill K, Regulski M, Kumari S, Olson A, Gent J, Schneider KL, Wolfgruber TK, May MR, Springer NM, Antoniou E, McCombie WR, Presting GG, McMullen M, Ross-Ibarra J, Dawe RK, Hastie A, Rank DR, Ware D (2017) Improved maize reference genome with single-molecule technologies. Nature 546:524–527. https://doi.org/10.1038/nature22971

Lu Y, Yan J, Guimarães CT, Taba S, Hao Z, Gao S, Chen S, Li J, Zhang S, Vivek BS, Magorokosho C, Mugo S, Makumbi D, Parentoni SN, Shah T, Rong T, Crouch JH, Xu Y (2009) Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. Theor Appl Genet 120:93–115. https://doi.org/10.1007/s00122-009-1162-7

Mayer M, Unterseer S, Bauer E, de Leon N, Ordas B, Schön CC (2017) Is there an optimum level of diversity in utilization of genetic resources? Theor Appl Genet 130:2283–2295. https://doi.org/10.1007/s00122-017-2959-4

McCouch S, Baute GJ, Bradeen J, Bretting PK, Buckler E, Burke JM, Charest D, Cloutier S, Cole G, Dempewolf H, Dingkuhn M, Feuillet C, Gepts P, Grattapaglia D, Guarino L, Jackson S, Knapp S, Langridge P, Lawton-Rauh A, Lijua Q, Lusty C, Michael T, Myles

S, Naito K, Nelson RL, Pontarollo R, Richards CM, Rieseberg L, Ross-Ibarra J, Rounsley S, Hamilton RS, Schurr U, Stein N, Tomooka N, van der Knaap E, van Tassel D, Toll J, Valls J, Varshney RK, Ward J, Waugh R, Wenzl P, Zamir D (2013) Agriculture: feeding the future. In: Nature. https://www.nature.com/articles/499023a. Retrieved 27 Jul 2018

Melchinger AE, Schopp P, Müller D, Schrag TA, Bauer E, Unterseer S, Homann L, Schipprack W, Schön CC (2017) Safeguarding our genetic resources with libraries of doubled-haploid lines. Genetics. https://doi.org/10.1534/genetics.115.186205

Menne MJ, Durre I, Korzeniewski B, McNeill S, Thomas K, Yin X, Anthony S, Ray R, Vose RS, Gleason BE, Houston TG (2012) Global historical climatology network—daily (GHCN-daily), version 3. NOAA Natl Cent Environ Inf. https://doi.org/10.7289/V5D21VHZ

Messmer MM, Melchinger AE, Boppenmaier J, Brunklaus-Jung E, Herrmann RG (1992) Relationships among early European maize inbreds: I. Genetic diversity among flint and dent lines revealed by RFLPs. Crop Sci 32:1301–1309

Monteiro F, Vidigal P, Barros AB, Monteiro A, Oliveira HR, Viegas W (2016) Genetic distinctiveness of rye in situ accessions from Portugal unveils a new hotspot of unexplored genetic resources. Front Plant Sci. https://doi.org/10.3389/fpls.2016.01334

Navarro JAR, Willcox M, Burgueño J, Romay C, Swarts K, Trachsel S, Preciado E, Terron A, Delgado HV, Vidal V, Ortega A, Banda AE, Montiel NOG, Ortiz-Monasterio I, Vicente FS, Espinoza AG, Atlin G, Wenzl P, Hearne S, Buckler ES (2017) A study of allelic diversity underlying flowering-time adaptation in maize landraces. Nat Genet 49:476–480. https://doi.org/10.1038/ng.3784

Ortiz R, Taba S, Tovar VHC, Mezzalama M, Xu Y, Yan J, Crouch JH (2010) Conserving and enhancing maize genetic resources as global public goods—a perspective from CIMMYT. Crop Sci 50:13–28. https://doi.org/10.2135/cropsci2009.06.0297

Peter R, Eschholz TW, Stamp P, Liedgens M (2009a) Early growth of flint maize landraces under cool conditions. Crop Sci 49:169–178. https://doi.org/10.2135/cropsci2007.10.0538

Peter R, Eschholz TW, Stamp P, Liedgens M (2009b) Swiss Flint maize landraces—a rich pool of variability for early vigour in cool environments. Field Crops Res 110:157–166. https://doi.org/10.1016/j.fcr.2008.07.015

Pollak LM (2003) The history and success of the public-private project on germplasm enhancement of maize (GEM). Adv Agron 78:46–89

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

Reif JC, Hallauer AR, Melchinger AE (2005a) Heterosis and heterotic patterns in maize. Maydica 50:215–223

Reif JC, Hamrit S, Heckenberger M, Schipprack W, Maurer HP, Bohn M, Melchinger AE (2005b) Trends in genetic diversity among European maize cultivars and their parental components during the past 50 years. Theor Appl Genet 111:838–845. https://doi.org/10.1007/s00122-005-0004-5

Revilla P, Butrón A, Malvar RA, Ordás RA (1999) Relationship among kernel weight, early vigor, and growth in maize. Crop Sci 39:654–658. https://doi.org/10.2135/cropsci1999.0011183X003900020007x

Revilla P, Rodríguez VM, Ordás A, Rincent R, Charcosset A, Giauffret C, Melchinger AE, Schön CC, Bauer E, Altmann T, Brunel D, Moreno-González J, Campo L, Ouzunova M, Álvarez Á, Ruíz de Galarreta JI, Laborde J, Malvar RA (2016) Association mapping for cold tolerance in two large maize inbred panels. BMC Plant Biol 16:127. https://doi.org/10.1186/s12870-016-0816-2

Röber FK, Gordillo GA, Geiger HH (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 VM, Romay MC, Ordás A, Revilla P (2010) Evaluation of European maize (*Zea mays* L.) germplasm under cold conditions. Genet Resour Crop Evol 57:329–335. https://doi.org/10.1007/s10722-009-9529-9

Salhuana W, Pollak L (2006) Latin American Maize Project (LAMP) and Germplasm Enhancement of Maize (GEM) project: generating useful breeding germplasm. Maydica 51:339–355

Schnell F (1983) Probleme der Elternwahl—Ein Überblick. In: Arbeitstagung der Arbeitsgemeinschaft der Saatzuchtleiter in Gumpenstein, Austria. 22–24 Nov. Verlag und Druck der Bundesanstalt für alpenländische Landwirtschaft, Austria, pp 1–11

Sood S, Flint-Garcia S, Willcox M, Holland J (2014) Mining natural variation for maize improvement: selection on phenotypes and genes. In: Tuberosa R, Graner A, Frison E (eds) Genomics of plant genetic resources. Springer, Dordrecht

Strigens A, Schipprack W, Reif JC, Melchinger AE (2013) Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding. PLoS ONE. https://doi.org/10.1371/journal.pone.0057234

Unterseer S, Bauer E, Haberer G, Seidel M, Knaak C, Ouzunova M, Meitinger T, Strom TM, Fries R, Pausch H, Bertani C, Davassi A, Mayer KFX, Schön CC (2014) A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. BMC Genom 15:823. https://doi.org/10.1186/1471-2164-15-823

Wang C, Hu S, Gardner C, Lübberstedt T (2017) Emerging avenues for utilization of exotic germplasm. Trends Plant Sci 22:624–637. https://doi.org/10.1016/j.tplants.2017.04.002

Warburton M, Reif J, Frisch M, Bohn M, Bedoya C, Xia X, Crossa J, Franco J, Hoisington D, Pixley K (2008) Genetic diversity in CIMMYT nontemperate maize germplasm: landraces, open pollinated varieties, and inbred lines. Crop Sci 48:617–624

Wilde K, Burger H, Prigge V, Presterl T, Schmidt W, Ouzunova M, Geiger HH (2010) Testcross performance of doubled-haploid lines developed from European flint maize landraces. Plant Breed 129:181–185. https://doi.org/10.1111/j.1439-0523.2009.01677.x

Wright S (1978) Evolution and the genetics of populations, vol 4. University of Chicago Press, Chicago

PNAS RESEARCH ARTICLE | GENETICS OPEN ACCESS

Theoretical and experimental assessment of genome-based prediction in landraces of allogamous crops

Armin C. Hölker[a,b,1], Manfred Mayer[a,1], Thomas Presterl[b], Eva Bauer[c], Milena Ouzunova[b], Albrecht E. Melchinger[a,d,2], and Chris-Carolin Schön[a,2]

Edited by Steven Cannon, US Department of Agriculture–Agricultural Research Service, Midwest Area, Ames, IA; received December 3, 2021; accepted March 25, 2022 by Editorial Board Member Douglas E. Soltis

Discovery and enrichment of favorable alleles in landraces are key to making them accessible for crop improvement. Here, we present two fundamentally different concepts for genome-based selection in landrace-derived maize populations, one based on doubled-haploid (DH) lines derived directly from individual landrace plants and the other based on crossing landrace plants to a capture line. For both types of populations, we show theoretically how allele frequencies of the ancestral landrace and the capture line translate into expectations for molecular and genetic variances. We show that the DH approach has clear advantages over gamete capture with generally higher prediction accuracies and no risk of masking valuable variation of the landrace. Prediction accuracies as high as 0.58 for dry matter yield in the DH population indicate high potential of genome-based selection. Based on a comparison among traits, we show that the genetic makeup of the capture line has great influence on the success of genome-based selection and that confounding effects between the alleles of the landrace and the capture line are best controlled for traits for which the capture line does not outperform the ancestral population per se or in testcrosses. Our results will guide the optimization of genome-enabled prebreeding schemes.

landraces | genomic selection | doubled haploids | gamete capture

Significance

Genetic variation inherent in landraces is essential for broadening the genetic diversity of our crops. This study pioneers the development of a theoretical framework to link molecular inventories of plant genetic resources to phenotypic variation, allowing an informed choice of landraces and their crossing partners. We show that genome-based prediction of genetic values can be implemented successfully in landrace-derived material, despite a strongly reduced level of relatedness compared with elite germplasm. Theoretical derivations are validated with unique experimental data collected on two different landraces. Our results are a pivotal contribution toward the optimization of genome-enabled prebreeding schemes.

Genetic improvement is essential to secure sustainable crop production. Future crops will have to combine high yield potential with major sustainability factors, such as stress tolerance and resource efficiency. To meet these demands, plant breeding will require a reservoir of genetic variation much larger than what is currently found in commercial varieties (1). For maize, it has been estimated that US breeding populations represent only 2% of the entire maize germplasm (2). In contrast, seed banks around the world harbor thousands of untapped landrace accessions (1, 3, 4). Revisiting this vast diversity of landraces is considered promising for elite germplasm improvement (1, 5–9), and developments in molecular, computational, and quantitative genetics open new avenues to make native diversity accessible.

Landraces have been shown to harbor beneficial alleles for traits with limited genetic variation in breeding populations (10), but for most agronomically important traits, they exhibit a substantial performance gap compared with elite germplasm (11–13). While for qualitative traits targeted introgression of favorable alleles discovered in landraces is possible, many traits have a polygenic foundation, which is determined by a large number of genes with small effects. Consequently, marker-based introgression of individual alleles is limited for those traits. Extracting inbred lines directly from landraces and selecting them for superior performance can close the performance gap only partially. Therefore, recurrent population improvement with additional rounds of recombination and selection is necessary to increase the frequency of favorable alleles before introducing landrace-derived genetic material into elite populations. Genome-based selection can accelerate this process, but the theoretical basis of its implementation in prebreeding still needs to be developed.

In outcrossing species, population improvement generally includes three distinct phases (14): 1) sampling candidates from the population to establish progeny for evaluation, 2) evaluating them in multienvironment field trials, and 3) recombining the best candidates to form the next cycle. In genome-based recurrent selection, genomic data are collected in the first phase, and together with data from the second phase, a statistical model is trained for prediction of breeding values of untested candidates from the same or future breeding cycles. The success of this approach depends strongly on the prediction accuracy that can be achieved with the training data. One major determinant is the type of progenies that can be derived from the ancestral landrace (e.g., inbred lines, full- or half-sib families). Additional factors are the quality of phenotyping

Author contributions: A.C.H., M.M., M.O., and C.-C.S. designed research; A.C.H., M.M., T.P., E.B., M.O., A.E.M., and C.-C.S. performed research; A.C.H. and M.M. analyzed data; T.P. and M.O. developed material; E.B., M.O., and C.-C.S. acquired funding; and A.C.H., M.M., A.E.M, and C.-C.S. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. S.C. is a guest editor invited by the Editorial Board.

[1]A.C.H. and M.M. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: melchinger@uni-hohenheim.de or chris.schoen@tum.de.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2121797119/-/DCSupplemental.

Published April 29, 2022.

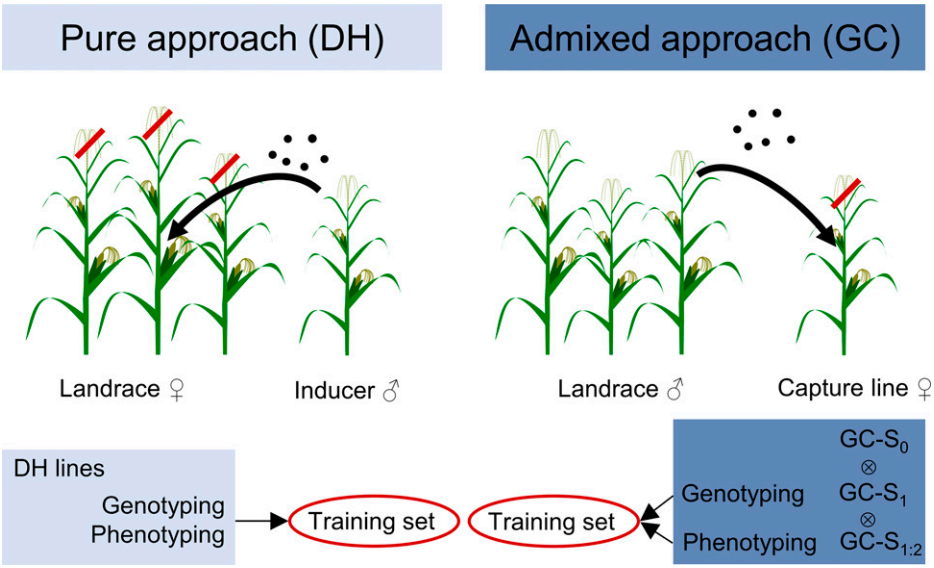


Fig. 1. Scheme of population development for the pure and admixed approaches.

expressed as the heritability (h^2) of the target traits, the sample size (N), and the number of markers (M).

Here, we developed the quantitative genetic framework for two fundamentally different concepts for establishing training populations from landraces. The two concepts are displayed in Fig. 1 and differ with respect to the proportion of landrace genome and technical steps for their production. The "pure" approach entails the production of fully homozygous doubled-haploid (DH) lines from the ancestral landraces. The DH lines exhibit twice the additive genetic variance of the ancestral landrace and allow high-precision phenotyping. The "admixed" approach captures gametes of the landraces in a cross with an inbred (capture) line of different genetic background followed by subsequent selfing of the offspring. When the aim of the prebreeding program is the immediate development of superior inbred lines, a natural choice for the capture line would be a high-performing elite line to increase the usefulness of the resulting population compared with the pure approach. However, the use of an elite capture line has been shown to carry a high risk of reconstructing the elite genome, associated with a loss of landrace alleles in later selection steps (15). We, therefore, investigated the role of the capture line for the genetic improvement of landrace-derived populations with a focus on genome-based recurrent selection. We link generic theory with population-specific molecular parameters and experimental results on several traits, including yield, in four unique populations representing the pure and the admixed approach as well as two ancestral landraces.

Results

Molecular Variances of DH and Gamete Capture Populations Can Be Predicted. We developed populations of DH and gamete capture (GC) lines from two flint maize landraces, Kemater Landmais Gelb (KE) and Petkuser Ferdinand Rot (PE). The French inbred FV2 derived from population Lacaune served as the capture line. Both populations were produced from the same seed batch of the respective landrace, which we defined as the ancestral landrace. In addition to the derived DH and GC populations, a random sample from the ancestral landrace (LS) was genotyped. Across populations, 85 and 92% of the total 472,169 single nucleotide polymorphisms (SNPs) were polymorphic in KE and PE, respectively. The majority of the polymorphic markers (80.9% for KE and 78.4% for PE) segregated in all three populations (Fig. 2*A* and *SI Appendix*, Fig. S1*A*). In both landraces, each population showed a small percentage of segregating markers that were fixed in one or both of the other two populations due to independent sampling from the ancestral landrace. The capture line FV2 carried a SNP allele not present in the LS and DH lines at 13,315 (KE) and 11,488 (PE) genomic positions, thus contributing about half of the private polymorphisms of the GC lines. For both landraces, allele frequencies observed in DH and GC corresponded with allele frequencies estimated from LS and FV2 (Fig. 2 *E* and *F* and *SI Appendix*, Fig. S1 *E* and *F*).

Mean pairwise genetic distances of genotypes ($\bar{X}_{GD}$) in the three types of populations are depicted in Fig. 2*B* and *SI Appendix*, Fig. S1*B* for landrace KE and PE, respectively. Under the assumption of Hardy–Weinberg equilibrium in the ancestral landrace and no selection, the expected $\bar{X}_{GD}$ in LS and DH is a function of the ancestral allele frequencies with $\bar{X}_{GD(LS)} = \bar{X}_{GD(DH)}$ (*SI Appendix*, *SI Text A1* and Table A1). In the GC populations, the allele frequencies of the capture line need to be accounted for. In the experimental LS and DH populations, mean and range of pairwise genetic distances were similar but not identical, with $\bar{X}_{GD(LS)} < \bar{X}_{GD(DH)}$ in KE and vice versa in PE. The more pronounced difference between LS and DH in KE was most likely the result of mild population admixture in the LS, which is reflected by an excess of closely related genotypes (Fig. 2*B*). Mean genetic distances between GC lines and the capture line FV2 calculated based on SNPs for which the LS and DH were monomorphic for the allele not carried by FV2 were in agreement with the expected value 0.5 in both landraces (Fig. 2*C* and *SI Appendix*, Fig. S1*C*). For this reduced set of markers, the variation of genetic distances to FV2 in GC reflects the effect of Mendelian sampling, as GC-S_0 plants are fully heterozygous, and the resulting genetic distances should be equivalent to what is expected in the F_2 generation of a biparental cross.

For all three types of populations, we derived expectations of the total molecular variance and its decomposition between and within genotypes assuming absence of selection. For a single locus, the total molecular variance calculated based on biallelic SNP allele frequencies is expected to be identical for DH and the sample from the ancestral landrace (LS) with $\varsigma^2_{DH} = \varsigma^2_{LS} = 2p(1-p)$,

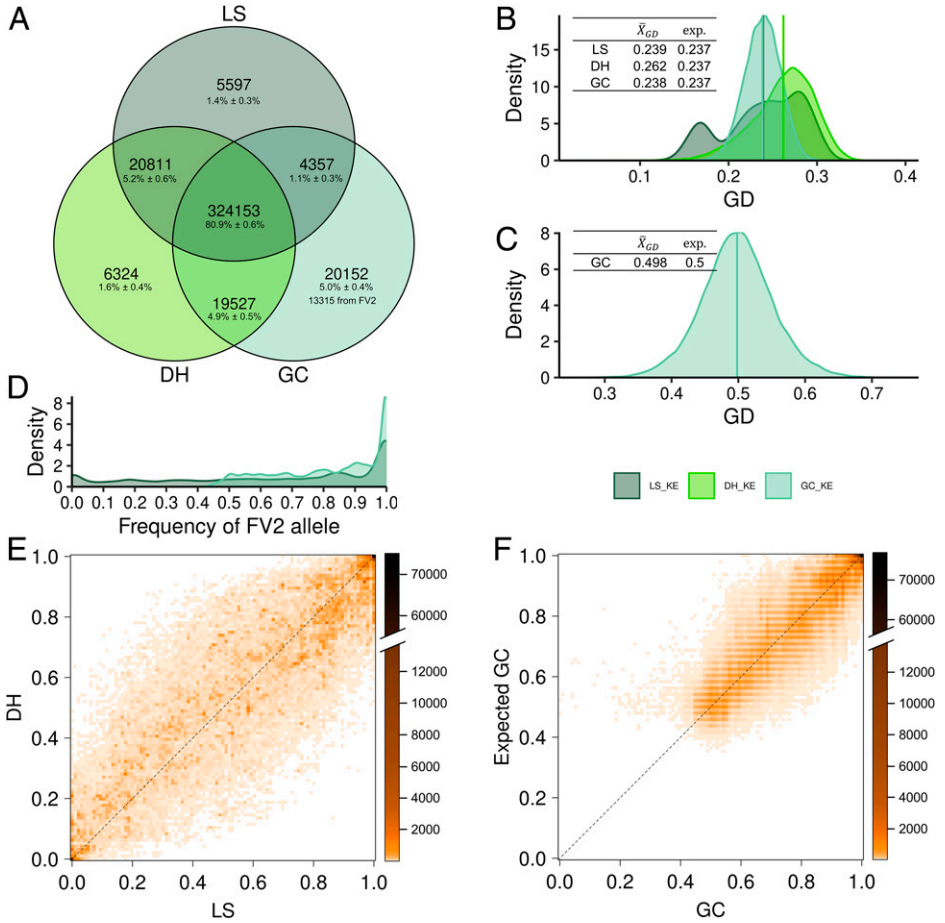


Fig. 2. Venn diagram of the number and percentage of marker polymorphisms shared by and exclusive to the sample of the ancestral landrace (LS), DH lines, and GC lines of KE (*A*). Means and estimated densities of genetic distances (GD) between genotypes within LS, DH, and GC using all markers (*B*) and between GC lines and FV2 using only markers for which DH and LS were monomorphic for the allele not carried by FV2 (*C*). Estimated density of the frequency of the FV2 allele in LS and GC (*D*). Allele frequencies in DH vs. LS (*E*) and expected frequencies in GC (calculated from LS and known FV2 genotype) vs. observed GC (*F*). The calculated numbers of marker polymorphisms (*A*) are the result of sampling 80 gametes per population with 500 replications and are shown as the absolute number and percentage of polymorphic markers (± SD). In GC, the number of polymorphic markers resulting from the cross with FV2 (LS and DH monomorphic for the allele not carried by FV2) is shown as the average across 500 sampling replications. The tables in *B* and *C* show the means of the genetic distances and their expected values (calculated from LS allele frequencies). *B*–*F* are based on the whole set of lines (i.e., $N = 48$ [LS], $N = 471$ [DH], and $N = 274$ [GC]).

with p being the frequency of the allele carried by the capture line (*SI Appendix*, Table A3). For GC, the expected total molecular variance amounts to $\varsigma^2_{GC-S_1} = 0.5(1+p)(1-p)$ (*SI Appendix*, Table A3). Consequently, for a given locus, $\varsigma^2_{GC-S_1} \geq \varsigma^2_{LS}$ if and only if the allele present in the capture line has frequency $p \leq 1/3$ in the ancestral landrace. Loci where the ancestral landrace and therefore also its sample are fixed for an allele different from FV2 (i.e., $p = 0$) contribute maximally to $\varsigma^2_{GC-S_1}$ but not to ς^2_{LS} and ς^2_{DH}. Our theoretical results demonstrate the importance of the genetic makeup of the capture line for building the GC. Here, the capture line FV2 contributed new alleles ($p = 0$ in LS and DH) at 2% (PE) and 3% (KE) of all polymorphic SNP positions, and the proportion of SNPs with $p \leq 1/3$ in LS was about 25% in both landraces (Fig. 2*D* and *SI Appendix*, Fig. S1*D*). Thus, the observed molecular variances for LS, DH, and GC (*SI Appendix*, Table S1) meet expectations.

The linkage disequilibrium (LD) decay distance (δ), for which the pairwise LD of markers on the same chromosome was greater than $r^2 > 0.2$, was slightly higher in the GC than in the DH lines and was higher in populations derived from KE ($1{,}032 \leq \delta \leq 1{,}263$ kb) than from PE ($399 \leq \delta \leq 660$ kb) (Fig. 3*A*). Within landraces, linkage-phase similarities (LPS) were high for the pairwise comparison of LS and DH but substantially reduced for LS and GC (Fig. 3*B*). Across landraces, LPS for the pairwise comparison of the same type of population was low for LS and DH but moderate to high for GC (Fig. 3*C*). Average LD between markers on different chromosomes was negligible in all populations and both landraces.

Experimental and Theoretical Results Are in Good Agreement for All Populations. The conceptual differences between the pure and the admixed approach with respect to means and genetic variances in DH and GC are visualized in *SI Appendix*, Fig. A1 based on the theoretical expectations given in Table 1 and *SI Appendix*, Table A2. In hybrid breeding, selection candidates are evaluated not only for their per se performance (PP) but mainly, for their combining ability with a tester from a different heterotic group. We, therefore, considered both the PP of the LS and of GC-$S_{1:2}$ and DH lines as well as their testcross performance (TP) with an inbred line from the dent heterotic pool. Assuming absence of epistasis, the PP of fully inbred generations (DH lines, GC-$S_{1:\infty}$ lines) and of all testcrosses can be described with a purely additive model (Table 1 and *SI Appendix*, *SI Text A2* and Table A2). For PP of the LS and for GC-$S_{1:2}$ lines, the mean and dispersion of the genotypic values depend on unknown landrace- and capture line specific dominance effects $[d]$ and $[d^*]$, respectively (Table 1).

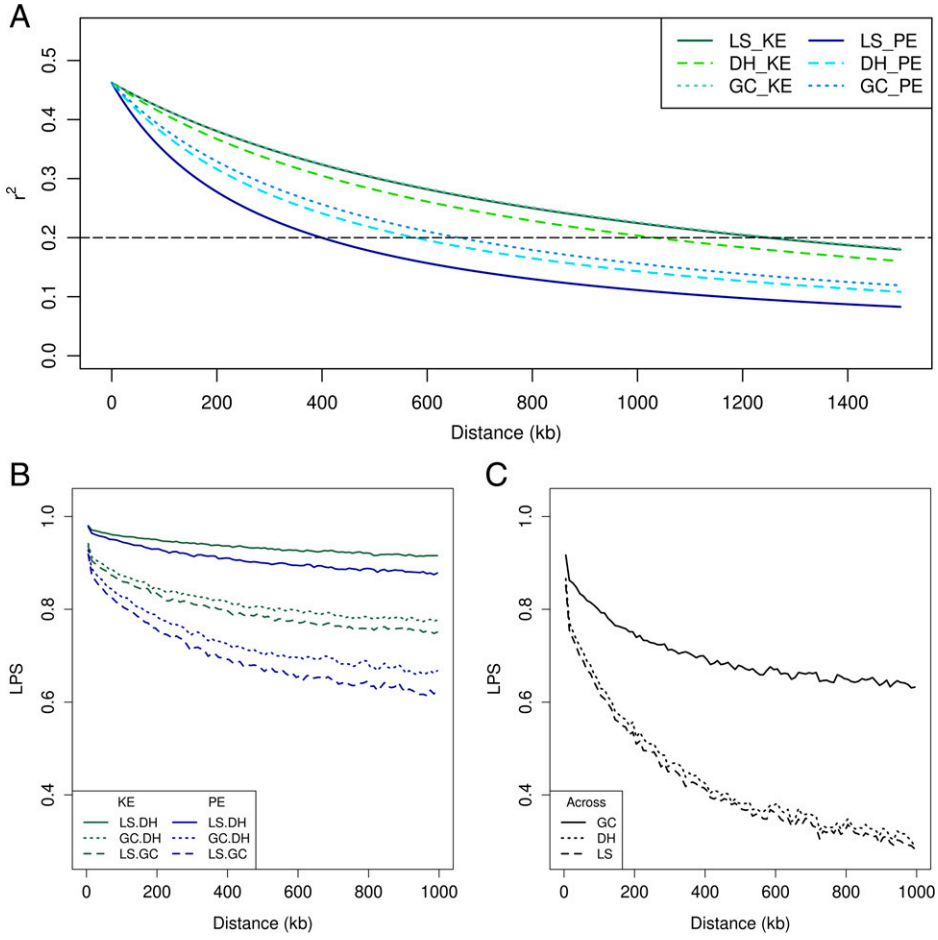


Fig. 3. Decay of LD with physical distance for the sample of the ancestral landrace (LS), the DH lines, and the GC lines of landraces KE and PE (*A*). Linkage phase similarities (LPS) for pairwise comparisons of the three types of populations within each landrace (*B*) and LPS for pairwise comparisons of the same type of population across the two landraces (*C*). For all calculations, 94 gametes were randomly sampled for each group.

In the following, we use the trait flowering time exemplarily to link theoretical and experimental results (Fig. 4 and *SI Appendix*, Fig. S2). Phenotypic values should be indicative of genotypic values, as heritabilities were high, ranging between 0.85 and 0.93 (Fig. 4 *A* and *B* and *SI Appendix*, Fig. S2 *A* and *B*). The LS flowered significantly earlier than most DH lines, and estimates of $[d]$ amounted to about 7% of the LS performance for both landraces. The inbred capture line FV2 flowered significantly earlier than the mean of the DH lines from KE and PE, pointing to an enrichment of early flowering

Table 1. Quantitative-genetic expectations of means and genetic variances for per se (PP) and testcross (TP) performance in the sample of the ancestral landrace (LS), derived DH, and GC lines

	Coefficient of parameters[†]									
	Population mean[‡] $\bar{x}$				Genetic variances					
					Primary variance		Variance within families		Total variance	
Population	$a+\Delta$	$(p-(1-p))a$	$[d]=2p(1-p)d_{12}$	$[d^*]=pd_{1x}+(1-p)d_{2x}$	σ^2_A	$\sigma^2_{A^*}$	σ^2_A	$\sigma^2_{A^*}$	σ^2_A	$\sigma^2_{A^*}$
LS	0	1	1	0	1	0	—	—	1	0
DH	0	1	0	0	2	0	0	0	2	0
GC-$S_{1:2}$	1/2	1/2	0	1/4	3/4	1/4	1/8	1/8	7/8	3/8
GC-$S_{1:\infty}$	1/2	1/2	0	0	3/4	1/4	1/4	1/4	1	1/2
FV2	1	0	0	0	—	—	—	—	—	—

For GC lines, the total genetic variance is decomposed into the primary variance between families as observed for GC-$S_{1:2}$ lines in this study and the variance within families.
[†]Parameters $[d]$ and $[d^*]$ are not required for TP.
[‡]p and $(1-p)$ refer to the frequencies of alleles A_1 and A_2 in LS, respectively. a and $a+\Delta$ refer to the additive effects in LS and the capture line, respectively, with different meanings for PP and TP. $[d]$ and $[d^*]$ refer to the contribution of dominance effects to the PP of LS and GC-$S_{1:2}$, respectively, where d_{12}, d_{1x}, and d_{2x} refer to the dominance effect of genotypes A_1A_2, A_1A_x, and A_2A_x, respectively, with A_x being the allele of the capture line. σ^2_A refers to the additive variance inherent in the ancestral landrace, with $\sigma^2_A = 2p(1-p)a^2$. $\sigma^2_{A^*}$ refers to the additive variance resulting from the effects of the capture line alleles, with $\sigma^2_{A^*} = 2((1-p)a+\Delta/2)^2$ (details are in *SI Appendix*, *SI Text A2*).

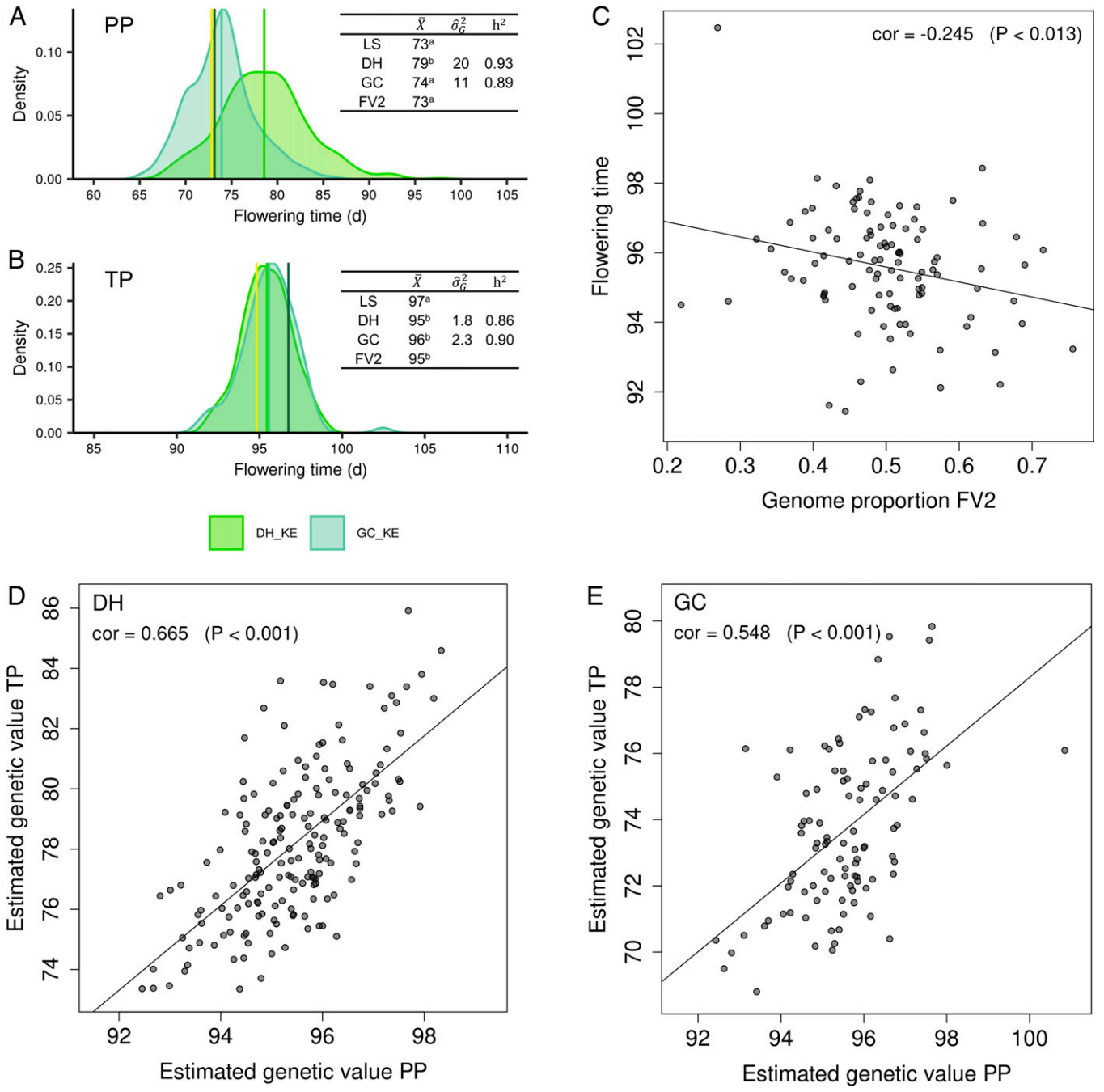


Fig. 4. Estimated densities showing the distribution of phenotypic values for per se performance (PP; *A*) and testcross performance (TP; *B*) of the DH and GC lines for landrace KE, scatterplots of proportions of FV2 genome vs. TP for flowering time (*C*), and estimated genetic values of PP vs. estimated genetic values of TP for flowering time in DH (*D*) and GC (*E*) lines. In *A* and *B*, the means (vertical lines) of the landrace sample (LS, dark green) and the capture line FV2 (yellow) are indicated, and the tables show the means ($\bar{X}$), genetic variances ($\hat{\sigma}_g 2$), and heritabilities (h^2). Means with a shared letter are not significantly different ($P > 0.05$). *C–E* indicate the Pearson correlation coefficients and corresponding *P* values of the shown correlations.

alleles in FV2. Mean flowering time of GC lines was significantly earlier than of DH lines. Under an additive model, the mean of the GC lines is expected to lie exactly between the mean of the DH lines and the capture line but was shifted toward FV2 in both landraces, indicating capture line specific dominance effects $[d^*]$ contributing to GC per se performance. In the testcrosses, differences between the capture line and the mean of the DH lines were attenuated, and consequently, mean DH and GC testcross performance was not significantly different in both landraces.

We derived expected genetic variances of the DH and GC populations (Table 1). The DH delivers the maximum additive genetic variance inherent in the respective landrace with $\sigma^2_{g(DH)} = 2\sigma^2_A$. The additive genetic variance among GC lines is $\sigma^2_{g(GC-S_{1:2})} = \frac{3}{4}\sigma^2_A + \frac{1}{4}\sigma^2_{A*}$ (Table 1), with $\sigma^2_{A*} = 2((1-p)a + \Delta/2)^2$ being a function of the effect and frequency of the alleles in the ancestral landrace and the effects of alleles originating from the capture line. If for a given trait the capture line contributes only alleles present in the ancestral landrace, as would be the case for a random DH line derived from the ancestral landrace, the genetic variance of the GC lines should be half the genetic variance among DH lines as $\sigma^2_{A*} = \sigma^2_A$ (*SI Appendix, SI Text A2*).

If the capture line contributes alleles not present in the landrace at many loci and the allelic effects at these loci differ substantially from the respective landrace alleles, the genetic variance among GC lines can be equal to or even larger than the variance among DH lines.

Estimates of genetic variances for flowering time were significant for PP and TP in both landraces (Fig. 4 *A* and *B* and *SI Appendix*, Fig. S2 *A* and *B*). For PP, the ratio $\sigma^2_{g(GC)}/\sigma^2_{g(DH)}$ was 0.55 in KE and 0.67 in PE. The results suggest that the additive genetic variance σ^2_{A*} generated by crossing the landrace with FV2 did not differ substantially from σ^2_A, despite the enrichment of earliness alleles in FV2. In the testcrosses, however, the ratio $\sigma^2_{g(GC)}/\sigma^2_{g(DH)}$ was 1.28 in KE and 1.35 in PE, indicating $\sigma^2_{A*} > \sigma^2_A$. The reduction in σ^2_g in the testcrosses compared with per se performance was much higher for DH than GC lines. These results indicate that dominance interactions with the tester allele differed for the DH and FV2, and consequently also for GC, in both landraces. Nevertheless, genetic covariances between PP and TP were significant for both types of populations and both landraces. Genotypic correlations as well as correlations of estimated genetic values between PP and TP for flowering time were higher for DH than for GC lines (Fig. 4 *D* and *E* and *SI Appendix*, Fig. S2 *D* and *E* and Table S2).

One concern with the GC population is the overrepresentation of the capture line genome in the progeny after selection. The proportion of FV2 genome of GC-S_1 plants determined with the reduced marker set (SNP alleles with $p = 0$ in LS and DH) ranged from 21.9 to 75.6% in KE and from 21.5 to 73.1% in PE, with averages of 50.3 and 50.2%, respectively, meeting expectations. A significant correlation of FV2 genome proportion and phenotypic performance was observed for flowering time only in the testcrosses of KE. As expected, the correlation was negative but weak ($r = -0.25$) (Fig. 4*C* and *SI Appendix*, Fig. S2*C*), demonstrating that GC lines enriched with earliness alleles can be selected without strong overrepresentation of the FV2 genome.

Results for the other traits are presented in *SI Appendix*, Figs. S3 and S4. In general, experimental results were in agreement with theoretical expectations and highly consistent across landraces. Estimates of $[d]$ for the two plant height traits amounted on average to about 26% relative to the performance of the LS. The mean of the GC lines for plant height was about halfway between LS and FV2, also indicating dominant type of gene action. For all traits, the ratio of genetic variances $\sigma^2_{g(GC)}/\sigma^2_{g(DH)}$ followed the same trend as shown for flowering time in both landraces for PP and TP. Correlations of FV2 genome proportion and observed phenotypic performance were not significant for all traits and both landraces (except flowering time in testcrosses as described above). For early plant height, the GC lines showed only low (PE) or nonsignificant (KE) genetic correlations between PP and TP, while for DH lines, they were intermediate to high (*SI Appendix*, Table S2).

Population Type Determines Accuracy of Genomic Prediction. The accuracy ρ of genome-based prediction is the success criterion for genomic selection. Increasing the sample size of the training set affected the magnitude and precision of ρ, and no plateau was reached up to $N = 250$ (Fig. 5*A* and *SI Appendix*, Fig. S5*A*). With respect to marker density, an increase in prediction accuracy could be observed up to 15,000 SNPs (*SI Appendix*, Fig. S6). Prediction accuracies ρ were consistently higher in DH lines compared with GC lines for the two plant height traits for all tested sample sizes of the training set. Differences were most pronounced for small sample sizes. For flowering time, differences between the two types of populations were negligible. Yield and dry matter content were assessed in testcrosses only. Accuracies for yield exceeded 0.5 in DH lines even with sample sizes $N < 200$ (Fig. 5*B* and *SI Appendix*, Fig. S5*B*). However, in the GC lines, prediction for yield failed ($\rho = -0.09$ in KE) and was very low for early plant height. The strong decrease in prediction accuracies of testcross traits in the GC can partially be accounted for by a combination of nonsignificant genetic variances in a high number of training sets (Fig. 5*D* and *SI Appendix*, Fig. S5*D*) and the limited size of the prediction sets in cross-validation ($N = 25$). In DH lines, however, testcross accuracies exceeded those of per se performance in some cases, despite lower genetic variances, lower heritabilities, and smaller training set size (plant height at V6 stage in KE, flowering time in PE).

Prediction accuracies across the two types of populations (Fig. 5 *C* and *D* and *SI Appendix*, Fig. S5 *C* and *D*) were low (0.20 to 0.49 for lines per se, 0.03 to 0.41 for testcrosses). The higher accuracies observed within DH lines (e.g., for final plant height) were not reflected in prediction across populations. Accuracies were similar irrespective if the prediction model was trained on DH to predict GC or vice versa. We also investigated if combining the two populations yielded a predictive advantage over within-population prediction. Despite a substantial increase in sample size, accuracies changed only marginally (from 0.53 to 0.55 on average across traits) (*SI Appendix*, Table S3), which was not expected considering the increase in prediction accuracy within populations with increasing N (Fig. 5*A* and *SI Appendix*, Fig. S5*A*).

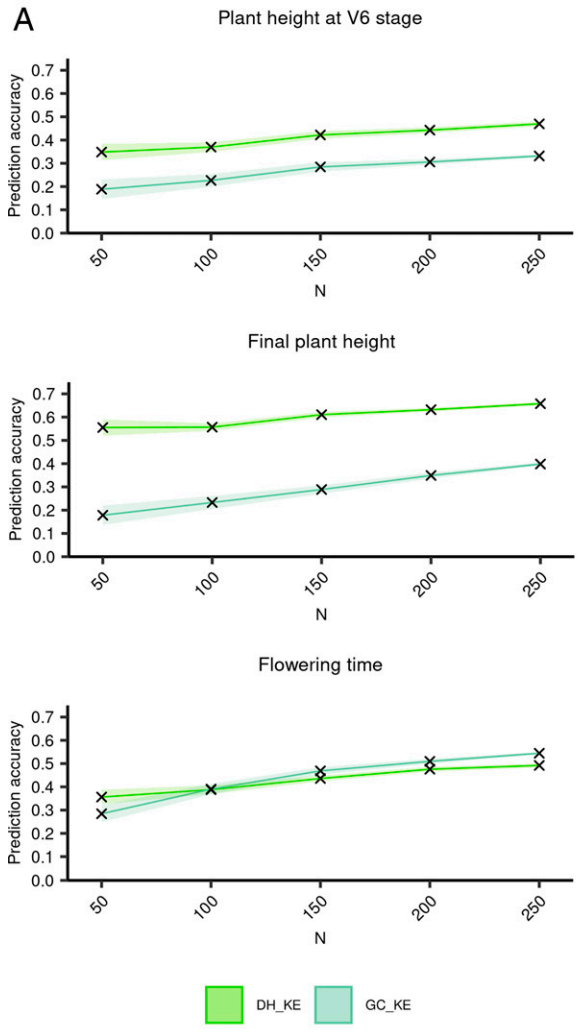
Prediction across landraces (e.g., using DH lines of KE for training and DH lines of PE for prediction or vice versa) yielded estimates of ρ close to zero for DH lines irrespective of which landrace was used for model training (Fig. 5*E* and *SI Appendix*, Fig. S5*E*). For GC lines, higher values were obtained especially for final plant height and flowering time ($0.25 \leq \rho \leq 0.29$), most likely due to shared haplotypes originating from FV2, resulting in much higher linkage phase similarities of GC compared with DH populations (Fig. 3*C*).

Estimates of ρ varied substantially for the different (cross-) validation runs (Fig. 5 and *SI Appendix*, Fig. S5) within and across populations as well as across landraces. In testcrosses with a training set size of $N = 75$ (per se $N = 200$), this was most pronounced. The variation was in part attributable to nonsignificant estimates of the genetic variance in the training set, which was more common in GC and most pronounced in testcross prediction.

Discussion

Extraction of beneficial haplotypes from landraces is a long-term endeavor. In landrace genomes, favorable alleles for one trait are often in high LD with unfavorable alleles for the same or other traits, and consequently several rounds of recombination and selection are required to close the performance gap to elite material and reduce linkage drag (16). With this study, we aimed to fill the knowledge gap on genome-based prediction accuracies that can be achieved in landrace-derived material in outcrossing species.

Prediction Accuracies Are High in Landrace-Derived DH Populations. Prediction accuracies achieved in this study with landrace-derived DH lines clearly demonstrate that genome-



B

		DH_PP	GC_PP	DH_TC	GC_TC
N_{max}		471	274	183	103
ρ	PH_V6	0.54	0.34	0.72	0.13
	PH_final	0.71	0.42	0.57	0.34
	FF	0.58	0.55	0.51	0.51
	DMC	-	-	0.44	0.38
	TDMY	-	-	0.51	-0.09

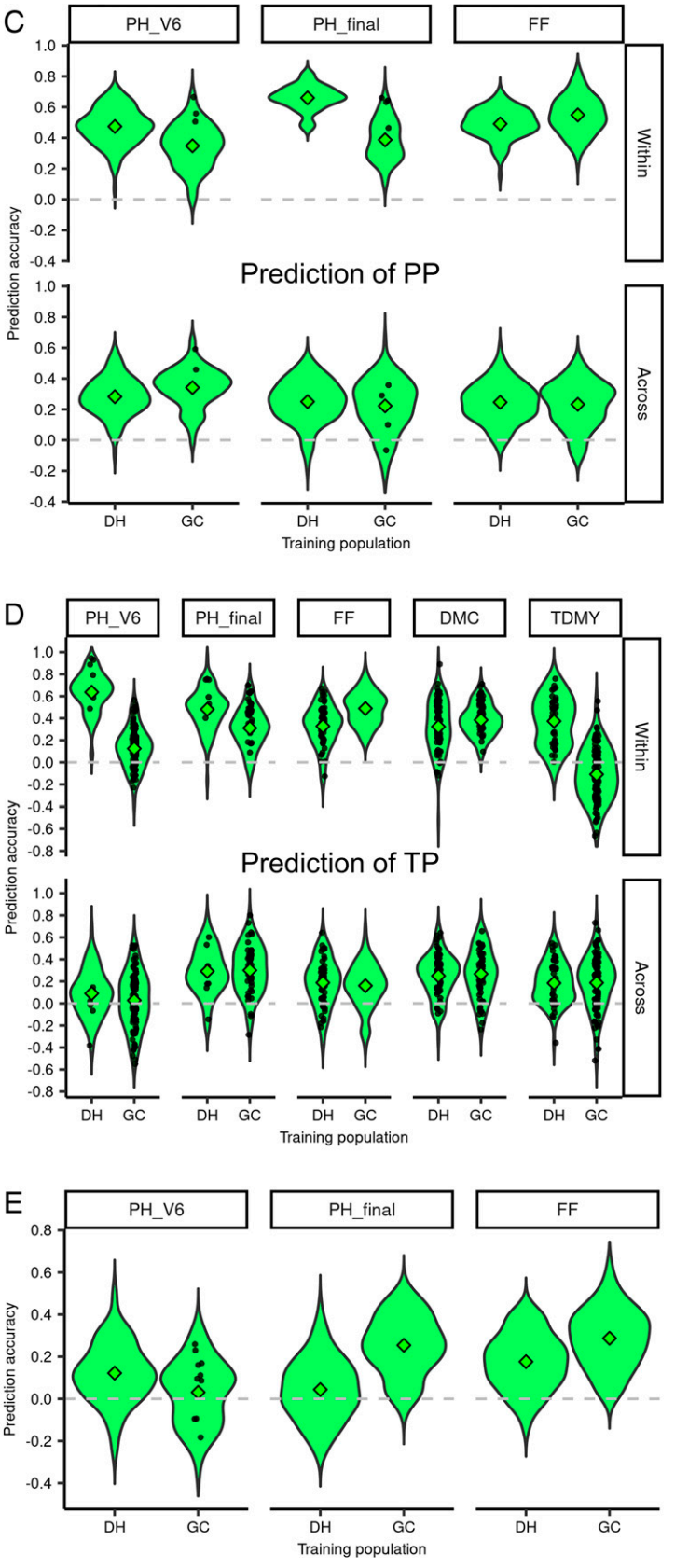


Fig. 5. Prediction accuracy (ρ) in landrace KE for per se performance (PP) in the DH and GC lines as a function of sample size N (*A*), for prediction of PP and testcross performance (TP) at the maximum available number of lines (N_{max}; *B*), for predictions within and across populations for PP (*C*) and TP (*D*) in DH and GC, and for across-landrace prediction for PP from KE (training on PE; *E*). Traits are plant height at V6 stage (PH_V6), final plant height (PH_final), and flowering time (FF) in PP and TP and dry matter content (DMC) and total dry matter yield (TDMY) in TP. For each N (*A*), sampling of lines was repeated 100 times, and 10 times fivefold cross-validation was carried out within each sample, yielding the basis for calculating the presented means and 95% quantiles (shaded areas around the curve). Prediction across and within populations as well as across landraces was carried out by randomly sampling $N = 200$ and $N = 75$ lines for training in PP (*C* and *E*) and TP (*D*), respectively, for predicting $N = 50$ (PP; *C* and *E*) or $N = 25$ (*D*) genotypes of the same or corresponding population (*C* and *D*) or the same population of the other landrace (*E*). Sampling was repeated 100 times. The violin plots (*C*–*E*) show all 100 values, with the diamonds indicating the means. Black dots show values of the prediction accuracy estimated from models where the genomic variance estimate was not significant (likelihood-ratio-test, $P > 0.05$).

based selection has great potential. Cross-validated accuracies for prediction of total dry matter yield were as high as 0.58 in testcrosses of PE DH lines (0.51 in KE) and even higher for other traits, despite sample sizes of less than 150 DH lines in the training set. Correlated estimated genetic values for PP and TP indicate effective genomic selection for traits like flowering time and plant height on the per se level, carrying over to correlated response for TP.

In landrace-derived DH populations, prediction accuracies should be merely a function of LD between markers and quantitative trait loci (QTL), as gametes are sampled at random from the ancestral population. Thus, it was surprising that accuracies were of similar magnitude as reported for elite maize germplasm with much higher LD and substantial relatedness between genotypes (17). Inflation of accuracies caused by DH lines with extreme values due to strong inbreeding depression could be ruled out by investigating cross-validation prediction sets manually. Hidden relatedness and population structure in the DH population, both factors that might inflate prediction accuracy, were not observed when investigating pairwise genetic distances of DH lines (Fig. 2 and *SI Appendix*, Fig. S1).

We, therefore, conclude that in DH populations derived from landraces preselected for molecular and phenotypic properties as suggested by Mayer et al. (18), prediction accuracies of 0.5 or higher can be considered a realistic benchmark in genome-based selection of complex traits with training set sizes of $N \geq 200$ due to large additive genetic variance, high heritabilities, and moderate LD.

Efficiency of the Admixed Approach. Some landrace populations carry high genetic load, leading to low efficiency of DH production. Thus, crossing the landrace with an inbred capture line from a different genetic background might be the only option to avoid homozygous deleterious allele combinations. So, what are the consequences for prediction accuracies in comparison with DH populations? As expected from theory and observed in the experimental populations of this study, average allele frequencies of polymorphic SNPs were shifted toward more unbalanced allele frequencies in the GC lines (Fig. 2 and *SI Appendix*, Fig. S1), affecting locus-specific contributions to the total genetic variance. If the capture line carries an allele present in the ancestral landrace, the locus-specific variance in the GC decreases compared with the DH population, except for loci with extreme allele frequencies in the ancestral landrace ($p \leq 1/6$) (*SI Appendix*, *SI Text A3*). If the capture line carries an allele not present in the ancestral landrace, the locus-specific variance in the GC will depend on the effect of this allele as well as on the frequencies and effects of the alleles in the ancestral landrace. If the allele of the capture line exhibits dominance over the landrace alleles (i.e., $[d^*] > 0$), the dominance variance might increase at this locus (*SI Appendix*, *SI Text A2*). Thus, when training the model on DH or GC lines, the weight assigned to individual SNPs can differ markedly between the two populations, explaining the fairly low prediction accuracies across populations, irrespective if model training was conducted on DH or GC lines.

Crossing with a capture line will affect linkage phases between markers and QTL and the extent of LD compared with the DH lines (Fig. 3). All GC-S_0 plants are half-sibs and share one identical gamete. Through the subsequent selfing process, haplotypes may arise with different linkage phases and LD decay compared with those of the ancestral landrace, compromising prediction accuracies within the GC and across populations. This effect will be trait-specific and will depend strongly on the genetic makeup of the capture line. As could be seen from the experimental data, linkage phase similarities with the LS were considerably reduced in GC compared with DH lines. Prediction accuracies for plant height and especially for testcross yield were substantially reduced in the GC populations, but not for flowering time or dry matter content. We hypothesize that for the two maturity-related traits, the capture line FV2 enriched the GC populations with alleles not present in either of the two landraces at a substantial number of loci. These alleles occur with frequency 0.5 in the GC population and thus, obtain high weight in prediction compensating for the negative effects of opposing linkage phases between markers and QTL at other loci.

When predicting across landraces, accuracies were close to 0 for DH populations but >0.2 for GC lines when predicting in KE onto PE and vice versa. These results corroborate the hypothesis that prediction in the GC populations was at least partially driven by additive effects of shared FV2 haplotypes and/or their dominance over the landrace alleles (Fig. 3*C*).

Genome-Based Improvement of Landraces. In this study, we investigated the potential of genome-based prediction to increase the frequency of favorable alleles of target traits in landrace-derived populations. We conclude that the pure approach is to be preferred over the admixed approach, because with the admixed approach a substantial reduction in prediction accuracy must be expected unless prediction is driven by capture line alleles. When implementing the admixed approach, the choice of capture line will have a major impact on the success of the prebreeding program. It determines the mean and genetic variance of the GC population and the risk of masking favorable landrace alleles. Molecular data can inform about locus-specific allele frequencies in the ancestral landrace and the capture line, and under certain assumptions, these allele frequencies translate directly into expectations for the molecular and genetic variance in the GC population (*SI Appendix*, *SI Text A3*). For quantitative traits, however, many loci contribute to the genetic variance, and unless a large proportion of causal variants for the traits of interest is known, molecular parameters will provide little guidance on the choice of capture line. In this study, the phenotypic per se performance of inbred line FV2 compared with the LS and the mean of the DH lines provided a first indication for which traits the capture line might contribute alleles not present in the ancestral landrace and which type of gene action to expect. It remains to be shown for other GC populations derived from different landraces and capture lines if this pattern holds. We could also show that dominance interactions with the tester alleles differed for landrace and capture line alleles, affecting prediction accuracies in the DH and the GC populations differently. Thus, not only the capture line per se but also its interaction with the tester had a direct effect on the genetic variance accessible for selection.

In summary, the results of this study show that the pure approach has clear advantages over the admixed approach for genome-based improvement of landraces. With continuous technological advances, the application of DH technologies is likely to become routine in many plant genetic resources (19). If the production of fully inbred lines either by the DH technology or by recurrent selfing is not possible, the admixed approach is still a good alternative. The risk of masking valuable variation present in the landrace needs to be minimized by an informed choice of capture line and tester. Our study shows that the confounding effects between the alleles of the landrace and the capture line are best controlled for traits for which the capture line does not outperform the ancestral population per se or in testcrosses.

Materials and Methods

Plant Material. We applied two different strategies (Fig. 1) for sampling gametes from European maize landraces. The landraces KE and PE of European flint maize were chosen of 35 landraces for this study on the basis of population-genetic analyses described by Mayer et al. (18) and phenotypic screening for variation in early-development traits assessed in field trials. DH lines were derived directly from the landraces for the first sampling strategy (pure approach) (11). For the second sampling strategy (admixed approach), we modified a scheme originally proposed by Stadler (20): pollen mixtures from the landraces were used to pollinate the capture line FV2. FV2 is an important founder line of the European flint heterotic group developed by INRA from the French landrace Lacaune and was intensively used as parent in commercial hybrids between the 1960s and 1990s. We termed this procedure "gamete capture" (GC). The GC-S_0 plants are half-sibs, with one gamete from FV2 and the other gamete from the landrace. Subsequently, the GC-S_0 plants were selfed to produce GC-S_1 ears. One GC-S_1 plant per ear was genotyped and selfed. Field evaluation was performed with the corresponding GC-S_2 lines planted ear to row, subsequently referred to as GC-$S_{1:2}$. For each landrace, all populations were derived from the same seed source, which we define as the ancestral landrace. Three different sets of seeds from this ancestral landrace were randomly sampled to obtain 1) the sample of the ancestral landrace (LS), 2) the landrace plants used for DH induction, and 3) the landrace plants used to pollinate the capture line. For production of testcross seed, randomly chosen lines from each population as well as FV2 and plants sampled from the ancestral landrace were hand-crossed as pollinators onto the inbred line F353 (INRA, France), a prominent line of the European dent heterotic group.

Field Experiments and Phenotypic Data Analysis. The DH and GC populations were evaluated in adjacent field trials connected through common checks. Field experiments for the DH populations were described in detail by Hölker et al. (11); phenotyping of the GC populations was performed analogously. Briefly, per se performance (PP) was evaluated in four environments in Germany: Roggenstein (ROG) and Einbeck (EIN) in 2017 and 2018. Two separate but adjacent sets of 8 (DH 2018) or 10 (DH and GC 2017, GC 2018) 10×10 lattice designs with two replicates per line were used in each environment. As common checks, we added plants sampled from the ancestral landrace (LS) as well as 15 (2017) or 4 (2018) inbred lines, including in both years the line FV2. Plots were single rows of 3 m length, with 0.75 m distance between rows, and planting density was 8.8 plants m^{-2}.

Testcross performance (TP) was evaluated in two environments (ROG and EIN) in 2019. Testcrosses of DH lines were grown in four 10×10 lattice designs; for GC lines, a generalized α-lattice design with 200 entries was used. Testcrosses of the LS and of two inbred lines together with six commercial hybrids were included as checks in all trials, and FV2 was included in GC trials only. Plots were double rows of 6 m length, with 0.75 m distance between rows and planting density of 9 or 11 plants m^{-2}. Sowing, fertilization, and plant protection in per se and testcross evaluation followed standard agricultural practice at the experimental stations.

The traits plant height at V6 stage (PH_V6, cm), final plant height (PH_final, cm), flowering time (FF, days from sowing until 50% of plants in the plot silked), dry matter content (DMC, percentage, only TP), and total dry matter yield (TDMY, dt/ha, only TP) at forage harvest were investigated.

We expanded the analysis described for the DH experiments in Hölker et al. (11) for joint analysis of the GC and DH experiments in a single step using the following model:

$$\begin{aligned} y_{ijkopst} = {} & \mu + m_i + \delta_{Checks} l_j + g_{k(ij)} + u_o + gu_{ko(ij)} \\ & + \delta_{DH}\{lu_{jo} + k_{p(o)} + r_{s(op)} + b_{t(ops)}\} \\ & + \delta_{GC}\{lu_{jo} + k_{p(o)} + r_{s(op)} + b_{t(ops)}\} + \varepsilon_{ijkopst}, \end{aligned} \quad [1]$$

where $i = 1, 2, 3, 4$ denotes four groups (GC, DH, LS, and checks); $j = 1, 2, 3, 4$ denotes the different populations (GC_KE, GC_PE, DH_KE, and DH_PE); μ is the overall mean; m_i is the effect of group i; l_j is the effect of population j in groups $i = 1$ and 2; δ_{Checks} is a dummy variable with $\delta_{Checks} = 1$ if the line belongs to DH or GC populations and $\delta_{Checks} = 0$ for LS or inbred lines used as checks; δ_{DH} (δ_{GC}) is a dummy variable with $\delta_{DH} = 1$ ($\delta_{GC} = 1$) if data belong to the DH (GC) experiment and $\delta_{DH} = 0$ ($\delta_{GC} = 0$) otherwise; $g_{k(ij)}$ is the genotypic effect of line k nested in group i and population j; u_o is the effect of environment o; lu_{jo} is the interaction of population j and environment o; and $gu_{ko(ij)}$ is the interaction of genotype k and environment o. The effects $k_{p(o)}$, $r_{s(op)}$, $b_{t(ops)}$, and $\varepsilon_{ijkopst}$ refer to the effect of the lattice (nested in environments), replicate (nested in lattices in environments), incomplete block (nested in replicates in lattices in environments), and the residual error, respectively. All effects except m_i and l_j were treated as random. Genotype [$g_{k(ij)}$] and genotype $\times$ environment [$gu_{ko(ij)}$] variance components were modeled individually for the populations ($j = 1, 2, 3, 4$), assuming that DH and GC lines across and within landraces were stochastically independent. Residuals were assumed to be normally distributed with mean zero and four heterogeneous variances, one each for $\delta_{Checks} = 1$ and $\delta_{Checks} = 0$ in GC and DH experiments, assigning the same residual variance to all GC and DH lines within all environments. Raw data and outliers were manually curated by inspection of residual plots. The model in Eq. **1** refers to the analysis of PP. TP was analyzed analogously, adjusting for the generalized α-lattice design used in GC trials. Variance components and their SEs were estimated with ASReml-R package 3.0 (21). Entry-mean heritabilities were calculated for each population following Hallauer et al. (14), and SEs of heritability estimates were derived using the delta method (22). Heritabilities (h^2) and variance component estimates exceeding twice their SEs were considered significant. For obtaining best linear unbiased estimates (BLUEs) of the genotypic value of each entry, the model from Eq. **1** was simplified, replacing factors m_i, $\delta_{Checks} l_j$ with a factor separating the two experiments (DH and GC), dropping $\delta_{DH} lu_{jo}$ and $\delta_{GC} lu_{jo}$ from the model, and treating genotype as a fixed effect. This model was also used to test for significant differences (t-tests) between LS, DH, GC, and FV2 in linear contrasts calculated with the package asremlPlus (23). For estimating genetic covariances and genetic correlations between PP and TP for a given phenotypic trait, we expanded the model from Eq. **1** to a bivariate model treating PP as one trait and TP as the other trait. Significance of genetic covariances was tested in likelihood-ratio-tests comparing the model including the covariance with the reduced model without the covariance.

Genetic Data Analysis. The inbred line FV2, samples from each ancestral landrace (LS), DH lines, and GC plants were genotyped with the 600k Affymetrix Axiom Maize Array (24). The quality filtering of the SNP data for the LS and DH populations was described in detail in Hölker et al. (11) and was done analogously for the GC populations. Briefly, markers were filtered according to the best quality class (24) and an unambiguously mapped physical position in the B73 reference sequence AGPv4 (25). Markers and individuals with >10% missing values were removed. For DH lines, markers and individuals with >5% heterozygous genotype calls were removed, and the remaining heterozygous calls (0.19%) were set to missing values. For DH lines, missing values were imputed separately for each population using Beagle version 5.0 (26) with default settings. Missing values in the LS and GC were imputed, and two gametes from each individual were phased using Beagle version 5.0, with parameters iterations = 50, phase-segment = 10, and phase-states = 500. Markers were coded as counts of the FV2 allele (0: homozygous for opposite allele of FV2; 1: heterozygous; 2: homozygous for FV2 allele). In total, 1,512 genotypes (LS_KE = 48, LS_PE = 47, DH_KE = 471, DH_PE = 402, GC_KE = 274, GC_PE = 270) with 472,169 polymorphic SNPs remained for further analysis. Thereof, all DH and GC have been evaluated for PP, and a subset (DH_KE = 183, DH_PE = 173, GC_KE = 103, GC_PE = 54) has also been evaluated for TP.

Analysis of Molecular Variance and Genetic Diversity. We sampled 80 gametes from each population (LS, DH, and GC) and landrace with 500 replicates for comparing the number and percentage of polymorphic markers across populations.

LD was measured using r^2 (27) for samples of 94 gametes within each population. We calculated r^2 for pairs of SNPs within a distance of 1 Mb and used nonlinear regression to investigate the r^2 decay with physical distance (28). The LD decay distance is defined as the physical distance δ for which the curve reaches $r^2 = 0.2$. For estimating LD across chromosomes, we sampled 5,000 markers per chromosome with replacement for all 45 pairwise combinations of chromosomes and calculated r^2 for all pairs of markers across chromosomes. Linkage phase similarities (LPS) between populations were calculated according to Schopp et al. (29). LPS according to physical distance was calculated grouping marker pairs into bins of 10 kb up to a maximum distance of 1 Mb.

Genetic distance (GD) between two genotypes was measured as $GD = 1 - SM$, where SM is the simple matching coefficient across all SNP loci calculated as detailed by Jacobson et al. (30). We also compared allele frequencies 1) between DH and LS and 2) between the experimental and expected GC, where the expected GC was obtained by $(p + 1)/2$, with p being the frequency of the

FV2 allele in the respective LS. An analysis of molecular variance (31) based on Euclidean distances was used to estimate the molecular variance within and between individuals of LS, DH, and GC for each landrace. Calculations of the proportion of markers with $p = 0$ and $p \leq 1/3$ as well as the average allele frequency for each population were based on 415,346 (KE) and 446,687 (PE) markers polymorphic across LS, DH, and GC.

Genome-Based Prediction Model. We performed genomic best linear unbiased prediction (GBLUP) in several scenarios for PP and TP, always applying the model

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Zu} + \mathbf{e}, \quad [2]$$

where $\mathbf{y}$ is a vector of BLUEs of the training set obtained from the phenotypic analysis, $\mathbf{1}$ is a vector of 1s, μ is the population mean, $\mathbf{u}$ is a vector of random estimated genetic values with the distribution $\mathbf{u} \sim N(0, \mathbf{U}\sigma_g^2)$, and $\mathbf{Z}$ is the corresponding incidence matrix. $\mathbf{U}$ is the realized relationship matrix calculated on the basis of marker data following method 1 of VanRaden (32), and σ_g^2 is the genetic variance pertaining to the GBLUP model. The matrix $\mathbf{U}$ was calculated considering all genotypes (both population types and landraces) as one population. The vector of residuals $\mathbf{e}$ is assumed to be normally distributed with a mean of zero and equal variance [$\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$], where $\mathbf{I}$ is the identity matrix and σ_e^2 denotes the residual variance pertaining to the GBLUP model. The relationship matrices were calculated using R [version 3.6.0 (33)] and the R-package synbreed version 0.12-9 (34). Variance components pertaining to the GBLUP model were estimated using the R-package ASReml-R version 3.0 (21).

Genomic prediction accuracy (ρ) is reported as the correlation between predicted and unobservable true genetic values. Estimates of ρ were obtained from the Pearson correlation between the observed phenotypes and the estimated genetic values divided by the square root of h^2 of the prediction set (35).

Scenarios for Genomic Prediction. We studied the influence of the number of markers M and sample size N on ρ within populations by randomly sampling M markers using all genotypes from the respective population or sampling N lines without replacement from the population using all markers and carrying out 10 times fivefold cross-validation. The number of markers M was increased from 1,000 to 250,000. Sample size N was increased from 50 lines to the maximum possible number for the respective population in increments of 50. Sampling was repeated 100 times for each M and N. Prediction accuracy ρ was averaged across replications. The 95% quantile of ρ was calculated from the sampling replications. With small sample sizes N, ρ was set to "missing value" if the mixed model algorithm for a particular training set did not converge.

For comparing the prediction accuracy ρ within and between the DH and GC populations from the same landrace, $N = 200$ (PP) or $N = 75$ (TP) lines were sampled randomly from one population (either DH or GC) for training the model. The prediction set always comprised a disjoint set of $N = 50$ (PP) or $N = 25$ (TP) lines either from the same or from a different population. Sampling was repeated 100 times. The same sampling procedure was applied for investigating across landrace predictions using the same type of population of the other landrace as the prediction set.

Data Availability. Seeds from all genotypes used in the study are available through material transfer agreements. The genotypic data of the inbred line FV2, 873 DH lines, 544 GC lines, and 95 landrace plants and all corresponding phenotypic data of PP and TP have been deposited in Figshare (https://doi.org/10.6084/m9.figshare.17014421) (36).

ACKNOWLEDGMENTS. We thank the technical staff at KWS SAAT SE & Co. KGaA (KWS) and the Technical University of Munich (TUM) for their excellent support in carrying out the phenotypic evaluation required for this study; the TUM Plant Technology Center for providing infrastructure and technical support for field trials; Therese Bolduan and Tanja Rettig (KWS) for their contribution in developing the genetic material as well as planning and conducting field trials; the technical staff at KWS for DNA extraction; Hans Rudolf Fries (TUM) for processing the genotyping arrays; and Sylwia Schepella and Stefan Schwertfirm (TUM) for technical assistance. This study was funded by the Federal Ministry of Education and Research (BMBF, Germany) within the scope of the funding initiative Plant Breeding Research for the Bioeconomy (funding ID: 031B0195, project MAZE), by the Bavarian State Ministry of the Environment and Consumer Protection within the project network BayKlimaFit (project TGC01GCUFuE69741, "Improving cold tolerance in maize"), and by KWS through PhD fellowships (A.C.H. and M.M.).

Author affiliations: [a]Plant Breeding, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany; [b]KWS SAAT SE & Co. KGaA, 37574 Einbeck, Germany; [c]Campus Office, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany; and [d]Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany

1. S. McCouch *et al.*, Agriculture: Feeding the future. *Nature* **499**, 23–24 (2013).
2. W. L. Brown, Broader germplasm base in corn and sorghum in *Proc. Annu. Corn Sor. Res.* **30**, 81–89 (1975).
3. D. Hoisington *et al.*, Plant genetic resources: What can they contribute toward increased crop productivity? *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5937–5943 (1999).
4. R. Ortiz *et al.*, Conserving and enhancing maize genetic resources as global public goods–a perspective from CIMMYT. *Crop Sci.* **50**, 13–28 (2010).
5. M. Warburton *et al.*, Genetic diversity in CIMMYT nontemperate maize germplasm: Landraces, open pollinated varieties, and inbred lines. *Crop Sci.* **48**, 617–624 (2008).
6. A. Strigens, W. Schipprack, J. C. Reif, A. E. Melchinger, Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding. *PLoS One* **8**, e57234 (2013).
7. L. M. Pollak, The history and success of the public-private project on germplasm enhancement of maize (GEM). *Adv. Agron.* **78**, 46–89 (2003).
8. W. Salhuana, L. Pollak, Latin American Maize Project (LAMP) and Germplasm Enhancement of Maize (GEM) project: Generating useful breeding germplasm. *Maydica* **51**, 339–355 (2006).
9. J. A. Romero Navarro *et al.*, A study of allelic diversity underlying flowering-time adaptation in maize landraces. *Nat. Genet.* **49**, 476–480 (2017).
10. M. Mayer *et al.*, Discovery of beneficial haplotypes for complex traits in maize landraces. *Nat. Commun.* **11**, 4954 (2020).
11. A. C. Hölker *et al.*, European maize landraces made accessible for plant breeding and genome-based studies. *Theor. Appl. Genet.* **132**, 3333–3345 (2019).
12. J. Böhm, W. Schipprack, H. F. Utz, A. E. Melchinger, Tapping the genetic diversity of landraces in allogamous crops with doubled haploid lines: A case study from European flint maize. *Theor. Appl. Genet.* **130**, 861–873 (2017).
13. A. R. Hallauer, J. H. Sears, Integrating exotic germplasm into corn belt maize breeding programs 1. *Crop Sci.* **12**, 203–206 (1972).
14. A. R. Hallauer, M. J. Carena, J. de Miranda Filho, *Quantitative Genetics in Maize Breeding* (Springer Science & Business Media, ed. 6, 2010).
15. G. Gorjanc, J. Jenko, S. J. Hearne, J. M. Hickey, Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. *BMC Genomics* **17**, 30 (2016).
16. S. Sood, S. Flint-Garcia, M. C. Willcox, J. B. Holland, "Mining natural variation for maize improvement: Selection on phenotypes and genes" in *Genomics of Plant Genetic Resources*, R. Tuberosa, A. Graner, E. Frison, Eds. (Springer, Dordrecht, the Netherlands, 2014), pp. 615–649.
17. H.-J. Auinger *et al.*, Calibration and validation of predicted genomic breeding values in an advanced cycle maize population. *Theor. Appl. Genet.* **134**, 3069–3081 (2021).
18. M. Mayer *et al.*, Is there an optimum level of diversity in utilization of genetic resources? *Theor. Appl. Genet.* **130**, 2283–2295 (2017).
19. J. Ren *et al.*, Novel technologies in doubled haploid line development. *Plant Biotechnol. J.* **15**, 1361–1370 (2017).
20. L. J. Stadler, Gamete selection in corn breeding. *J. Am. Soc. Agron.* **36**, 988–989 (1944).
21. D. Butler, B. R. Cullis, A. Gilmour, B. Gogel, *ASReml-R Reference Manual. The State of Queensland* (Department of Primary Industries and Fisheries, Brisbane, QLD, Australia, 2009).
22. J. B. Holland, W. E. Nyquist, C. T. Cervantes-Martínez, "Estimating and interpreting heritability for plant breeding: An update" in *Plant Breeding Reviews*, I. Goldman, Ed. (Wiley-Blackwell, 2010), pp. 9–112.
23. C. Brien, asremlPlus: Augments "ASReml-R" in Fitting Mixed Models and Packages Generally in Exploring Prediction Differences. R package version 4.1-22 (2019). https://CRAN.R-project.org/package=asremlPlus. Accessed 13 April 2022.
24. S. Unterseer *et al.*, A powerful tool for genome analysis in maize: Development and evaluation of the high density 600 k SNP genotyping array. *BMC Genomics* **15**, 823 (2014).
25. Y. Jiao *et al.*, Improved maize reference genome with single-molecule technologies. *Nature* **546**, 524–527 (2017).
26. B. L. Browning, Y. Zhou, S. R. Browning, A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* **103**, 338–348 (2018).
27. W. G. Hill, A. Robertson, Linkage disequilibrium in finite populations. *Theor. Appl. Genet.* **38**, 226–231 (1968).
28. W. G. Hill, B. S. Weir, Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Popul. Biol.* **33**, 54–78 (1988).
29. P. Schopp, D. Müller, F. Technow, A. E. Melchinger, Accuracy of genomic prediction in synthetic populations depending on the number of parents, relatedness, and ancestral linkage disequilibrium. *Genetics* **205**, 441–454 (2017).
30. A. Jacobson, L. Lian, S. Zhong, R. Bernardo, Minimal loss of genetic diversity after genomewide selection within biparental maize populations. *Crop Sci.* **55**, 783–789 (2015).
31. L. Excoffier, P. E. Smouse, J. M. Quattro, Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491 (1992).
32. P. M. VanRaden, Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).
33. R Core Team, R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2019).
34. V. Wimmer, T. Albrecht, H.-J. Auinger, C.-C. Schön, synbreed: A framework for the analysis of genomic prediction data using R. *Bioinformatics* **28**, 2086–2087 (2012).
35. J. C. M. Dekkers, Marker-assisted selection for commercial crossbred performance. *J. Anim. Sci.* **85**, 2104–2114 (2007).
36. A. C. Hölker *et al.*, Theoretical and experimental assessment of genome-based prediction in landraces of allogamous crops. Figshare. 10.6084/m9.figshare.17014421. Deposited 13 April 2022.

6 Acknowledgments

I would like to express my deepest gratitude to my supervisor Prof. Dr. Chris-Carolin Schön for giving me the opportunity to work on this PhD project. Thank you for your constant support, patience, and guidance in research during the last years. During the PhD years I learned a lot from you and very much appreciated your honest, critical, and encouraging feedback.

I thank Prof. Dr. Tobias Würschum for serving on my graduate committee and Prof. Dr. Aurélien Tellier for chairing it.

Thanks go to all project partners and co-authors. Thank you for your good collaboration and the many interesting discussions. Special thanks go to Prof. Dr. Albrecht Melchinger who, in addition to his role as co-author and collaborator, originally convinced me to pursue a PhD at all. Special thanks also go to Dr. Milena Ouzunova and Dr. Thomas Presterl for their invaluable contribution to the development of the plant material and the experimental data, without which the project would not have been possible. In addition, I want to thank Dr. Milena Ouzunova, Dr. Thomas Presterl and Dr. Pascal Schopp also for their support and flexibility during the final stages of the thesis.

I would like to thank all former colleagues at the Chair of Plant Breeding for help and assistance in many situations and for creating a great working environment. Special thanks go to my office mate at the chair, Manfred Mayer, for his great support and invaluable input during our many discussions over the desk.

A big thanks goes to the involved technical staff at the Chair of Plant Breeding, KWS SAAT SE & Co. KGaA, the University of Hohenheim and CSIC. Without their help the generation of the phenotypic and genotypic data that are part of this thesis would not have been possible. A special acknowledgement goes to Sylwia Schepella, Stefan Schwertfirm and Tanja Rettig for their excellent work.

This research was made possible through financial support from the Federal Ministry of Education and Research (BMBF) within the scope of the funding initiative Plant Breeding Research for the Bioeconomy (Funding ID: 031B0195, project MAZE) through the Bavarian State Ministry of the Environment and Consumer Protection within the project network BayKlimaFit (Project TGC01GCUFuE69741 “Improving cold tolerance in maize”). In addition, I gratefully acknowledge the Ph.D. fellowship by KWS SAAT SE & Co. KGaA.

Finally, I want to thank my beloved family. I am grateful to my wife and my parents. Thank you for your love, encouragement and support in whatever I do. Without you I would not be where I am today.