

OPEN

QTL mapping and genome-wide prediction of heat tolerance in multiple connected populations of temperate maize

Delphine Van Inghelandt¹, Felix P. Frey^{2,3}, David Ries¹ & Benjamin Stich^{1,2}

Climate change will lead to increasing heat stress in the temperate regions of the world. The objectives of this study were the following: (I) to assess the phenotypic and genotypic diversity of traits related to heat tolerance of maize seedlings and dissect their genetic architecture by quantitative trait locus (QTL) mapping, (II) to compare the prediction ability of genome-wide prediction models using various numbers of KASP (Kompetitive Allele Specific PCR genotyping) single nucleotide polymorphisms (SNPs) and RAD (restriction site-associated DNA sequencing) SNPs, and (III) to examine the prediction ability of intra-, inter-, and mixed-pool calibrations. For the heat susceptibility index of five of the nine studied traits, we identified a total of six QTL, each explaining individually between 7 and 9% of the phenotypic variance. The prediction abilities observed for the genome-wide prediction models were high, especially for the within-population calibrations, and thus, the use of such approaches to select for heat tolerance at seedling stage is recommended. Furthermore, we have shown that for the traits examined in our study, populations created from inter-pool crosses are suitable training sets to predict populations derived from intra-pool crosses.

Maize (*Zea mays* L.), compared to other crop species which grow in temperate Europe, is heat tolerant due to its C4 metabolism and its tropical origin¹. Nevertheless, temperate maize cultivars can experience substantial damages when encountering heat stress². Especially during flowering and grain filling, heat stress has severe impacts on maize plants³. Phenotypic consequences of heat stress at adult stage on phenotypic traits of maize are, among others, a reduction of the time to flowering⁴, an increased anthesis-silking interval⁵, a reduction of photosynthetic tissue due to leaf scorching⁴, and a reduction of grain and whole plant yield⁶.

However, temperate maize is also significantly affected by heat events during seedling stage⁷. This is of particular practical importance for biogas production for which maize is the most important crop⁸ in temperate Europe. One of the cultivation practices is that the planting of biogas maize is postponed until the harvest of the winter cereals in early summer. With this cropping system, sensitive maize seedlings are exposed to heat stress⁹.

In the future, heat stress is expected to become an even more critical threat to crop cultivation in temperate regions than it is today¹⁰ as the mean temperature and the severity of heat events will rise due to climate change¹¹. Therefore, breeding heat-tolerant cultivars is crucial to sustain crop production in the future¹².

The tolerance to heat stress in maize was studied on a molecular level with a focus on natural variation by Ottaviano *et al.*¹³, Prova and Sari-Gorla¹⁴, and Reimer *et al.*⁹. These studies focused on isolated plant characteristics such as the cellular membrane stability, pollen germination, and root architecture. Alam *et al.*⁵ estimated variance components for traits involved in heat tolerance in field trials under natural heat stress condition. However, we are not aware of systematic studies characterising genetic variability for heat tolerance at seedling stage on a whole plant level.

The classical approach to improve a trait by breeding is to screen genetic material in one or several environments in which the conditions are such that the phenotypic trait of interest shows heritable variation. The

¹Institute for Quantitative Genetics and Genomics of Plants, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University, 40225, Düsseldorf, Germany. ²Max Planck Institute for Plant Breeding Research, 50829, Köln, Germany. ³Present address: Crop Functional Genomics, Institute of Crop Science and Resource Conservation, University of Bonn, Bonn, 53113, Germany. Delphine Van Inghelandt and Felix P. Frey contributed equally. Correspondence and requests for materials should be addressed to B.S. (email: benjamin.stich@hhu.de)

issue with an evaluation of heat tolerance in natural environments is that the timing and the strength of the heat event are typically unpredictable¹⁵. One way to circumvent this problem is to screen the genetic material of interest in an artificial environment such as greenhouses or growth chambers. The efficiency of such approaches can be increased by combining them with marker-assisted selection approaches. For many years, the markers for such approaches have been identified by quantitative trait loci (QTL) mapping or genome-wide association mapping. Although numerous QTL have been identified for maize (for review see Sehgal *et al.*¹⁶), the impact of marker-assisted selection for improving truly quantitative traits in maize breeding is limited¹⁷. This is primarily attributed to the small effects of many of the detected QTL. An alternative promising approach for such traits is genome-wide prediction (GWP) because it captures not only the variance of the QTL but also all genetic variance. However, to the best of our knowledge, such an approach has not been tested previously for traits related to heat tolerance.

Several studies have shown that moderate-to-high genomic prediction accuracies can be obtained in bi-parental populations for a trait with high heritability, even by using low marker density and a relatively small training population^{18–20}. Other studies have illustrated that genotypic characterisation using high-density genotyping platforms might improve the prediction accuracy^{21–23}. To the best of our knowledge, only few experimental studies till now have compared GWP based on low density genotyping and genotyping by sequencing (GBS)^{24–26}. Furthermore, the composition and size of the training and validation sets are crucial for GWP. Technow *et al.*²⁷ analysed the possibility of combining training sets across heterotic pools. However, no earlier study has examined the suitability of segregating populations derived from inter-pool crosses as training set for the prediction of the two original heterotic pools.

The objectives of this study were the following: (I) to assess the phenotypic and genotypic diversity of traits related to heat tolerance of maize seedlings and dissect their genetic architecture by quantitative trait locus (QTL) mapping, (II) to compare the prediction ability of genome-wide prediction models using various numbers of KASP (Kompetitive Allele Specific PCR genotyping) single nucleotide polymorphisms (SNPs) and RAD (restriction site-associated DNA sequencing) SNPs, and (III) to examine the prediction ability of intra-, inter-, and mixed-pool calibrations.

Methods

Plant material and phenotypic evaluation. This study was based on six segregating populations derived from pairwise crosses of four Dent (S067 = D_1 , P040 = D_2 , S058 = D_3 , S070 = D_4) and four Flint (L012 = F_1 , L017 = F_2 , L043 = F_3 , L023 = F_4) maize inbred lines from the University of Hohenheim²⁸. The eight inbred lines were previously characterised in detail for their heat tolerance at seedling stage⁷. The inbreds were crossed pairwise to create two Dent x Dent (DxD), two Flint x Flint (FxF), and two Dent x Flint (DxF) F_1 genotypes (Supplementary Fig. 1). The F_1 genotypes were further self-pollinated in an ear to row manner, resulting in six segregating populations ($P_{D_1D_2}$, $P_{D_3D_4}$, $P_{F_1F_2}$, $P_{F_3F_4}$, $P_{D_1F_1}$, $P_{D_4F_4}$) comprising between 75 and 107 $F_{3:4}$ progenies with a total of 607 genotypes.

Seeds were sown in soil (50% ED73, 50% Mini Tray (Einheitserde- und Humuswerke, Gebr. Patzer GmbH & Co. KG, Sinntal-Altengronau, Germany)) in single pots (9 cm edge length) as described previously. The experiment was replicated three times. The experimental design was a lattice design comprising 32 incomplete blocks per replication which were distributed on four tables in a walk-in growth chamber (Bronson Incubator Services B.V., Nieuwkuijk, Netherlands). The parental inbred lines were included as checks once on each table.

The plants were grown at 25 °C during a 16 h light period and at 20 °C during a 8 h dark period for a total of three weeks in the growth chamber; the relative humidity was set to 60% during this period. Photosynthetic active radiation emitted by fluorescent tubes was between 270–280 $\mu\text{mol m}^{-2} \text{s}^{-1}$ in the canopy of the plants. The plants were watered every morning with an automatic irrigation system (Itec DC station Multi Program, I.T. Systems Ltd., Bazra, Israel) to avoid drought stress.

The leaf growth rate was assessed as follows: fourteen days after sowing, the length of the fourth leaf from the shoot base to the leaf tip was measured daily for a period of three days during the stage of linear growth. The slope of a linear trend line of leaf length measurements vs. time represented the leaf growth rate (LR). Twenty days after sowing, leaf greenness (SD) (SPAD-502, Minolta Corporation, Ramsey, NJ, USA) was assessed as the maximum value of four readings on the leaf blade of the latest fully developed leaf. Furthermore, leaf scorching of young leaves (SC) and leaf senescence of old leaves (SN) were recorded on a scale of 1 (weak damage) to 9 (strong damage). The length of the fourth leaf (LL), the plant height (PH) from the shoot base to the point where the youngest leaf detached from the older leaf's sheath, and the number of leaves (NL) with visible leaf ligule were recorded per plant. 21 days after sowing, shoot dry weight (DW) and the shoot water content (WC) of the fresh material were determined. The above outlined experiment was repeated at a higher heat level, where the temperature was increased, six days after planting, to 38 °C at day and 33 °C at night to induce heat stress for a two-week period. The study was, thus, based on two experiments with different heat levels, which will be referred to hereafter as standard and heat conditions.

Genotyping. *KASP SNPs.* The parental inbred lines of the segregating populations were genotyped using a 50 K SNP array²⁹. Out of 56,110 SNPs, 170 SNP markers were selected to genotype the individuals of the six segregating populations. For each population, between 47 and 77 markers were chosen (60 for $P_{D_1D_2}$, 47 for $P_{D_3D_4}$, 75 for $P_{F_1F_2}$, 64 for $P_{F_3F_4}$, 67 for $P_{D_1F_1}$, and 77 for $P_{D_4F_4}$) as they were polymorphic between the two parental inbreds of each population and showed no heterozygosity in any of the parental inbreds. SNP marker selection was optimised for equal distribution across the physical map (due to the unavailability of a genetic map at that time) and the overlapping of markers between populations. The selected SNP markers were genotyped using KASP SNP technology by TraitGenetics GmbH (Gatersleben, Germany) on a bulk of 6 to 10 $F_{3:4}$ plants per genotype in the respective populations. This data set with 170 SNPs on 607 progenies will be designated hereafter as KASP₆₀₇.

RAD SNPs. The RAD libraries of the segregating populations were prepared for single-end sequencing according to Baird *et al.*³⁰ with the following modifications: barcodes were 5 bp long, and were at least two mutational steps apart from each other with regard to the first four bases, followed by the fifth checksum base. A total of 2 μg genomic DNA of the same pool of plants that was used for KASP genotyping was digested for 30 min at 37 °C in a 50 μl reaction with 20 units of *KpnI* (New England Biolabs). Samples were inactivated by purification with Qiaquick spin columns (Qiagen, Hilden, Germany). Libraries were 96-fold barcoded, each genotype at two barcodes, resulting in 7 libraries. The sequencing of the RAD libraries was performed on a HiSeq2000 with 100 bp single end reads by the Max Planck-Genome-centre Cologne, Germany (<https://mpgc.mpipz.mpg.de/home/>), following the manufacturer's protocol.

Demultiplexing of raw sequencing data by barcode was performed using the Stacks software pipeline³¹. Parameters were chosen to allow barcode rescue with a distance of two, where reads were discarded according to default settings. In the next steps, version 0.4.2 of Trim Galore! (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) was used for adapter and quality trimming. In addition to the standard parameters, reads were filtered for a length of at least 75 bp, and the threshold for trimming low quality ends from reads was increased to 30 for higher stringency.

The trimmed reads were mapped genotype-wise to the repeat masked and concatenated reference sequence, using BWA-MEM. In accordance to previous studies³², apart from an increased sensitivity parameter of -r 1.0, standard parameters were used.

For genotype calling, GATK's HaplotypeCaller³³ was applied to each genotype's bam file independently, where the allowed maximum number of alternative alleles was set to three. The minimum base and mapping quality for calling were set to 30 and 40, respectively. Finally, soft-clipped bases were excluded from the analysis. All other parameters were set to their default values. Finally, the resulting files were combined into one and each position was recalculated and re-genotyped, considering information of all samples, by GATK's genotypeGVCFs. To ensure a high quality of the genotype calls, many filtering steps (see SM1) were applied to the resulting genotype call file.

After genotype calling, 53,579 SNP loci remained on 489 genotypes comprising 482 progenies and 7 parental inbred lines (Supplementary Fig. 1). Missing genotype calls were imputed per population using Beagle³⁴. Three cutoff values were chosen to filter per genotype the genotype calls based on the genotype probabilities (GP). For each of the three GP cutoff values, one of the three possible genotypes calls (minor allele homozygous, heterozygous, and major allele homozygous) was assigned to a genotype only when its predicted probability was greater than X for a given imputed genotype call, and all others were set to "NA". The three cutoff values were X = 1, 0.98³⁵, and the allele with maximum genotype probability (MAX). The such created missing genotype calls were mean imputed. The data sets associated with the three cutoff values will be designated hereafter as RAD_{482-GP,1}, RAD_{482-GP,0.98}, and RAD_{482-GP,MAX}, respectively. As there were less genotypes in the RAD than in the KASP data set, we defined an additional KASP data set with the same 482 progenies as for the RAD data set. This will be referred to hereafter as KASP₄₈₂.

Diversity. Gene diversity D³⁶ and population differentiation G_{ST}³⁷ were calculated. The average modified Roger's distances (MRD) between and within populations were calculated according to Wright *et al.*³⁸. Associations among genotypes were revealed with a principal component analysis (PCA) based on MRD estimates.

Phenotypic data analysis. *Adjusted entry mean calculation.* To assess the significance of the heat level effect, the following model was used:

$$Y_{icrb} = \mu + G_i + C_c + (G.C)_{ic} + R_{cr} + B_{crb} + e_{icrb} \quad (1)$$

where Y_{icrb} is the phenotypic observation of the i^{th} genotype in the b^{th} block within the r^{th} replication nested within the c^{th} heat level. μ is the general mean, G_i is the effect of the i^{th} genotype, C_c is the effect of the c^{th} heat level, $(G.C)_{ic}$ is the interaction between the i^{th} genotype and the c^{th} heat level, R_{cr} is the effect of the r^{th} replication nested within the c^{th} heat level, B_{crb} is the effect of the b^{th} block in the r^{th} replication nested within the c^{th} heat level and e_{icrb} the residual. C_c was regarded as fixed and all other effects were set as random. Traits with a significant C_c effect were considered heat-dependent traits. The adjusted entry means (AEM) of standard and heat conditions across genotypes were estimated and will be referred to as AEM_s and AEM_H, respectively.

To calculate AEM for each assessed trait of each genotype in each condition (AEM_{IS} and AEM_{IH} for standard and heat condition, respectively), the phenotypic observations of each heat level were separately analysed using the following model:

$$Y_{irb} = \mu + G_i + R_r + B_{rb} + e_{irb} \quad (2)$$

where Y_{irb} is the phenotypic observation of the i^{th} genotype in the r^{th} replication and the b^{th} block, R_r is the effect of the r^{th} replication, B_{rb} is the effect of the b^{th} block nested within the r^{th} replication, and e_{irb} the residual. The genotype effect G_i was of primary interest in this analysis and was considered a fixed effect. R_r was considered fixed and B_{rb} random.

Heat susceptibility index. A heat susceptibility index (HSI) for each trait and each genotype was calculated according to Mason *et al.*³⁹:

$$\text{HSI}_i = \frac{1 - \text{AEM}_{IH}/\text{AEM}_{IS}}{1 - \text{AEM}_H/\text{AEM}_S} \quad (3)$$

where HSI_i is the HSI for genotype i . The HSI of the individual traits will be designated as follows: HSI_{LL} , HSI_{PH} , HSI_{NL} , HSI_{SC} , HSI_{SN} , HSI_{SD} , HSI_{DW} , HSI_{WC} , and HSI_{LR} . Pairwise Pearson correlation coefficients were calculated between the HSI of all assessed traits across all genotypes. Further, HSI of this study were correlated with HSI assessed during adult stage under field conditions⁴.

Variance components and heritability. Genotypic (σ_g^2) and error (σ_e^2) variance components for each heat level were calculated using model (2) with a random G_i effect. For each trait, the broad sense heritability (h^2)⁴⁰ of the observations at each heat level was calculated separately for each (h_{pop}^2) as well as across (h_a^2) the six populations. Heritability can not be assessed for HSI directly as it is not measured but calculated on a mean basis with no replication³⁹. To calculate an approximate heritability for the HSI of each trait, we applied the following procedure, which is an extension of the approach of Ouk *et al.*⁴¹. For each of the two heat levels, we estimated, based on model (2), the fixed effects for R_i and B_{rb} . In a second step, these effects were subtracted from all observations from the corresponding replicate and block to correct for differences in replicate and block effects. The adjusted data thereby comprised three replications per genotype for each of the two conditions. Subsequently, we separately created random pairs for each genotype between one replication from the heat condition and one from the standard condition. Across all genotypes, 3^{2607} combinations of replications from the heat and standard conditions are possible. In the next step, for one possible combination of replications, the HSI was calculated as described previously, resulting in three HSI replicates for each genotype. These data were then analysed using the following model:

$$Y_{ir} = \mu + G_i + e_{ir}, \quad (4)$$

where Y_{ir} is the phenotypic observation of the i^{th} genotype in the r^{th} replication, μ is the general mean, G_i is the effect of the i^{th} genotype, and e_{ir} the residual. Based on genotypic (σ_g^2) and error (σ_e^2) variance components across all populations and for each population, the heritability was calculated. This procedure was repeated 100 times, and the median of the heritability was designed as heritability of HSI (h_{HSI}^2).

The genetic variance among and within populations was estimated for each HSI by partitioning the genotype effect of model (4) into the effect of the population and that of the genotype nested within the population. This procedure was repeated 100 times with random assignments between the different replications as explained for the h_{HSI}^2 calculation, and the median of the genetic variance among and within populations was calculated. The genetic differentiation observed for the HSI was then calculated⁴²⁻⁴⁴ and will be designed hereafter as Q_{ST} .

QTL analyses with KASP SNPs. SNP markers with a significant ($P < 0.001$) observed deviation from the expected allele frequency were excluded from the analysis (cf. Benke *et al.*⁴⁵). To improve the mapping of markers, information of five additional segregating populations which had been genotyped with the same set of SNP markers by Horn *et al.*⁴⁶ were considered during map creation. A consensus genetic linkage map was calculated chromosome-wise using the software CarthaGène⁴⁷.

QTL associated with heat tolerance were detected based on the consensus genetic linkage map using an iterative composite interval mapping approach⁴⁸, implemented in the software MCQTL^{49,50}. QTL analyses were performed for the HSI of the individual traits across all six populations. We took into account connections between populations using an additive kinship matrix. As the studied bi-parental $F_{3:4}$ populations have an expected heterozygosity of 25%, dominance effects between parental alleles of each bi-parental population were considered in the QTL analysis. Genotypic probabilities were computed every 5 centiMorgan (cM), taking into account information from neighboring markers. F thresholds to detect QTL for each trait were determined by 1 000 permutations, to adhere to a global type I error of 5% across populations and the entire genome⁵¹. F thresholds used to select co-factors were fixed at 90% of the F threshold values for QTL detection.

SNP markers associated with the respective HSI were selected as co-factors by forward regression, where the minimal distance between two co-factors was 10 cM. At the end of the detection process, confidence intervals were estimated on the basis of a 1.5 LOD unit fall⁵².

The dominance effect of each QTL was tested for its significance ($P < 0.05$) in each population using a two-sided t-test. The difference between the additive effects of pairs of parental alleles on the respective QTL was tested a posteriori using a Tukey test.

Furthermore, we overlapped the QTL detected in this study with those identified in a companion study at adult stage⁴. To identify candidate genes for heat tolerance in terms of the assessed traits, we mined genes, identified by Frey *et al.*⁷ as heat responsive or heat tolerance genes. We determined the position of these genes on the genetic map by linear regression with information of the nearest two SNP markers. Heat tolerance genes of Frey *et al.*⁷ mapping to QTL confidence intervals were designated in the following as heat tolerance candidate genes.

GWP. *Genetic model.* GWP was performed by genomic best linear unbiased prediction (GBLUP⁵³):

$$y = \mathbf{X}\beta + \mathbf{Z}u + e, \quad (5)$$

where y is the vector of the HSI of the corresponding trait, \mathbf{X} is a design matrix assigning fixed effects to the genotypes, β is a vector of fixed effects, \mathbf{Z} is the design matrix that assigned the random effects to the genotypes, and u the vector of the random effects that were assumed to be normally distributed with: $u \sim N(0, \mathbf{K}\sigma_u^2)$, where \mathbf{K} denotes the realised kinship matrix calculated on the basis of the molecular marker data⁵⁴ with σ_u^2 being the variance pertaining to the GBLUP model. Residuals e were assumed to be independent and normally distributed with $e \sim N(0, \mathbf{I}\sigma^2)$, where \mathbf{I} is the identity matrix and σ^2 the residual variance.

Two different genetic models were examined for each type of molecular marker that differed in the \mathbf{K} matrix used: M_A considering only additive effects with \mathbf{K}_A (the additive kinship matrix) and M_{AD} considering additive and dominance effects with \mathbf{K}_A and \mathbf{K}_D (the dominance kinship matrix)⁵⁵. GBLUP method was used as implemented in the R package Sommer⁵⁶.

Model training and performance assessment. The effects were estimated in a training set (TS) and used in a second step to predict the breeding values of the genotypes of a validation set (VS). As measure of model performance, prediction ability was calculated as the Pearson correlation coefficient between observed and predicted phenotypes of the VS, $r(y, \hat{y})$ ⁵⁷. In our study, different TS were used to establish the prediction model: in the TS_{pop} scenario, the model was trained using one or two populations; for the TS_a scenario all populations were used together as TS. For later scenario, prediction abilities were calculated both across all populations (r_a) as well as for each population (r_{pop}).

Resampling schemes. Different types of resampling schemes were used to evaluate the model performance as a function of the examined TS*VS combination:

- For TS_a , we performed 20 replications of five-fold cross validation (CV) to assess the model performance across all populations. Accordingly, the entire data set was subdivided into five mutually independent subsets where four formed the training set (TS) and one the validation set (VS). To account for the structure of our data set with six different segregating populations, the proportion of genotypes from the different segregating populations in the individual subsets was kept identical to that observed in the entire data set. The median of the Pearson correlation coefficient between observed and predicted phenotypes of the validation set $r(y_{VS}, \hat{y}_{VS})$ was used to assess model performance. Prediction abilities were calculated both across all populations ($r_{a,cv}$) as well as for each population ($r_{pop,cv}$).
- In the context of prediction within populations with TS_{pop} , the influence of the size of the training set on the prediction ability was analysed by simulating, for each segregating population, a TS size N from 10 to 100 in steps of 10 using a bootstrapping approach. The median of the prediction ability across 500 replications was calculated for the HSI of each trait, genetic model, and type of SNP marker within each population and is designated in the following as r_{pop,BSP_n} .
- For the between population prediction using TS_{pop} , the influence of the composition of the TS on the prediction ability was examined using three primary scenarios of TS compositions TS1, TS2, and TS3. In case of using the FxF and DxD populations as VS, TS1 comprised genotypes from populations for which the parental genotypes were from the same heterotic pool as those of the VS. TS2 comprised genotypes from one (TS_2_s) or two ($TS2_c$) populations in which the parental genotypes belonged to the opposite heterotic pool as those of the VS. TS3 comprised genotypes from the DxF populations with three different scenarios: $TS3_{sr}$ used the related mixed population, $TS3_{su}$ used the unrelated mixed population, and $TS3_c$ combined both mixed populations in the TS (Fig. 1, bottom).

In case of using the DxF populations as VS, the three scenarios were selected accordingly: TS1 comprised genotypes of the other DxF segregating population; TS2 comprised genotypes of the two combined FxF and DxD populations which did not have a parental inbred in common with the population in the VS. TS3 comprised genotypes of the two combined FxF and DxD populations which had a parental inbred in common with the population in the VS. To avoid differences in the prediction ability due to sample size effects, the TS were randomly reduced to a size of 50 such that the structure of the original TS was maintained. To obtain stable estimates for the prediction ability, the median of 500 independent bootstrapping runs of the TS construction was calculated for the HSI of each trait, genetic model, type of molecular marker, and TS*VS combination. The prediction abilities based on this bootstrapping strategy between populations will be referred to hereafter as r_{pop,BSP_b} .

Number of molecular markers. In order to examine the influence of the number of molecular markers on the prediction ability, X random RAD SNPs were sampled from the $RAD_{482-GP.0.98}$, where X ranged from 170 to 43 520 in steps of $2 * X$. Based on the selected SNPs, genomic predictions were obtained for the HSI of each trait using the M_A genetic model, and the median of the prediction abilities across 100 replications was calculated.

Comparison between observed and expected within-population prediction ability. We used the formula suggested by Daetwyler *et al.*⁵⁸ to estimate the expected within-population prediction ability. Following Meuwissen *et al.*⁵⁹, M_e was estimated as $M_e = 2N_eL$ where L is the genome size in Morgans. $L = 22.36$ was adopted from a previous linkage-mapping study, using an F_2 population which was characterised by genotyping by sequencing⁶⁰ similarly to our study. The effective population size (N_e) was calculated using the harmonic mean approximation for two generations (Hartl and Clark, p. 291⁶¹), resulting in $Ne = 2.92, 2.93, 2.94, 2.93, 2.92$, and 2.92 for populations $P_{D_1D_2}, P_{D_3D_4}, P_{F_1F_2}, P_{F_3F_4}, P_{D_1F_1}$, and $P_{D_4F_4}$, respectively.

Results

Phenotypic diversity. The condition effect (standard vs. heat) was significant ($P < 0.01$) across all populations and for all traits. Across the two examined conditions, h_a^2 was medium to high (0.50–0.83) for all assessed traits (Table 1). For all traits except NL and SD, h_a^2 was higher at heat compared to standard condition. h_{HSI}^2 ranged from 0.41 (NL) to 0.75 (SC) and was with the exception of SC, WC, and LR lower than those of the traits in the two conditions. The heritability values observed on a population level (h_{pop}^2) were very low for a few trait*population*condition levels (Supplementary Table 1). This effect was the strongest for SC for which, at standard

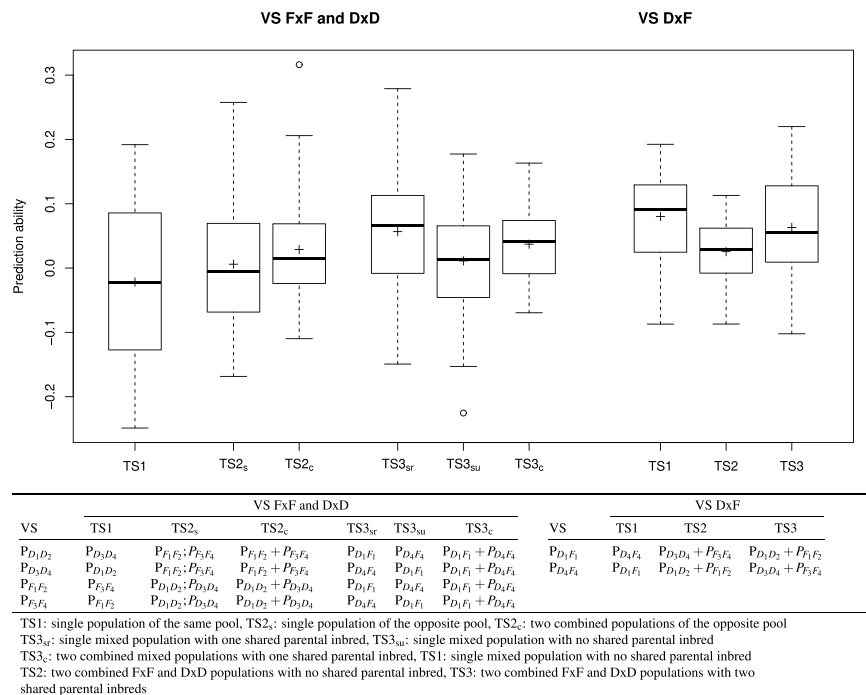


Figure 1. Boxplot of the between-population prediction abilities r_{pop, BSP_b} across nine heat susceptibility indexes (HSI) and different validation sets (VS) using a bootstrapping procedure with 50 genotypes for three different types of training set (TS) compositions. The analyses were based on RAD_{482-GP.0.98} and the M_A model.

Trait	Standard condition (25 °C)			Heat condition (38 °C)			HIS h_{HSI}^2	Condition effect	Q_{ST}
	Mean (AEM)	Range (AEM)	h_a^2	Mean (AEM)	Range (AEM)	h_a^2			
Leaf length	59.69	35.40–84.20	0.73	41.90	12.38–59.00	0.79	0.56	***	0.18
Plant height	18.52	9.45–24.18	0.58	12.99	2.86–20.00	0.73	0.52	**	0.16
Number of leaves	3.47	2.44–4.35	0.71	4.48	2.41–6.01	0.63	0.41	***	0.23
Leaf scorching	1.05	1.00–6.34	0.49	3.51	0.78–9.18	0.82	0.75	***	0.37
Leaf senescence	1.80	0.75–6.60	0.57	3.36	1.07–8.38	0.70	0.50	***	0.07
Leaf greenness	50.09	37.38–63.87	0.83	38.79	23.44–53.44	0.67	0.64	***	0.25
Shoot dry weight	0.91	0.31–1.54	0.69	0.70	0.12–1.64	0.76	0.67	***	0.17
Shoot water content	0.93	0.90–0.95	0.54	0.89	0.81–0.93	0.59	0.56	***	0.26
Leaf growth rate	0.25	0.12–0.39	0.50	0.22	–0.01–0.37	0.70	0.54	***	0.17

Table 1. Mean and range of the adjusted entry mean (AEM), broad sense heritability (h_a^2) of the studied traits for each condition, as well as for the heat susceptibility index (h_{HSI}^2), significance of the environmental condition (standard vs heat) effect for each trait, and degree of genetic differentiation among the six populations (Q_{ST}). *, **, *** Significant at the 0.05, 0.01, and 0.001 probability level, respectively.

condition, a h_{pop}^2 of 0 was observed for four of the six populations and smaller than 0.25 for the remaining two populations.

For LL, SC, DW, and LR, the mean and variation of the HSI were significantly ($P < 0.05$) higher in the two DxX populations ($P_{D_1D_2}$ and $P_{D_3D_4}$) than in the FxF and DxX populations. Especially, the population $P_{D_3D_4}$ differed significantly from the other populations in the mean HSI for all traits except SD (Supplementary Fig. 2). Q_{ST} calculated for the HSI varied from 0.07 (SN) to 0.37 (SC; Table 1).

The first two principal components (PC) of the PCA of the HSI explained 45% and 14% of the total variance, respectively (Supplementary Fig. 3). PC1 was significantly ($P < 0.01$) correlated with each HSI where the correlation coefficient was low ($<|0.25|$) for HSI_{SD} and between 0.36 and 0.85 for the other HSI (Supplementary Fig. 4). PC2 was significantly correlated with the HSI of each trait except HSI_{NL} and HSI_{WC} where the correlation coefficient was low ($<|0.25|$) for all HSI except for HSI_{SC} and HSI_{SD}. The cluster of the two DxX populations overlapped only weakly with the cluster of the two FxF populations (Supplementary Fig. 3).

Genetic diversity. The consensus genetic linkage map for KASP SNPs had a total length of 1 823.5 cM with an average distance of 11.3 cM and a maximum distance of 83.2 cM between two adjacent markers.

Trait	QTL	Chr	Pos	LOD	Interval	R ²	Additive effect of parent								Dominance effect of population					
							D ₁	D ₂	D ₃	D ₄	F ₁	F ₂	F ₃	F ₄	P _{D₁D₂}	P _{D₃D₄}	P _{F₁F₂}	P _{F₃F₄}	P _{D₁F₁}	P _{D₄F₄}
HSI _{LL}	Q _{HSI_{LLa}}	2	37	9.64	0–42	0.07	0.08 ^B	-0.11 ^C	0.07 ^{AB}	-0.06 ^{AC}	0.02 ^{ABC}	0.01 ^{ABC}	0.03 ^{ABC}	-0.04 ^{AC}	-0.00 ^{ns}	-0.23 ^{ns}	-0.02 ^{ns}	-4.40 ^{ns}	-0.03 ^{ns}	0.78 ^{ns}
	Q _{HSI_{LLb}}	10	30	10.55	14–38	0.07	-0.00 ^{AB}	-0.05 ^{AB}	-0.03 ^{AB}	-0.10 ^B	0.03 ^A	0.02 ^{AB}	0.06 ^{AB}	0.07 ^A	0.00 ^{ns}	-0.06 ^{ns}	-0.19 ^{**}	0.05 ^{ns}	-0.07 ^{ns}	0.29 ^{ns}
	Simultaneous fit					0.13														
HSI _{PH}	Q _{HSI_{PH}}	9	3	13.20	0–17	0.09	-0.11 ^{BD}	0.01 ^{ABCD}	0.10 ^A	-0.13 ^{CD}	0.05 ^A	0.05 ^{AB}	0.01 ^{ABCD}	0.03 ^{AC}	0.10 ^{ns}	-0.00 ^{ns}	0.12 ^{ns}	0.20 ^{ns}	0.05 ^{ns}	0.15 ^{ns}
HSI _{SC}	Q _{HSI_{SC}}	9	74	8.88	0–74	0.07	-0.33 ^A	0.50 ^A	-0.02 ^A	0.03 ^A	-0.35 ^A	0.18 ^A	-0.02 ^A	0.01 ^A	1.34 ^{***}	1.05 ^{**}	-1.19 ^{ns}	0.29 ^{ns}	-2.51 ^{ns}	-0.28 ^{ns}
HSI _{SD}	Q _{HSI_{SD}}	5	101	10.06	70–151	0.08	0.10 ^B	-0.27 ^A	-0.22 ^A	-0.23 ^A	-0.01 ^{AB}	0.18 ^B	0.26 ^B	0.19 ^B	-0.17 ^{ns}	0.13 ^{ns}	-0.16 ^{ns}	-0.58 ^{ns}	-0.00 ^{ns}	-0.11 ^{ns}
HSI _{LR}	Q _{HSI_{LR}}	2	82	10.06	60–141	0.07	-0.24 ^{AB}	-1.49 ^C	0.12 ^{AB}	-0.06 ^{AB}	1.00 ^B	0.73 ^{AB}	0.13 ^{AB}	-0.19 ^A	0.99 ^{ns}	-3.63 ^{ns}	0.24 ^{ns}	-0.69 ^{ns}	-0.87 ^{ns}	0.66 ^{ns}

Table 2. Quantitative trait loci (QTL) detected for the heat susceptibility index (HSI) of five traits (Leaf length: LL, Plant height: PH, Leaf scorching: SC, Leaf greenness: SD, Leaf growth rate: LR) at a significance level of $P < 0.05$, with genetic map position [cM], logarithmic odds ratio (LOD) and their support interval, proportion of explained phenotypic variance (R^2), additive effects of each parental genotype and dominance effects of the six populations. A, B, C, D Additive effects of parental inbreds with the same letters are not significantly ($P < 0.05$) different from each other. *, **, *** Significant at the 0.05, 0.01 and 0.001 probability level, respectively. ^{ns}Not significant.

In the PCA based on MRD estimates calculated from KASP₆₀₇, the first and second PC explained 21.87% and 11.87% of the molecular variance, respectively (Supplementary Fig. 5). For RAD_{482-GP.0.98}, PC1 and PC2 explained 25.94 and 13.50% of the molecular variance. In the PCA based on KASP₆₀₇, five distinct clusters were observed where one cluster was always constituted by the individuals of one segregating population except the individuals of the two Fx_F populations which were located in one overlapping cluster. For the RAD SNPs, the same trend was observed, but the two Fx_F populations (P_{F₁F₂} and P_{F₃F₄}) were assigned to two distinct clusters.

The lowest gene diversity was observed for P_{D₃D₄} irrespective of the considered marker type (Supplementary Table 2). The ranking between the populations for D and G_{st} was different when calculated based on KASP or RAD SNPs. The correlation between the MRD distance matrices calculated with KASP and RAD SNPs was with 0.84 significantly ($P = 1.67 \times 10^{-4}$) different from 0. The correlation between the K_A matrices calculated with KASP and RAD SNPs was with 0.56 considerably lower but also significantly ($P = 6 \times 10^{-4}$) different from 0.

QTL mapping and overlapping region with QTL detected at adult stage. We identified a total of six QTL for the HSI of five of the nine traits (Table 2), each explaining between 7% and 9% of the phenotypic variance (R^2). The detected QTL were not randomly distributed across the genome, but a total of three QTL hot spots were observed. The QTL detected by Frey *et al.*⁴ for adult traits in field trials collocated with these hotspots (Fig. 2). Five of the heat tolerance genes identified by Frey *et al.*⁷ were located within the six QTL confidence intervals detected in our study (Supplementary Table 3).

GWP. The square root of the proportion of phenotypic variance explained by the QTL for the HSI of all traits was lower than the prediction abilities of the GWP models, irrespective of the type of molecular marker and genetic model used (Table 3). Furthermore, the correlation between R_{QTL} and r_a across the nine examined traits for KASP₆₀₇ was approximately 0.33.

For the HSI of all traits, prediction abilities r_a were equal or lower for KASP₄₈₂ than for KASP₆₀₇ (Table 3). r_a calculated based on RAD₄₈₂ were consistently higher than those for KASP₄₈₂, independent of the GP cutoff value which was used to filter the genotype matrices. Despite this mean difference in the prediction abilities r_a of the different data sets and genetic models, the trend across the nine traits was the same. The prediction abilities $r_{a,cv}$ varied for KASP₆₀₇ between 0.31 (SN) and 0.70 (SC), whereas they ranged from 0.34 (SN) to 0.72 (SC) for RAD_{482-GP.0.98} (Fig. 3 and Supplementary Fig. 6). We observed a higher $r_{a,cv}$ for RAD_{482-GP.0.98} than for the two other GP cutoff values (data not shown) and, therefore, used the former data set for all further analyses. The correlation between Q_{ST} and r_a calculated for KASP₄₈₂ was with 0.82 significantly different from 0 ($P = 0.0068$). In contrast, the correlation between Q_{ST} and r_a calculated based on RAD_{482-GP.0.98} was 0.19 and not significant ($P = 0.62$).

Based on KASP₄₈₂, the prediction ability across populations r_a was, apart from some exceptions, higher than the prediction ability calculated for each population r_{pop} (Supplementary Fig. 6). The opposite trend was observed for RAD_{482-GP.0.98} (Fig. 3). For both molecular marker types, the prediction ability $r_{a,cv}$ of the M_A and M_{AD} genetic model did not differ significantly. However, the within-population prediction ability r_{pop,BSP_w} of the individual population*trait combinations, was either similar or higher for the M_A compared to the M_{AD} model (Supplementary Figs 7–12). Therefore, we focused in this study on the former model.

An increase in the prediction ability with an increasing number of molecular markers was observed (Fig. 4). The number of markers for which the prediction ability reached a plateau differed between the examined traits.

An increase of the within-prediction ability with an increasing size of the TS was observed for most population*trait*genetic model*molecular marker type combinations (Supplementary Figs 7–12). With a few exceptions, which showed a linear increase, the detected increase followed a logarithmic trend line (Supplementary Figs 13–15). We observed within-population prediction abilities for three traits (WC, PH and SD) that were similar to or higher than the expected abilities (Fig. 5). This was not true for the other traits. Especially, NL was predicted worse than expected. Furthermore, two populations P_{D₁D₂} and P_{D₄F₄} were predicted better than expected, whereas the other four were predicted worse than expected. The within-population prediction abilities r_{pop,BSP_w}

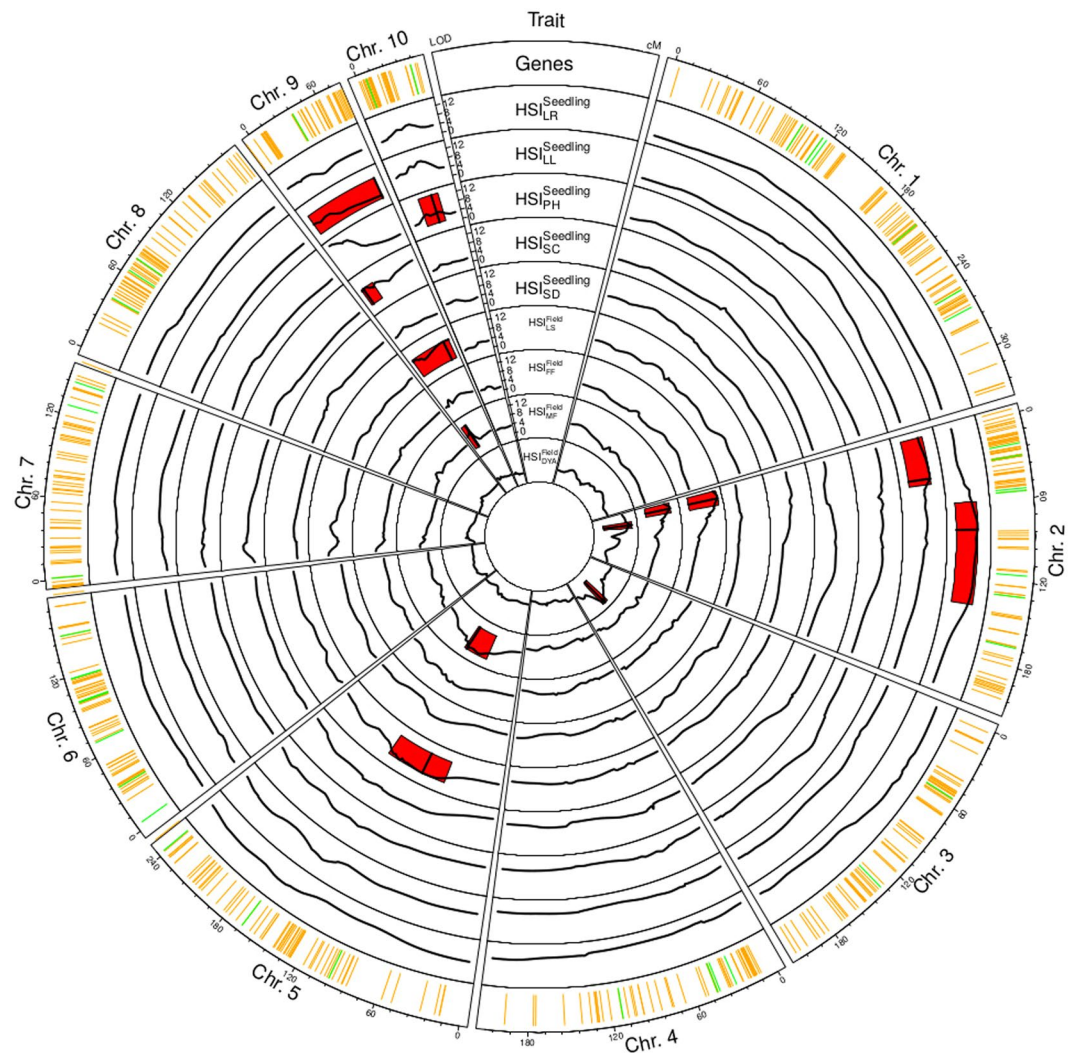


Figure 2. Circle plot showing the location of quantitative trait loci (QTL) affecting heat tolerance of maize. Heat tolerance (green) and heat responsive (orange) candidate genes⁷ are represented in the first track. Tracks 2–10 show logarithmic odds ratio (LOD) scores (black), detected QTL and confidence intervals (red) of the QTL for the heat susceptibility indexes (HSI) of the traits: leaf elongation rate (LR), leaf length (LL), plant height (PH), leaf scorching (SC) and leaf greenness (SD) at seedling stage, and leaf scorching (LS), time to female (FF) and male flowering (MF) and adjusted dry yield (DYA) at adult stage⁴.

Data set	R_{QTL}	r_a	r_a	r_a	r_a	r_a
	KASP ₆₀₇	KASP ₆₀₇	KASP ₄₈₂	RAD _{482-GP-MAX}	RAD _{482-GP-0.98}	RAD _{482-GP-1}
Trait						
HSI _{LL}	0.36	0.66	0.64	0.88	0.92	0.93
HSI _{PH}	0.30	0.64	0.63	0.88	0.94	0.94
HSI _{NL}	0.00	0.58	0.58	0.69	0.70	0.70
HSI _{SC}	0.26	0.80	0.81	0.85	0.88	0.88
HSI _{SN}	0.00	0.56	0.52	0.68	0.71	0.76
HSI _{SD}	0.28	0.69	0.66	0.80	0.90	0.83
HSI _{DW}	0.00	0.63	0.62	0.88	0.93	0.94
HSI _{WC}	0.00	0.65	0.68	0.89	0.88	0.86
HSI _{LR}	0.26	0.71	0.70	0.89	0.92	0.93

Table 3. Square root of the proportion of the explained phenotypic variance of the QTL (R_{QTL}), genome-wide prediction ability based on the different KASP and RAD SNP sets for the heat susceptibility index (HSI) of nine traits (leaf length (LL), plant height (PH), number of leaves (NL), leaf scorching (SC), leaf senescence (SN), leaf greenness (SD), shoot dry weight (DW), shoot water content (WC) and leaf growth rate (LR)) across the six populations (r_a) using the additive genetic model M_A .

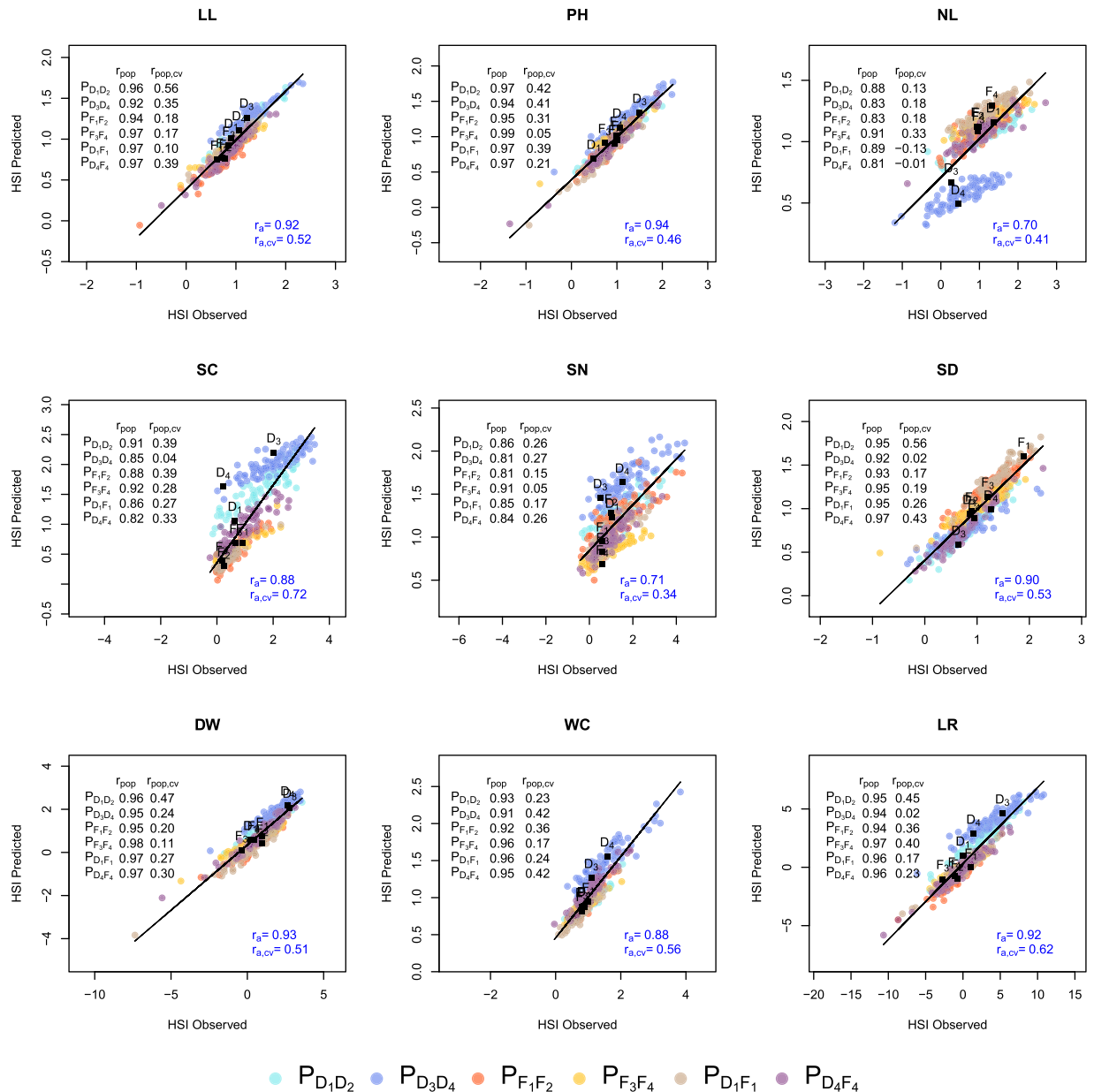


Figure 3. Observed versus genome-wide predicted heat susceptibility index (HSI) for each trait using $RAD_{482-GP,0.98}$. The prediction across all populations was based on the additive model M_A applied across populations without (r_a) or with cross validation ($r_{a,cv}$). For the prediction within-populations, the model was built across populations but the prediction was performed within each population without (r_{pop}) and with ($r_{pop,cv}$) cross validation.

were for 80% of the population*trait combinations higher than those found based on a model built across the six populations ($r_{pop,cv}$) when considering the same TS size of 50 (Fig. 6A). However, this was only true for 25% of the combinations when considering the original TS size of 385 genotypes (Fig. 6B). Across all traits and populations, the between-population predictions in the $TS2_c$ scenario (TS built from two combined populations) resulted in higher prediction abilities than those of $TS2_s$ (TS built from one population; Fig. 1, left). Furthermore, we observed that $TS3$ and especially the $TS3_{sr}$ scenario resulted most often in the highest prediction abilities for FxF or DxD VS (Fig. 1, left). For DxV populations as VS, $TS1$ resulted in the highest prediction ability (Fig. 1, right).

Discussion

For most traits, we observed a higher broad sense heritability (h^2_x) at heat compared to standard condition. This observation was due to an increased genotypic variance at heat condition while the error variance was not notably increased (data not shown). Our findings are in contrast with field-based studies in which the heritability is mostly lower at heat compared to standard condition due to an increased error variance (e.g. Cairns *et al.*⁶²). This difference might be explained by field environmental factors which are more important under heat conditions,

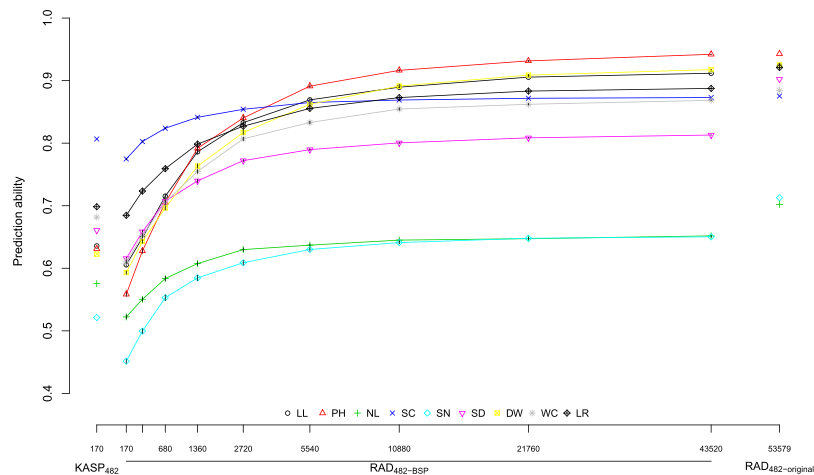


Figure 4. Prediction abilities (r_a) for different numbers of RAD SNPs as well as for the entire KASP₄₈₂ (left) and RAD_{482-GP.0.98} (right) for HSI of nine phenotypic traits. Different numbers of RAD SNPs set sizes were simulated using a resampling procedure. The black vertical bars at each points indicate the standard deviation of the prediction abilities over the 100 replications. The analyses are based on the M_A model.

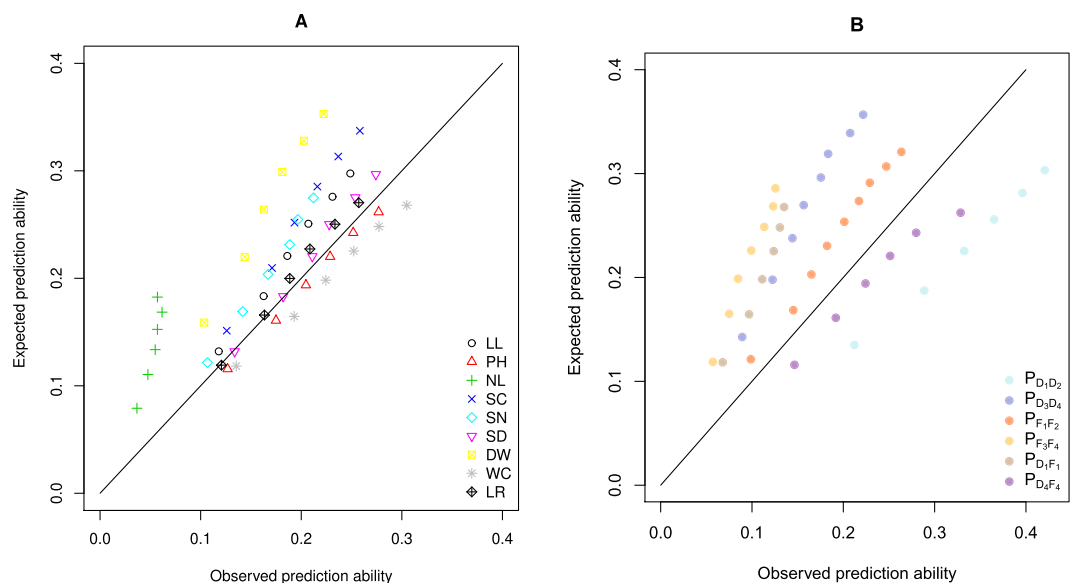


Figure 5. Observed vs. expected within-population prediction abilities averaged across populations (A) and averaged across the HSI of the different traits (B). The analyses are based on RAD_{482-GP.0.98} and the M_A model.

e.g. soil heterogeneity. However, these factors have low relevance under controlled conditions, e.g. in growth chambers used in our study.

The h_a^2 values for the assessed traits were medium to high for both environmental conditions (Table 1). h_a^2 was comparable with heritability estimates observed by Frey *et al.*⁴ and Naveed *et al.*⁶³ under heat stress, and by Cerrudo *et al.*¹⁷ under drought stress. This observation suggests that an adequate estimation of the AEM for each genotype was achieved, which is the prerequisite for a high power to detect QTL, for genome-wide prediction, as well as to interpret differences between heterotic pools.

We observed significant ($P < 0.05$) differences in the HSI between populations derived from DxD crosses compared to populations derived from FxF crosses (Supplementary Fig. 2). Except for SD, the HSI of FxF populations were lower than that of DxD populations. This suggests that FxF populations are more tolerant to heat stress than DxD populations at seedling stage. The findings of Hallauer *et al.*⁶⁴ suggested that the Flint pool contributed an improved chilling tolerance to cultivars bred for temperate Europe. Moreover, the results from Strigens *et al.*⁶⁵ suggested an improved morphological and physiological adaptation of the Flint pool to chilling temperatures compared to the Dent pool. These observations along with ours suggest that genotypes of the Flint pool have a higher tolerance to temperature stress during seedling stage than genotypes of the Dent pool. This

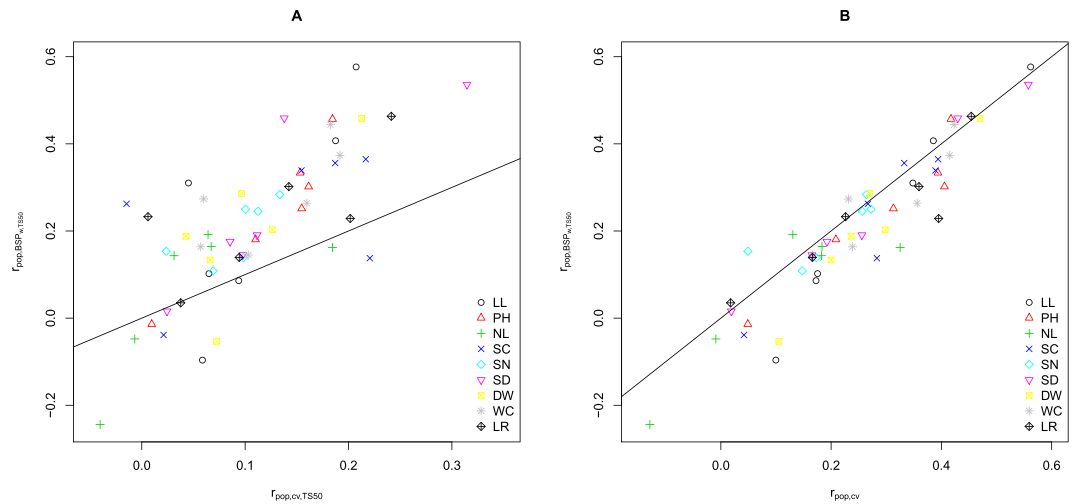


Figure 6. Comparison of the prediction abilities of within-population calibration ($r_{pop,BSP_w,TS50}$) with a training set (TS) of 50 genotypes with the prediction abilities based on an across-population calibration ($r_{pop,cv}$) with a TS size of 50 (A) and 385 (B) genotypes. The analyses are based on $RAD_{482-GP,0.98}$ and the M_A model.

conclusion is in accordance with the results of Reimer *et al.*⁹ who observed a higher tolerance to temperature extremes during seedling stage of genotypes of the Flint pool compared to genotypes of the Dent pool.

The correlation between heat tolerance at seedling and adult stages, which was examined in a companion study⁴, was low (Supplementary Fig. 4). This finding was expected as plant performance in young stages might have limited implications on plant performance after transition from the vegetative to the generative phase. Similar trends were observed for salinity tolerance in wheatgrass⁶⁶ or tolerance to defoliation intensity in maize⁶⁷. Additionally, Gibert *et al.*⁶⁸ observed that for some traits, the correlation between trait and growth changes with plant size and physiological stage. In summary, these results indicate that plant growth strategies should not be considered as constant over the entire life but is stage-dependent.

The negative correlation between the heat tolerance of maize at seedling and adult stages was also manifested on the level of heterotic pools. Genotypes derived from DxD crosses had a lower heat tolerance at seedling stage but a high heat tolerance at adult stage compared to genotypes derived from FxF crosses⁴. Therefore, both heterotic pools should be considered when increasing heat tolerance across developmental stages.

Segregating populations derived from DxF crosses are only rarely created in commercial maize breeding programs. Nevertheless, because of their different behaviour regarding heat tolerance, such populations were included in our study to test their suitability for QTL mapping and GWP. The confidence intervals of the QTL for heat tolerance at seedling stage detected in our study (Table 2) overlapped not only with QTL confidence intervals for the same trait at adult stage⁴ (Fig. 2) but also with QTL regions associated with other abiotic stresses. The confidence intervals of QTL for HSI_{LL} on chromosome 2 overlapped with a QTL associated with cold tolerance, which was identified in a metaQTL-analysis⁶⁹. Furthermore, a QTL identified for the shoot and root dry weight and the leaf area under drought stress conditions⁷⁰ as well as a QTL for the leaf chlorophyll content at drought stress identified by Messmer *et al.*⁷¹ mapped to the same genomic region on top of chromosome 2. These findings might suggest that different abiotic stresses might have similar genetic regulation mechanisms⁷. Furthermore, four out of five heat tolerance genes detected by Frey *et al.*⁷, in an RNAseq experiment, were located within the QTL confidence intervals on chromosome 2 (Supplementary Table 3, Frey *et al.*⁷²). This over-representation of heat tolerance genes in a particular region and their collocation with several QTL for heat tolerance illustrates the importance of genetic mechanisms for heat tolerance available on this chromosome.

Phenotypic evaluation of genetic material for heat tolerance is difficult due to irregular and uncontrolled appearance of heat stress conditions in field experiments and is technically demanding when performed in plant growth chambers. Marker-assisted selection (MAS) using QTL for traits related to heat tolerance could supplement phenotypic selection. However, each QTL detected in our study explained with <10% a relatively small part of the phenotypic variance of the respective trait. It suggests that the heat tolerance at seedling stage is inherited as a true quantitative trait. This in turn makes the use of MAS for heat tolerance inefficient. Therefore, in a subsequent step, we evaluated the utility of GWP to improve heat tolerance of temperate maize.

With the GWP approach, we were able to explain an approximately threefold higher proportion of the genetic variance of the HSI of each trait compared to the detected QTL (Table 3). This is in accordance with earlier reports on other traits in maize^{73–75} and in other crops²⁴, and illustrates the superiority of GWP over MAS for truly quantitative traits.

To quantify the potential advantage of GWP over phenotypic selection, the former can be seen as an indirect selection compared to direct phenotypic selection. The relative merit of indirect vs. direct selection can be calculated as the indirect selection response divided by the direct selection response⁷⁶. This ratio can be rearranged as the ratio $r_{a,cv}/h_{HSI}^2$. GWP is superior to phenotypic selection if this ratio is >1. Assuming the same selection intensity for GWP and phenotypic selection and considering the prediction abilities ($r_{a,cv}$ with $RAD_{482-GP,0.98}$ and M_A)

of our study, the maximum relative cycle can be calculated for which identical selection gains are realised with indirect and direct selection^{76,77}. The maximum relative cycle lengths for GWP were 92% (HSI_{LL}), 88% (HSI_{PH}), 100% (HSI_{NL}), 96% (HSI_{SC}), 68% (HSI_{SN}), 82% (HSI_{SD}), 76% (HSI_{DW}), 100% (HSI_{WC}), and 114% (HSI_{LR}) of those of phenotypic selection. Therefore, GWP for heat stress tolerance is already equivalent or superior to phenotypic selection for three traits even without considering potential advantages due to higher selection intensities, which is very promising in light of earlier studies^{27,78}. In the following, various factors influencing the prediction ability are discussed.

We found a good agreement between observed and expected prediction abilities across all studied traits (Fig. 5). However, some traits were systematically over- (NL, DW) or under-estimated (WC). One of the potential reasons for a difference between observed and expected predicted ability could be the low heritability of the traits. This was especially true for the HSI of NL which had the lowest heritability (Table 1), and which prediction ability deviated substantially from the expected values (Fig. 5). Another reason for the systematic over- and underestimations could be deviations from the highly polygenic genetic architecture of the traits which is assumed in the applied deterministic formula⁵⁸. For such traits, one would expect that the detected QTL explained the highest proportion of the phenotypic variance. However, it is striking that the traits for which the most considerable difference between observed and expected prediction ability was found were the traits for which no QTL were detected in our study. This warrants further research.

The superiority of GWP over phenotypic selection was evaluated as described previously based on an additive genetic model. However, the F_{3,4} populations used in our study had an expected heterozygosity of 25%. In this case, it can be beneficial to investigate genetic models covering not only additive but also dominance effects when performing GWP. No obvious trend was observed regarding the superiority of one of the genetic models for the across-population prediction strategies r_a , r_{pop} , $r_{a,cv}$, and $r_{pop,cv}$ (data not shown). However, the $r_{pop,BSP}$ predictions abilities based on the M_A model were at least equal and, for some population*trait combinations, higher than that observed for the M_{AD} model (Supplementary Figs 7–12). This suggests that dominance effects were either not relevant for the studied traits or were not captured by our GBLUP model. This result was in accordance with observations made in the animal breeding field that generally, the increase of the accuracy of additive breeding values by including dominance effects was scarce⁷⁹. Therefore, we considered only the M_A genetic model for all further analyses.

Other factors that potentially influence the prediction ability are the number and type of molecular markers that are used to characterise the genetic material. From the three GP cutoff values examined for the RAD SNPs, RAD_{482-GP:0.98} resulted in the highest $r_{a,cv}$ values for most of the traits (data not shown). Our finding suggests that the mean imputation of genotypes calls with low GP is less error-prone than the genotype calls obtained by Beagle. Therefore, we chose the GP cutoff value of 0.98 for all further analyses with the RAD SNPs.

The prediction abilities r_a and r_{pop} calculated for RAD₄₈₂ were higher than those obtained for KASP₄₈₂, independent of the considered GP cutoff value (Table 3, Fig. 3, and Supplementary Fig. 6). Our observation is in accordance with the results of Elbasyoni *et al.*²⁵ in wheat and could be due to the considerably higher number of markers that were available in the RAD data set compared to the KASP data set. This was confirmed by the increasing prediction abilities observed for increasing numbers of RAD SNPs (Fig. 4). A high number of markers increases the precision of the K estimates, increases LD between markers and QTL, as well as ensures a better genome representation.

A second observation leading to the same conclusion was that Q_{ST} was significantly correlated with r_a calculated for KASP₄₈₂ but not with r_a calculated for RAD_{482-GP:0.98}. This finding suggests that the prediction abilities obtained with the KASP SNPs are predominantly due to the modeling of genetic relationships and are therefore higher when the ratio of genetic variance among- vs. within-populations increases. In contrast, the prediction abilities obtained with RAD SNPs are to a higher extent due to LD between marker and QTL alleles and to a lower extent due to the modeling of genetic relationship and thus do not correlate with Q_{ST}.

Finally, the prediction abilities r_a calculated for KASP₄₈₂ were with 0.72 also significantly ($P = 0.028$) correlated with h^2_{HSI} , which is in accordance with the results of Poland *et al.*⁸⁰, Hayes *et al.*⁸¹, and De Moraes *et al.*²⁶. However, the prediction ability r_a calculated for RAD_{482-GP:0.98} were not significantly ($P = 0.14$) correlated with h^2_{HSI} . This agrees theoretical considerations⁸² and empirical studies⁸³ and can be explained by an increased proportion of the prediction abilities caused by LD between markers and QTL when a high number of molecular markers was used. Therefore, all discussions hereafter are based on RAD_{482-GP:0.98}.

Further factors that have a high impact on the prediction abilities are the size and composition of the training set and, thus, were examined in our study. For most trait*population combinations, we observed an increase of the prediction ability $r_{pop,BSP}$ with an increasing size of the training set (Supplementary Figs 7–13). This is in accordance with the results of Van Raden *et al.*⁸⁴ and Technow *et al.*²⁷. However, in contrast to these authors, we did not observe a plateau with an increasing size of the TS. One reason for this could be that our training set size is small and the plateau was not yet reached. Nevertheless, the TS size of this study corresponds to that typically used in commercial maize breeding programs.

We observed significant ($P < 0.05$) differences in the heat tolerance between the two heterotic pools. Therefore, we were interested in examining the potential of prediction between populations (inter-population calibration) and heterotic pools (inter-pool calibration) in comparison with the previously discussed within-population calibration. The latter resulted across all population*trait combinations in higher prediction abilities $r_{pop,BSP}$ than the inter-population calibration, independent of the TS used for the $r_{pop,BSP}$. This is in accordance with the results of Technow *et al.*²⁷. The most likely reason for this is that LD patterns are not consistent between heterotic pools^{85,86}. To identify markers that are in LD with QTL across different heterotic pools, an even higher number of markers might be required^{86,87}. Nevertheless, as situations in which within-population prediction rarely appear in commercial breeding programs, we evaluated various scenarios for $r_{pop,BSP}$.

We observed a considerably lower prediction ability for intra-pool calibration (r_{pop,BSP_b} based on TS1; Fig. 1, left) than for the intra-population calibration (r_{pop,BSP_w}). This was more pronounced than expected according to Habier *et al.*⁵⁴. Our observation can be due to the fact that the parental inbreds used to develop segregating populations in our study were selected such that they show a high variation for heat tolerance related traits not only across all eight inbreds but also across the inbreds of one heterotic pool. Therefore, a high variation between the populations was observed even when the parental inbreds were from the same heterotic pool.

Compared to TS1, we observed a minor increase of the prediction ability when using an inter-pool calibration (TS2; Fig. 1, left). This was especially true when the TS was composed of two populations (TS2_c). Moreover, this finding might be explained by the fact that the four Flint and Dent parental inbred lines used to develop the segregating populations were chosen to be as diverse as possible.

Compared to the results of TS2 (s and c), the prediction abilities for the mixed-pool calibration (TS3) were even higher. A part of this increase was due to the fact that the DxF populations shared parental inbreds with the DxD and FxF populations. Our observation that the prediction ability for a scenario with shared parental inbred (TS3_{sr}) were higher than for a scenario with no shared parental inbreds between TS and VS (TS3_{su}; Fig. 1, left) supported this explanation. In addition to the fact that TS and VS have one common parental inbred, the higher prediction ability for TS3 compared to TS2 (s and c) might be due to the higher segregation variance of the mixed populations (Supplementary Table 2 and Supplementary Fig. 5).

This aspect was studied more in detail by examining the prediction ability for mixed populations (DxF) using different types of TS (Fig. 1, right). The TS with combined FxF and DxD populations which were partially derived from common parents with the VS (TS3) performed better than combined FxF and DxD populations that had no parental inbred in common with the VS (TS2). This further emphasises parental inbreds shared between the populations of the TS and VS is an important component for the success of genomic prediction. However, we observed the highest prediction ability for mixed populations when using other mixed populations as TS even if when had no common parents (TS1). One hypothesis to explain this result is the high diversity of such mixed populations (Supplementary Table 2). Our findings suggest that commercial maize breeding companies who create DxF populations to e.g. increase the per-se performance of the Flint pool with Dent material or to introgress earliness in Dent inbreds with Flint material can use such populations as TS to predict the original heterotic pools.

Generally, the critical aspect when selecting the TS in routine plant breeding programs is the limited size of the segregating populations which hampers either the building of a TS of suitable size or reduces relatedness between TS and VS if several segregating populations are used to build the TS. Therefore, we have evaluated the potential of across-pools and -populations calibrations. r_{pop,BSP_w} was for 80% of the population*trait combinations better than a model built on the six populations ($r_{pop,cv}$) if considering the same TS size of 50. However, this was only the case for 25% of the combinations if considering the original TS size of 385 genotypes (Fig. 6). This implies that if genomic prediction is used to select among the genotypes of one single population and if the data for the TS must be generated de novo, the use of within-population prediction is the most promising approach. Nevertheless, in many cases, a commercial maize breeding company has multiple connected segregating populations for which predictions should be obtained. For such a scenario, our results clearly indicate that a combination of populations in one TS, also across different heterotic pools, increases the prediction ability compared to the use of within-population calibration (Fig. 6).

Conclusion

In this study, we provided a basis for future research on genome regions and candidate genes involved in heat stress tolerance of maize seedlings. Antagonistic pleiotropy between heat tolerance at seedling and adult stages was observed in genomic hot-spot regions, especially on chromosome 2. Although six QTL were detected in our study, these explained a very small proportion of the phenotypic variance. Therefore, marker-assisted selection is not promising for the traits evaluated in our study. On the contrary, the prediction abilities observed for GWP encourages the use of such approaches for heat tolerance at seedling stage, especially if genotypic characterisation was carried out with a high number of markers, e.g. from RAD sequencing. Furthermore, our results suggest that the combination of populations from different heterotic pools in one training set is a promising approach to increase the prediction ability for heat tolerance traits, especially if the population in the TS share parental genotypes with those of the VS. Finally, we demonstrated that for the examined traits, segregation populations derived from inter-pool crosses are very suitable to predict populations derived from intra-pool crosses.

Data Availability

The sequencing datasets generated and analysed in the current study are available in the NCBI Sequence Read Archive (SRA) repository, <https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA564701>. All phenotyping and genotyping data generated or analysed during this study are included as supplementary information files.

References

1. Sage, R. F., Kocacinar, F. & Kubien, D. S. C4 photosynthesis and temperature. In Raghavendra, A. S. & Sage, R. F. (eds) *C4 Photosynthesis and Related CO2 Concentrating Mechanisms*, chap. 10, 161–195 (Springer, Dordrecht, The Netherlands, 2011).
2. Giaveno, C. & Ferrero, J. Introduction of tropical maize genotypes to increase silage production in the central area of Santa Fe, Argentina. *Crop Breeding and Applied Biotechnology* **3**, 89–94 (2003).
3. Barnabás, B., Jäger, K. & Fehér, A. The effect of drought and heat stress on reproductive processes in cereals. *Plant, Cell and Environment* **31**, 11–38, <https://doi.org/10.1111/j.1365-3040.2007.01727.x> (2008).
4. Frey, F. P., Presterl, T., Lecoq, P., Orlik, A. & Stich, B. First steps to understand heat tolerance of temperate maize at adult stage: identification of QTL across multiple environments with connected segregating populations. *Theoretical and Applied Genetics* **129**, 945–961, <https://doi.org/10.1007/s00122-016-2674-6> (2016).

5. Alam, M. A. *et al.* Dissecting heat stress tolerance in tropical maize (*Zea mays* L.). *Field Crops Research* 110–119, <https://doi.org/10.1016/j.fcr.2017.01.006> (2017).
6. Wahid, A., Gelani, S., Ashraf, M. & Foolad, M. Heat tolerance in plants: An overview. *Environmental and Experimental Botany* 61, 199–223 (2007).
7. Frey, F. P., Urbany, C., Hüttel, B., Reinhardt, R. & Stich, B. Genome-wide expression profiling and phenotypic evaluation of European maize inbreds at seedling stage in response to heat stress. *BMC Genomics* 16, 123, <https://doi.org/10.1186/s12864-015-1282-1> (2015).
8. Deutsches Maiskomitee, http://www.maiskomitee.de/web/public/Fakten.aspx/Statistik/break_Deutschland/Statistik/Biogas (2013).
9. Reimer, R. *et al.* Root response to temperature extremes: association mapping of temperate maize (*Triticum aestivum* L.). *Maydica* 58, 156–168 (2013).
10. Lobell, D. B. & Field, C. B. Global scale climate-crop yield relationships and the impacts of recent warming. *Environmental Research Letters* 2, 14002 (2007).
11. IPCC. Summary for policymakers. In Stocker, T. F. *et al.* (eds) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013).
12. Chen, J., Xu, W., Velten, J., Xin, Z. & Stout, J. Characterization of maize inbred lines for drought and heat tolerance. *Journal of Soil and Water Conservation* 67, 354–364 (2012).
13. Ottaviano, E., Sari-Gorla, M., Pè, E. & Frova, C. Molecular markers (RFLPs and HSPs) for the genetic dissection of thermotolerance in maize. *Theoretical and Applied Genetics* 81, 713–719 (1991).
14. Frova, C. & Sari-Gorla, M. Quantitative trait loci (QTLs) for pollen thermotolerance detected in maize. *Molecular and General Genetics* 245, 424–430 (1994).
15. Buckley, L. B. & Huey, R. B. How extreme temperatures impact organisms and the evolution of their thermal tolerance. *Integrative and Comparative Biology* 56, 98–109, <https://doi.org/10.1093/icb/icw004> (2016).
16. Sehgal, D., Singh, R. & Rajpal, V. R. Quantitative Trait Loci Mapping in Plants: Concepts and Approaches. In Rajpal, V., Rao, S. & Raina, S. (eds) *Molecular Breeding for Sustainable Crop Improvement*, vol. 2, chap. 2, 31–60, <https://doi.org/10.1007/978-3-319-27090-6> (Springer, 2016).
17. Cerrudo, D. *et al.* Genomic Selection Outperforms Marker Assisted Selection for Grain Yield and Physiological Traits in a Maize Doubled Haploid Population Across Water Treatments. *Frontiers in Plant Science* 9, 1–12, <https://doi.org/10.3389/fpls.2018.00366> (2018).
18. Lorenzana, R. E. & Bernardo, R. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theoretical and Applied Genetics* 120, 151–161, <https://doi.org/10.1007/s00122-009-1166-3> (2009).
19. Albrecht, T. *et al.* Genome-based prediction of testcross values in maize. *Theoretical and Applied Genetics* 123, 339–350, <https://doi.org/10.1007/s00122-011-1587-7> (2011).
20. Lian, L., Jacobson, A., Zhong, S. & Bernardo, R. Genomewide prediction accuracy within 969 maize biparental populations. *Crop Science* 54, 1514–1522, <https://doi.org/10.2135/cropsci2013.12.0856> (2014).
21. Poland, J. A., Brown, P. J., Sorrells, M. E. & Jannink, J. L. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One* 7, <https://doi.org/10.1371/journal.pone.0032253> (2012).
22. Crossa, J. *et al.* Genomic Prediction in Maize Breeding Populations with Genotyping-by-Sequencing. *G3 Genes Genomes Genetics* 3, 1903–1926, <https://doi.org/10.1534/g3.113.008227> (2013).
23. Massman, J. M., Gordillo, A., Lorenzana, R. E. & Bernardo, R. Genomewide predictions from maize single-cross data. *Theoretical and Applied Genetics* 126, 13–22, <https://doi.org/10.1007/s00122-012-1955-y> (2013).
24. Zhang, X. *et al.* Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. *Heredity* 114, 291–299, <https://doi.org/10.1038/hdy.2014.99> (2015).
25. Elbasyoni, I. S. *et al.* A comparison between genotyping-by-sequencing and array-based scoring of SNPs for genomic prediction accuracy in winter wheat. *Plant Science* 270, 123–130, <https://doi.org/10.1016/j.plantsci.2018.02.019> (2018).
26. de Moraes, B. F. X. *et al.* Genomic selection prediction models comparing sequence capture and SNP array genotyping methods. *Molecular Breeding* 38, <https://doi.org/10.1007/s11032-018-0865-3> (2018).
27. Technow, F., Bürger, A. & Melchinger, A. E. Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. *G3 Genes Genomes Genetics* 3, 197–203, <https://doi.org/10.1534/g3.112.004630> (2013).
28. Andersen, J. R., Schrag, T., Melchinger, A. E., Zein, I. & Lübberstedt, T. Validation of Dwarf8 polymorphisms associated with flowering time in elite European inbred lines of maize (*Zea mays* L.). *Theoretical and Applied Genetics* 111, 206–217 (2005).
29. Ganai, M. W. *et al.* A large maize (*Triticum aestivum* L.) SNP genotyping array: Development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS One* 6, <https://doi.org/10.1371/journal.pone.0028334> (2011).
30. Baird, N. A. *et al.* Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3, 1–7, <https://doi.org/10.1371/journal.pone.0003376> (2008).
31. Paris, J. R., Stevens, J. R. & Catchen, J. M. Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution* 8, 1360–1373, <https://doi.org/10.1111/2041-210X.12775> (2017).
32. Navarro, J. A. R. *et al.* A study of allelic diversity underlying flowering-time adaptation in maize landraces. *Nature Genetics* 49, 476–480, <https://doi.org/10.1038/ng.3784> (2017).
33. Auwera, G. A. V. D. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 11, <https://doi.org/10.1002/0471250953.bi1110s43> (2014).
34. Browning, B. L. & Browning, S. R. Genotype Imputation with Millions of Reference Samples. *American Journal of Human Genetics* 98, 116–126, <https://doi.org/10.1016/j.ajhg.2015.11.020> (2016).
35. Kuo, F. *Tools for genetic data management and strategies for optimized imputation of missing genotypes*. Dissertation (2014).
36. Nei, M. *Molecular Evolutionary Genetics* (Columbia University Press, 1987).
37. Gerlach, G., Jueterbock, A., Kraemer, P., Deppermann, J. & Harmand, P. Calculations of population differentiation based on GST and D: Forget GST but not all of statistics. *Molecular Ecology* 19, 3845–3852, <https://doi.org/10.1111/j.1365-294X.2010.04784.x> (2010).
38. Wright, S. *Evolution and the Genetics of Populations. Volume 4: Variability Within and Among Natural Populations* (Chicago, 1978).
39. Mason, R. E. *et al.* QTL associated with heat susceptibility index in wheat (*Triticum aestivum* L.) under short-term reproductive stage heat stress. *Euphytica* 174, 423–436 (2010).
40. Hallauer, A. R., Carena, M. J. & Miranda Filho, J. B. *Quantitative Genetics in Maize Breeding*. (Springer New York, New York, NY, 2010).
41. Ouk, M. *et al.* Use of drought response index for identification of drought tolerant genotypes in rainfed lowland rice. *Field Crops Research* 99, 48–58, <https://doi.org/10.1016/j.fcr.2006.03.003> (2006).
42. Lande, R. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution* 46, 381–389 (1992).
43. Prout, T. & Barker, J. S. F. F statistics in *Drosophila buzzatii*: Selection, population size and inbreeding. *Genetics* 134, 369–375, <https://doi.org/10.1007/978-94-009-0619-8> (1993).
44. Spitz, K. Population structure in *Daphnia obtusa*: Quantitative genetic and allozymic variation. *Genetics* 135, 367–374, <https://doi.org/10.1007/s10709-011-9564-2> (1993).

45. Benke, A. *et al.* The genetic basis of natural variation for iron homeostasis in the maize IBM population. *BMC Plant Biology* **14**, 12 (2014).
46. Horn, F., Habekuß, A. & Stich, B. Linkage mapping of barley yellow dwarf virus resistance in connected populations of maize. *BMC Plant Biology* **15**, 29 (2015).
47. de Givry, S., Bouchez, M., Chabrier, P., Milan, D. & Schiex, T. CARTHAGENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics* **21**, 1703–1704 (2005).
48. Charcosset, A. *et al.* Heterosis in maize investigated using connected RIL populations. In *Quantitative genetics and breeding methods: the way ahead. Les colloques no 96*, 89–98 (INRA, Paris, France, 2001).
49. Mangin, B. *et al.* MCQTL: A reference manual (2010).
50. Bardol, N. *et al.* Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. *Theoretical and Applied Genetics* **126**, 2717–2736 (2013).
51. Churchill, G. A. & Doerge, R. W. Empirical Threshold Values for Quantitative Trait Mapping. *Genetics* **138**, 963–971, <https://doi.org/10.1534/genetics.107.080101> (1994).
52. Lander, E. S. & Botstein, S. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185, <https://doi.org/10.1038/hdy.2014.4> (1989).
53. VanRaden, P. M. Efficient methods to compute genomic predictions. *Journal of Dairy Science* **91**, 4414–4423, <https://doi.org/10.3168/jds.2007-0980> (2008).
54. Habier, D., Fernando, R. L. & Dekkers, J. C. M. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* **177**, 2389–2397, <https://doi.org/10.1534/genetics.107.081190> (2007).
55. Muñoz, P. R. *et al.* Unraveling Additive from Nonadditive Effects Using. *Genetics* **198**, 1759–1768, <https://doi.org/10.1534/genetics.114.171322> (2014).
56. Covarrubias-Pazarán, G. Genome-Assisted prediction of quantitative traits using the r package sommer. *PLoS One* **11**, 1–15, <https://doi.org/10.1371/journal.pone.0156744> (2016).
57. Legarra, A., Robert-Granié, C., Manfredi, E. & Elsen, J. M. Performance of genomic selection in mice. *Genetics* **180**, 611–618, <https://doi.org/10.1534/genetics.108.088575>, 1303.3997 (2008).
58. Daetwyler, H. D., Pong-Wong, R., Villanueva, B. & Woolliams, J. A. The impact of genetic architecture on genome-wide evaluation methods. *Genetics* **185**, 1021–1031, <https://doi.org/10.1534/genetics.110.116855> (2010).
59. Meuwissen, T. & Goddard, M. Accurate prediction of genetic values for complex traits by whole-genome resequencing. *Genetics* **185**, 623–631, <https://doi.org/10.1534/genetics.110.116590> (2010).
60. Su, C. *et al.* High Density Linkage Map Construction and Mapping of Yield Trait QTLs in Maize (*Zea mays*) Using the Genotyping-by-Sequencing (GBS) Technology. *Frontiers in Plant Science* **8**, 1–14, <https://doi.org/10.3389/fpls.2017.00706> (2017).
61. Hartl, D. & Clark, A. *Principles of population genetics* (Sinauer associates Sunderland, 1997).
62. Cairns, J. E. *et al.* Identification of Drought, Heat, and Combined Drought and Heat Tolerant Donors in Maize. *Crop Science* **53**, 1335 (2013).
63. Naveed, M., Ahsan, M., Akram, H. M., Aslam, M. & Ahmed, N. Genetic Effects Conferring Heat Tolerance in a Cross of Tolerant x Susceptible Maize (*Zea mays* L.) Genotypes. *Frontiers in Plant Science* **7**, 1–12, <https://doi.org/10.3389/fpls.2016.00729> (2016).
64. Hallauer, A. R. Methods used in developing maize inbreds. *Maydica* **35**, 1–16 (1990).
65. Strigens, A. *et al.* Association mapping for chilling tolerance in elite flint and dent maize inbred lines evaluated in growth chamber and field experiments. *Plant, Cell and Environment* **36**, 1871–1887, <https://doi.org/10.1111/pce.12096> (2013).
66. Pearen, J. R., Pahl, M. D., Wolynetz, M. S. & Hermesh, R. Association of salt tolerance at seedling emergence with adult plant performance in slender wheatgrass. *Canadian Journal of Plant Science* **77**, 81–89, <https://doi.org/10.4141/P95-159> (1997).
67. Denton, E. M., Smith, B. S., Hamerlynck, E. P. & Sheley, R. L. Seedling Defoliation and Drought Stress: Variation in Intensity and Frequency Affect Performance and Survival. *Rangeland Ecology and Management* **71**, 25–34, <https://doi.org/10.1016/j.rama.2017.06.014> (2018).
68. Gibert, A., Gray, E. F., Westoby, M., Wright, I. J. & Falster, D. S. On the link between functional traits and growth rate: meta-analysis shows effects change with plant size, as predicted. *Journal of Ecology* **104**, 1488–1503, <https://doi.org/10.1111/1365-2745.12594> (2016).
69. Rodríguez, V. M., Butrón, A., Rady, M. O. A., Soengas, P. & Revilla, P. Identification of quantitative trait loci involved in the response to cold stress in maize (*Zea mays* L.). *Molecular Breeding* **33**, 363–371 (2013).
70. Ruta, N., Stamp, P., Liedgens, M., Fracheboud, Y. & Hund, A. Collocations of QTLs for Seedling Traits and Yield Components of Tropical Maize under Water Stress Conditions. *Crop Science* **50**, 1385 (2010).
71. Messmer, R., Fracheboud, Y., Bänziger, M., Stamp, P. & Ribaut, J.-M. Drought stress and tropical maize: QTLs for leaf greenness, plant senescence, and root capacitance. *Field Crops Research* **124**, 93–103 (2011).
72. Frey, F. P. *The genetic basis of heat tolerance in temperate maize (Zea mays L.)*. Ph.D. thesis (2015).
73. Peiffer, J. A. *et al.* The Genetic Architecture of Maize Stalk Strength. *PLoS One* **8**, <https://doi.org/10.1371/journal.pone.0067066> (2013).
74. Peiffer, J. A. *et al.* The genetic architecture of maize height. *Genetics* **196**, 1337–1356, <https://doi.org/10.1534/genetics.113.159152> (2014).
75. Foiada, F. *et al.* Improving resistance to the European corn borer: a comprehensive study in elite maize using QTL mapping and genome-wide prediction. *Theoretical and Applied Genetics* **128**, 875–891, <https://doi.org/10.1007/s00122-015-2477-1> (2015).
76. Falconer, D. S. & Mackay, T. F. C. *Introduction to quantitative genetics*. fourth edn. (Longman Group, Harlow, UK, 1996).
77. Stich, B. & Van Inghelandt, D. Prospects and potential uses of genomic prediction of key performance traits in tetraploid potato. *Frontiers in Plant Science* **9**, 1–12, <https://doi.org/10.3389/fpls.2018.00159> (2018).
78. Heffner, E. L., Lorenz, A. J., Jannink, J. L. & Sorrells, M. E. Plant breeding with Genomic selection: Gain per unit time and cost. *Crop Science* **50**, 1681–1690, <https://doi.org/10.2135/cropsci2009.11.0662> (2010).
79. Varona, L., Legarra, A., Toro, M. A. & Vitezica, Z. G. Non-additive effects in genomic selection. *Frontiers in Genetics* **9**, 1–12, <https://doi.org/10.3389/fgene.2018.00078> (2018).
80. Poland, J. *et al.* Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome* **5**, 103–113, <https://doi.org/10.3835/Plantgenome2012.06.0006> (2012).
81. Hayes, B. J. *et al.* Genomic Prediction from Whole Genome Sequence in Livestock: the 1000 Bull Genomes Project. *Proceedings of the 10th World Congress in Genetics Applied to Livestock Production*, <https://doi.org/10.7910/DVN/KHBDWU.0505333v2> (2014).
82. Meuwissen, T. H. Accuracy of breeding values of ‘unrelated’ individuals predicted by dense SNP genotyping. *Genetics Selection Evolution* **41**, 1–9, <https://doi.org/10.1186/1297-9686-41-35> (2009).
83. Heffner, E. L., Jannink, J.-L. & Sorrells, M. E. Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. *The Plant Genome* **4**, 65, <https://doi.org/10.3835/plantgenome2010.12.0029> (2011).
84. VanRaden, P. *et al.* Invited Review: Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science* **92**, 16–24, <https://doi.org/10.3168/jds.2008-1514> (2009).
85. Goddard, M., Hayes, B., McPartlan, H. & Chamberlain, A. Can the same markers be used in multiple breeds? In *8th World Congress on Genetics Applied to Livestock Production* (2006).

86. Van Inghelandt, D., Reif, J. C., Dhillon, B. S., Flament, P. & Melchinger, A. E. Extent and genome-wide distribution of linkage disequilibrium in commercial maize germplasm. *Theoretical and Applied Genetics* **123**, 11–20, <https://doi.org/10.1007/s00122-011-1562-3> (2011).
87. De Roos, A. P., Hayes, B. J., Spelman, R. J. & Goddard, M. E. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. *Genetics* **179**, 1503–1512, <https://doi.org/10.1534/genetics.107.084301> (2008).

Acknowledgements

We would like to thank the Gemeinschaft zur Förderung der privaten deutschen Pflanzenzüchtung eV (GFP) for supporting the grant application. We thank the Max Planck-Genome-centre Cologne <http://mpgc.mpipz.mpg.de/home/> for sequencing the RAD libraries of this study. For technical assistance with the lab work and the growth chamber experiments we thank Andrea Lossow, Isabell Scheibert, Nele Kaul, Nicole Kamphaus and Frederike Horn. Further, we thank the team of gardeners at the MPIPZ, especially Frank Eikermann, Manfred Pohe, Sybille Richter, and Caren Dawidson for their help with the realisation of the growth experiments and with setting up the irrigation system. This research was funded by the Federal Ministry of Food, Agriculture and Consumer Protection - Project ID: 28-1-45.068-10 in the frame of the support of the innovation program as well as the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2048/1 - Project ID: 390686111. A subset of the results of this manuscript are part of the cumulative dissertation "*The genetic basis of heat tolerance in temperate maize (Zea mays L.)*" by Felix Frey.

Author Contributions

B.S. conceived the study, F.F. collected the phenotypic data and performed the QTL analysis, D.R. analyzed the RAD sequencing data, D.V.I. performed the RAD imputation and the genome-wide prediction analyses. D.V.I., B.S. and F.F. wrote the manuscript. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-50853-2>.

Competing Interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019