Footprints of Selection Derived From Temporal Heterozygosity Patterns in a Barley Nested Association Mapping Population

Andreas Maurer* and Klaus Pillen

Institute of Agricultural and Nutritional Sciences, Chair of Plant Breeding, Martin Luther University Halle-Wittenberg, Halle, Germany

Nowadays, genetic diversity more than ever represents a key driver of adaptation to climate challenges like drought, heat, and salinity. Therefore, there is a need to replenish the limited elite gene pools with favorable exotic alleles from the wild progenitors of our crops. Nested association mapping (NAM) populations represent one step toward exotic allele evaluation and enrichment of the elite gene pool. We investigated an adaptive selection strategy in the wild barley NAM population HEB-25 based on temporal genomic data by studying the fate of 214,979 SNP loci initially heterozygous in individual BC_1S_3 lines after five cycles of selfing and field propagation. We identified several loci exposed to adaptive selection in HEB-25. In total, 48.7% (104,725 SNPs) of initially heterozygous SNP calls in HEB-25 were fixed in $BC_1S_{3:8}$ generation, either toward the wild allele (19.9%) or the cultivated allele (28.8%). Most fixed SNP loci turned out to represent gene loci involved in domestication and flowering time as well as plant height, for example, *btr1/btr2*, *thresh-1*, *Ppd-H1*, and *sdw1*. Interestingly, also unknown loci were found where the exotic allele was fixed, hinting at potentially useful exotic alleles for plant breeding.

Keywords: barley, heterozygosity, temporal genomic data, adaptive evolution, selection, natural selection, artificial selection, unconscious selection

INTRODUCTION

Around 10,000 years ago, domestication of crops enabled mankind to settle and to commence agriculture. Year after year, early farmers and breeders selected the best performing plants for the next season. Domestication and selection were accompanied by a progressive depletion of genetic diversity, known as bottleneck effect (Tanksley and McCouch, 1997). To cope with future agricultural challenges, there is a need to replenish the limited elite gene pools with favorable exotic alleles from the wild progenitors of our crops (Zamir, 2008; McCouch et al., 2013; Zhang et al., 2017). However, the identification of beneficial exotic material can be laborious and challenging (Dempewolf et al., 2017).

NAM populations can be developed to investigate a multitude of exotic allele effects in an adapted background. They are created by crossing a diverse set of wild progenitors with one recurrent elite cultivar. This way high allele richness and statistical power are combined to evaluate complex traits through genome-wide association studies (GWAS).

Subsequent backcross steps with the elite cultivar may serve as a first step to integrate exotic alleles in adapted breeding material and increase the reliability of estimated wild allele effects. The wild barley NAM population HEB-25 is backcross-based and comprises 1,420 BC_1S_3 lines resulting from crosses of 25 highly divergent wild barley accessions (*Hordeum vulgare* ssp. *spontaneum* and ssp. *agriocrithon*) with the elite barley cultivar Barke (Maurer et al., 2015). In this population recombination rate (Dreissig et al., 2020) and several important agronomic traits like plant development (Maurer et al., 2015, 2016; Herzig et al., 2018; Wiegmann et al., 2019a; Pham et al., 2020), yield formation (Maurer et al., 2016; Sharma et al., 2018; Wiegmann et al., 2019a), grain nutrient concentration (Herzig et al., 2019; Wiegmann et al., 2019b), as well as tolerance to abiotic (Saade et al., 2016; Merchuk-Ovnat et al., 2018; Pham et al., 2019) and biotic (Vatter et al., 2017, 2018; Büttner et al., 2020; Mehnaz et al., 2021) stresses were investigated. Although all of these studies proved useful to find genomic regions controlling the investigated traits, in many cases, it remains unclear which effects can truly be designated as beneficial across certain environments. For instance, a clear statement about the usefulness of flowering time affecting wild alleles is often not possible due to their environment-dependent phenotypic plasticity (Guo et al., 2020; Li et al., 2021), complicating the transfer of beneficial alleles from one environment to another.

In this context, outsourcing the job of selection to mother nature may help to define what is beneficial in a certain environment. Over the long term, this enables the prevalence of certain ideotypes with a complex of optimally coordinated properties rather than focusing on single traits in classical selection. This principle of natural selection is key of the evolutionary plant breeding concept (Suneson, 1956; Phillips and Wolfe, 2005; Döring et al., 2011).

In the present study, we investigated an adaptive selection strategy in HEB-25 by temporal screening of initially heterozygous loci (where 6.25% of the genome is expected to be heterozygous in each HEB line in generation BC_1S_3) after 5 years of selfing and field propagation without conscious selection as a by-product of population conservation. A clear fixation of exotic alleles could hint at potentially useful exotic alleles that were more successful in contributing to the next field generation.

MATERIALS AND METHODS

Plant Material

The wild barley nested association mapping (NAM) population HEB-25 resulted from parallel crosses of 25 highly divergent wild barley accessions (*Hordeum vulgare* ssp. *spontaneum* and ssp. agriocrithon, hereafter named *Hsp*) with the German elite barley cultivar Barke (*Hordeum vulgare* ssp. *vulgare*, hereafter named *Hv*). F1 plants of the initial crosses were backcrossed with Barke and after three subsequent rounds of selfing the population comprised 1,420 BC_1S_3 plants (Maurer et al., 2015). Due to the mating design, each BC_1S_3 plant is expected to

harbor 71.875% homozygous Barke loci, 6.25% heterozygous loci, and 21.875% homozygous wild loci under the assumption of no selection.

Population Conservation

In 2011, BC_1S_3-derived lines were created by growing the complete progeny of each single BC_1S_3 plant (i.e., $BC_1S_{3:4}$) in small plots (double rows of 1.50 m length) in the field (Halle, Germany, 51°29′46.47″N; 11°59′41.81″E) and 20 randomly chosen ears of each line, displaying a representative subset of the whole plot, were harvested at maturity. After threshing and manual seed processing, 60 seeds thereof were randomly selected for sowing in 2012 (i.e., $BC_1S_{3:5}$). The same process of harvesting, processing, and sowing was repeated with 60 $BC_1S_{3:6}$ seeds in 2013 and 100 $BC_1S_{3:7}$ seeds in 2014. Then, 20 $BC_1S_{3:8}$ seeds (harvested from the $BC_1S_{3:7}$ plants in 2014) were grown and leaf material from 12 randomly chosen seedlings per line (50–100 mg) was harvested to form pooled samples for DNA extraction. This way, initial heterozygosity could be reconstructed through heterogeneity within the 12 plants (**Figure 1**).

DNA Extraction and SNP Genotyping

DNA was extracted according to the manufacturer's protocol, using the BioSprint 96 DNA Plant Kit and a BioSprint96 work station (Qiagen, Hilden, Germany), and finally dissolved in distilled water at approximately 50 ng/μl. The original 1,420 BC_1S_3 plants were genotyped with the barley 9 k Infinium iSelect SNP array (Maurer et al., 2015), consisting of 7,864 SNP markers as reported in Comadran et al. (2012). Pooled DNA samples of 12 $BC_1S_{3:8}$ seedlings per HEB-25 line were genotyped with the barley 50 k Infinium iSelect SNP array (Bayer et al., 2017; Maurer and Pillen, 2019) at TraitGenetics GmbH, Gatersleben, Germany to reconstruct original heterozygosity. SNP calling in Illumina genotyping assays is based on the concept of hybridization technology with specifically designed oligonucleotide probes, where the intensity of two distinct fluorescently labeled target sequences represents the signal strength for each of the two alleles. Then, a cluster algorithm is applied to distinguish the two contrasting homozygous classes and heterozygous calls (Zhao et al., 2018). At TraitGenetics GmbH, the cluster files determining the thresholds for allelic discrimination have been manually revised to improve the call quality, both for the 9 k and the 50 k array. After SNP calling, SNP markers that did not meet the quality criteria (polymorphic in at least one HEB family, < 10% failure (i.e., no call) rate, and < 12.5% heterozygous calls, which is twice the expectancy in BC_1S_3) were removed from the dataset. Furthermore, 256 SNPs were removed as they revealed exact segregation among all HEB lines, indicating that they were in complete linkage disequilibrium (LD). Only one of these duplicates was kept. Altogether, 4,717 SNPs, genotyped in both SNP arrays, met the quality criteria. In the present study, 57 of the initial 1,420 lines were eliminated due to clearly inconsistent genotypes between BC_1S_3 and $BC_1S_{3:8}$.
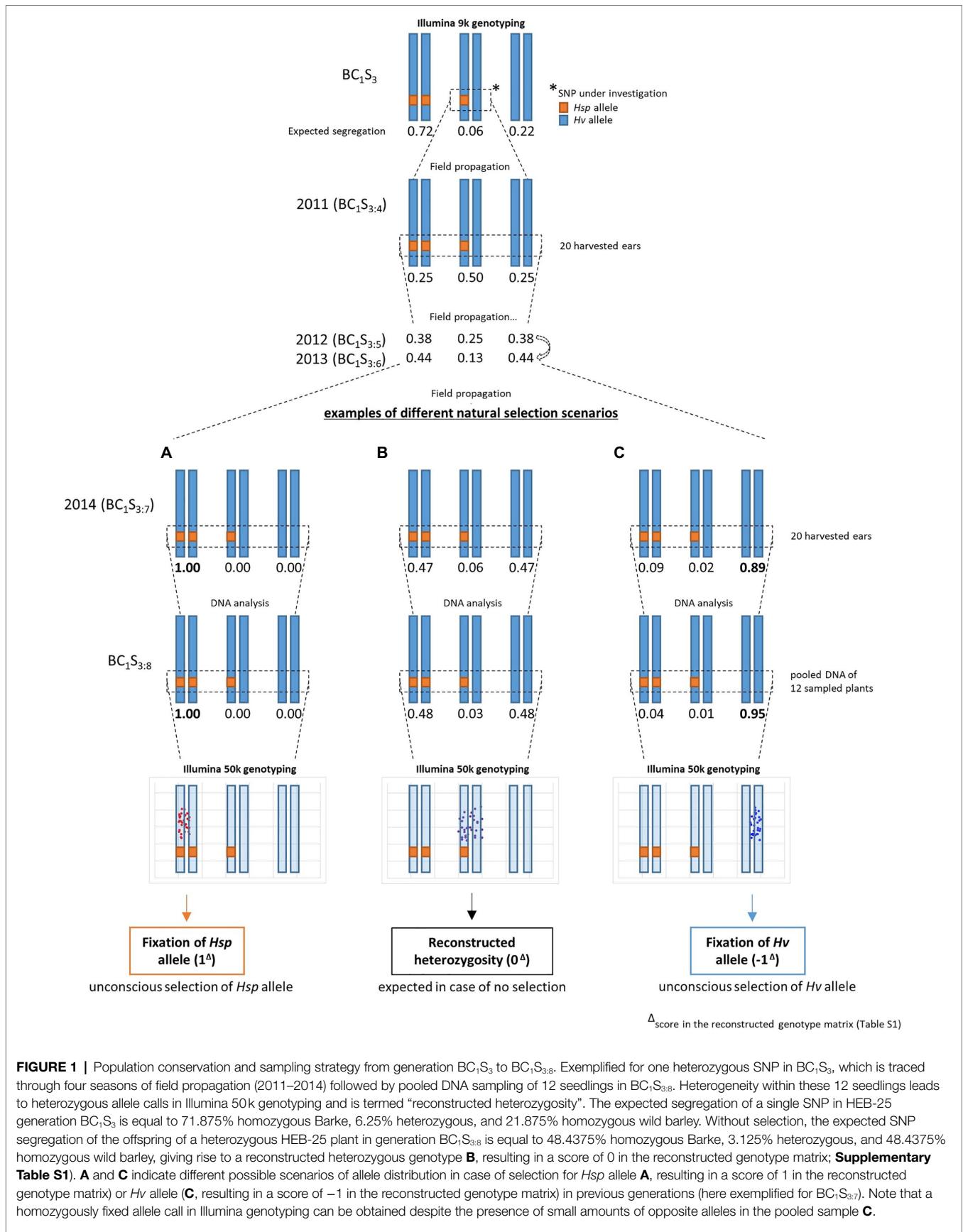
FIGURE 1 | Population conservation and sampling strategy from generation BC_1S_3 to $BC_1S_{3:8}$. Exemplified for one heterozygous SNP in BC_1S_3, which is traced through four seasons of field propagation (2011–2014) followed by pooled DNA sampling of 12 seedlings in $BC_1S_{3:8}$. Heterogeneity within these 12 seedlings leads to heterozygous allele calls in Illumina 50k genotyping and is termed "reconstructed heterozygosity". The expected segregation of a single SNP in HEB-25 generation BC_1S_3 is equal to 71.875% homozygous Barke, 6.25% heterozygous, and 21.875% homozygous wild barley. Without selection, the expected SNP segregation of the offspring of a heterozygous HEB-25 plant in generation $BC_1S_{3:8}$ is equal to 48.4375% homozygous Barke, 3.125% heterozygous, and 48.4375% homozygous wild barley, giving rise to a reconstructed heterozygous genotype **B**, resulting in a score of 0 in the reconstructed genotype matrix; **Supplementary Table S1**). **A** and **C** indicate different possible scenarios of allele distribution in case of selection for *Hsp* allele **A**, resulting in a score of 1 in the reconstructed genotype matrix) or *Hv* allele (**C**, resulting in a score of −1 in the reconstructed genotype matrix) in previous generations (here exemplified for $BC_1S_{3:7}$). Note that a homozygously fixed allele call in Illumina genotyping can be obtained despite the presence of small amounts of opposite alleles in the pooled sample **C**.

Determination of Allele Fixation at Originally Heterozygous SNPs

Only heterozygous loci in BC_1S_3 generation were screened for their allele state in $BC_1S_{3:8}$. The allele fixation rate (AFR) for each SNP was determined as

$$AFR_{SNP} = \frac{\text{Number of homozygous calls across all initially heterozygous HEB lines in } BC_1S_{3:8}}{\text{Number of heterozygous calls across all HEB lines in } BC_1S_3}.$$

To determine the relative fixation direction (RFD) in which an originally heterozygous SNP allele moved, a reconstructed genotype matrix was created (**Supplementary Table S1**) containing fixation values for each SNP*genotype combination where "0" represents SNPs that retained their (reconstructed) heterozygosity state, i.e., both alleles are still present in the 12 sampled plants of a $BC_1S_{3:8}$ line, "-1" represents SNPs that were fixed for the homozygous *Hv* allele, and "+1" represents SNPs that were fixed for the homozygous *Hsp* allele. The RFD for each SNP was then determined as

$$RFD_{SNP} = \frac{\text{Sum of fixation values}\left(\text{i.e.} -1,0,1\right)\text{ across all initially heterozygous HEB lines in } BC_1S_{3:8}}{\text{Number of heterozygous calls across all HEB lines in } BC_1S_3}.$$

Both AFR and RFD were only calculated for SNPs containing at least 10 heterozygous HEB lines in generation BC_1S_3 (n = 3,872 SNPs) to avoid strong bias.

To test for significance of RFD, chi-square goodness-of-fit tests were conducted to measure deviations from a 1:1 ratio of homozygous *Hv* to homozygous *Hsp* genotypes in $BC_1S_{3:8}$. For this purpose, a sliding window approach summarizing allele counts of 10 consecutive SNPs was applied. A significant deviation was accepted at a Bonferroni-corrected value of $p < 0.01$.

Pearson's correlations of RFD with *Hsp* allele SNP effects estimated in the whole HEB-25 population were calculated. For this purpose, SNP effects were obtained from a simple linear model regressing published data on plant height and flowering time (Maurer et al., 2016), ear number, grain number per ear, grain yield, and threshability (Herzig et al., 2019), as well as unpublished data of powdery mildew susceptibility and brittleness of rachis on the quantitative SNP scores (matrix D of Maurer and Pillen (2019)).

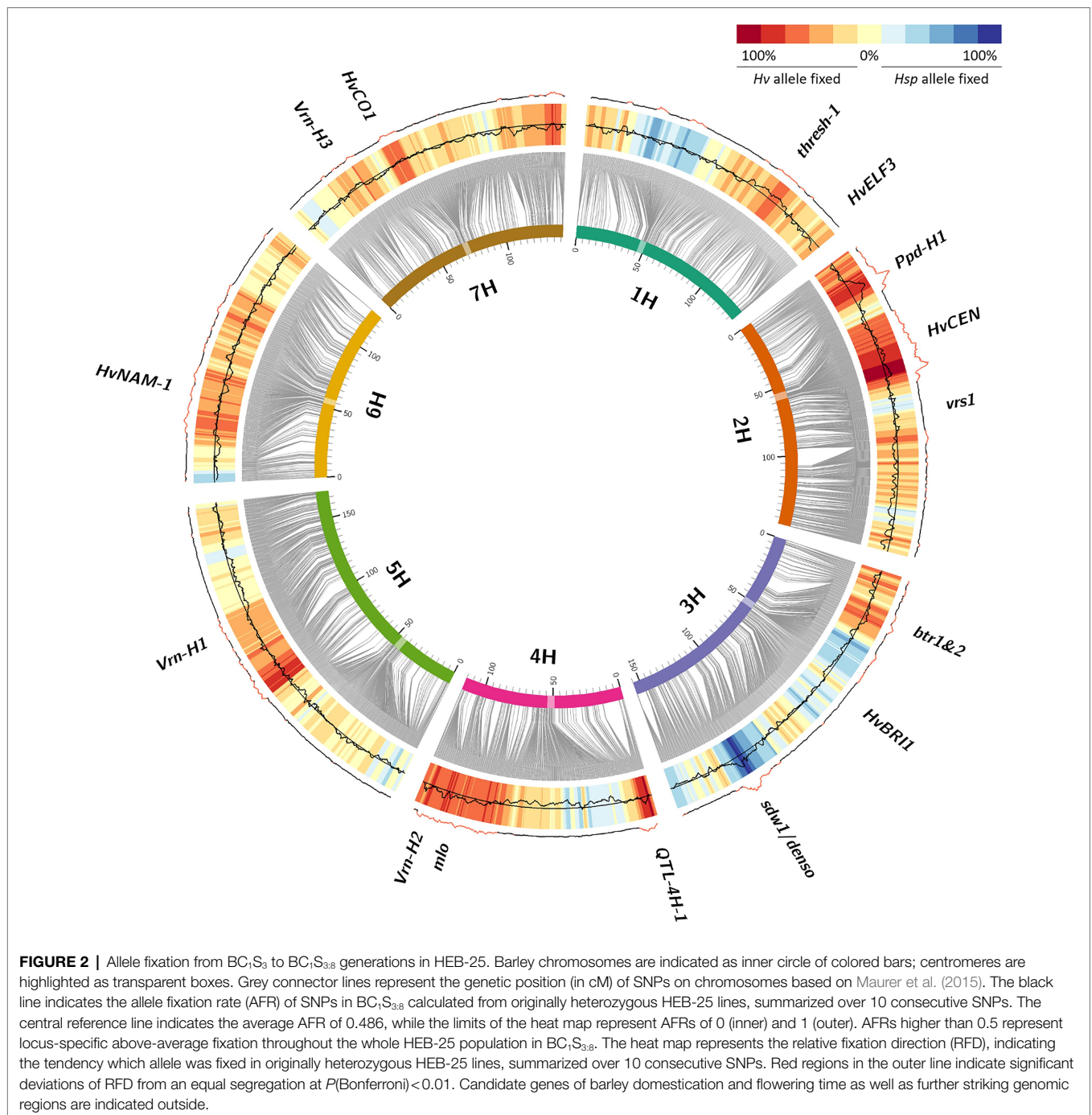## Expected Probabilities of Allele Fixation During Population Conservation

The expected segregation ratio in the BC_1S_3 generation of HEB-25 is 0.71875: 0.0625: 0.21875 for homozygous *Hv* allele, heterozygous, and homozygous *Hsp* allele, respectively. Those 6.25% of heterozygous SNPs will segregate in advanced generations. In each selfing generation, the heterozygous rate is halved and a quarter is going to be fixed in each of the two homozygous classes. The progeny of a single BC_1S_3 plant, which is termed $BC_1S_{3:4}$, is therefore expected to segregate in a 0.25: 0.5: 0.25 ratio at an initially heterozygous SNP. By following this rule, a segregation of 0.484375: 0.03125: 0.484375

at initially heterozygous loci is expected for $BC_1S_{3:8}$ generation, if no selection occurs during reproduction and plant cultivation (**Figure 1B**). Consequently, at an initially heterozygous locus, an unbiased reconstructed genotype in $BC_1S_{3:8}$ should consist of a pooled DNA sample of ~6 *Hv* and ~6 *Hsp* plants, giving rise to a heterozygous allele call in 50k genotyping. This way, the initial heterozygosity can be reconstructed. In case of no selection, 100% heterozygous calls are expected in the pooled sample of 12 $BC_1S_{3:8}$ plants originating from an initially heterozygous BC_1S_3 plant.

To estimate the impact of accidentally biased sampling on genotype calling in $BC_1S_{3:8}$ generation, the probabilities of obtaining a biased pooled sample of 12 plants, used for 50k genotyping, were determined in 1,000,000 binomial trials with regard to different simulated allele segregations (0.05–0.95) of the previously harvested generation. The probabilities of accidental sampling of 12 or ≥ 9 plants of the same homozygous allele class, leading to homozygous genotype calls in Illumina genotyping, were then estimated for each preset allele segregation.

RESULTS

In total, 214,979 (6,02%) heterozygous SNP calls were obtained from 3,573,466 polymorphic SNP assays in generation BC_1S_3 of population HEB-25 (**Supplementary Table S1**), which is close to the expected frequency of 6.25%. This represents an average number of ≈ 46 heterozygous HEB lines per SNP. Altogether, 104,725 (48.7%) SNP calls out of those initially heterozygous SNPs were homozygously fixed after five cycles of selfing in $BC_1S_{3:8}$ generation, based on SNP analysis of a pooled sample of 12 plants for each $BC_1S_{3:8}$ line. This is also visible in the frequency distribution of the allele fixation rate (AFR) of each single SNP (**Supplementary Figure S1**; **Supplementary Table S1**), indicating that on average originally heterozygous SNPs did not segregate equally into both allele classes in half of the lines (average AFR = 48.6%). In 42,752 (19.9%) cases, the SNP was fixed toward the *Hsp* allele, while in 61,973 (28.8%) cases, the SNP was fixed toward the *Hv* allele (**Supplementary Table S1**). Mapping these tendencies on the genome revealed clear patterns of genomic regions where the *Hsp* allele and the *Hv* allele were favored, respectively (**Figure 2**), enabling to separate the observed effects from genetic drift, which would not favor a specific allele and therefore would result in light yellow colors on the heat map. We expected that Barke (Hv), as a semi-dwarf European spring barley cultivar with resistance to powdery mildew, should be well adapted to the environmental conditions during propagation in the spring-sown trials in Halle. The obviously relevant loci *Ppd-H1*, *Vrn-H2*, mlo, and sdw1/denso were all under selection (**Figure 2**). AFR was highest for *Ppd-H1* (>90%) and showed a clear tendency of fixation for the Hv allele (negative RFD). This was also true for the Vrn-H2 and mlo loci conferring vernalization independency and resistance to powdery mildew, respectively. However, at denso/sdw1, the Hsp allele was clearly favored (positive RFD).

FIGURE 2 | Allele fixation from BC_1S_3 to $BC_1S_{3:8}$ generations in HEB-25. Barley chromosomes are indicated as inner circle of colored bars; centromeres are highlighted as transparent boxes. Grey connector lines represent the genetic position (in cM) of SNPs on chromosomes based on Maurer et al. (2015). The black line indicates the allele fixation rate (AFR) of SNPs in $BC_1S_{3:8}$ calculated from originally heterozygous HEB-25 lines, summarized over 10 consecutive SNPs. The central reference line indicates the average AFR of 0.486, while the limits of the heat map represent AFRs of 0 (inner) and 1 (outer). AFRs higher than 0.5 represent locus-specific above-average fixation throughout the whole HEB-25 population in $BC_1S_{3:8}$. The heat map represents the relative fixation direction (RFD), indicating the tendency which allele was fixed in originally heterozygous HEB-25 lines, summarized over 10 consecutive SNPs. Red regions in the outer line indicate significant deviations of RFD from an equal segregation at P(Bonferroni) < 0.01. Candidate genes of barley domestication and flowering time as well as further striking genomic regions are indicated outside.

DISCUSSION

Technical Factors Affecting Allele Fixation

One challenge of conserving extensive experimental plant populations with remaining heterozygosity is to develop a strategy to maintain the population in an effective manner. We applied a method of maintaining the barley NAM population HEB-25 by harvesting 20 randomly selected representative ears from each of the 1,420 lines during four seasons of field propagation. By comparing genotype data of the original

BC_1S_3 plants with a pooled DNA sample of the resulting progeny in $BC_1S_{3:8}$, we observed that the original heterozygosity present in the HEB lines was halved as indicated by the reconstructed heterozygosity in $BC_1S_{3:8}$. In other words, half of the lines segregating for a specific SNP locus were fixed toward a homozygous allele during population conservation. To interpret this number, one has to consider that the obtained genotype score is the product of leaf sampling for DNA extraction and subsequent microarray-based SNP genotyping and allele calling. Since the initial heterozygosity declines

after five selfing generations, we collected leaf material of 12 plants to reconstruct heterozygosity. Noteworthy, the selection of 12 plants itself harbors the potential to bias the true heterozygosity score if by chance only plants with the same homozygous genotype are collected. However, the probability that either only *Hv* or only *Hsp* genotypes are collected in this way is <0.2% for each homozygous allele class (**Supplementary Table S2**), assuming that both alleles occur in an equal expected proportion of 0.484375 in $BC_1S_{3:8}$ (**Supplementary Figure S2**). In microarray-based SNP genotyping, allele classes are defined with a certain tolerance for the fluorescence signals of both tested nucleotides. This means that a homozygous genotype call can be obtained even if there are up to ~20% of the opposite allele in the pooled DNA sample. If taking this into account, the probability of allele fixation is technically raised, though still rather low (<6% for each homozygous allele class; **Supplementary Table S2**). The observed AFR of ~0.5 could only be realized if the fractions of harvested seeds from the previous generation(s) are skewed (**Supplementary Figure S2**), either by natural or artificial selection. It must be noted that in addition to the previous remarks also the fact matters that the cluster file for allele discrimination was revised for the 50k array, possibly leading to a *per se* reduced rate of heterozygous calls. This can explain the relatively high AFR of ~0.5 throughout the whole genome without indication of a specifically selected locus.

Sources of Selection Pressure During Population Conservation

Deviations from this background noise may hint on a certain selective pressure. During field propagation of a population, many natural sources of selection pressure may exist. For instance, under extreme environmental conditions, specific genotypes may be lost if they cannot cope with existing selection pressure. Furthermore, plant–plant competition occurs in a plot, leading to an unequal contribution of alleles to subsequent generations. Lower sowing densities could help to mitigate this competition to maintain the genotype during plot propagation. However, also artificial selection can occur during maintaining a population, for instance by unconsciously harvesting the most vital looking ears or by chance harvesting unequal fractions of ears of both genotype classes. Therefore, for future studies, we recommend harvesting whole plots rather than a sample of each plot to avoid a potential source of artificial selection. In our case, we assume that a mixture of both sources of selection acted on HEB-25 lines during field cultivation.
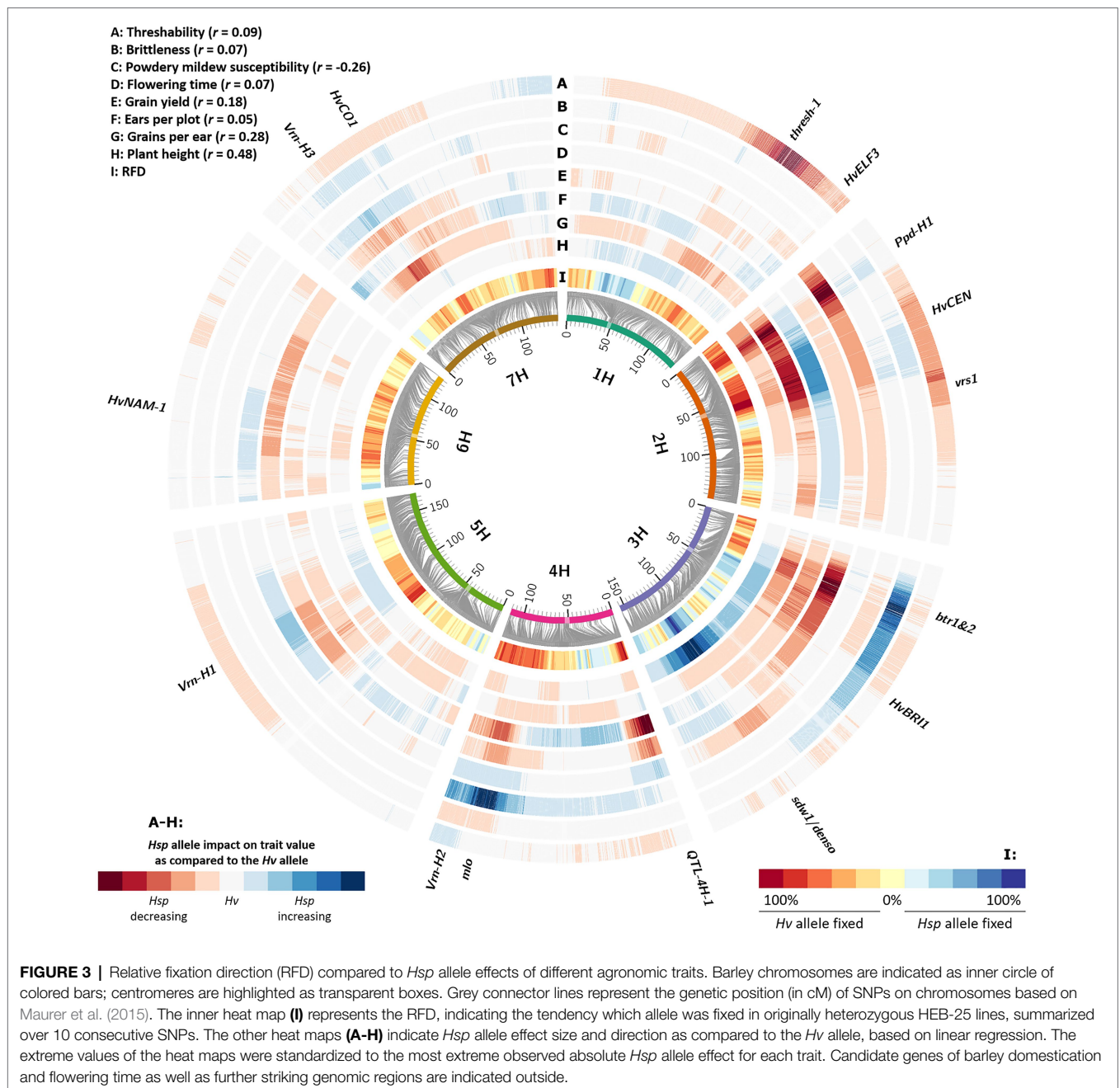
Genomic Regions Under Selection

Interestingly, we observed a systematic pattern of genomic regions being fixed either toward *Hv* or *Hsp* alleles (**Figure 2**; **Supplementary Table S1**). Those genomic regions are likely the reasons for selection events that occurred during population conservation. Assigning candidate genes to these regions revealed that most of the loci correspond to well-known

genes of barley domestication. Since HEB-25 results from crosses of wild barley and the domesticated elite barley cultivar Barke, this finding can be interpreted as a short story of domestication in barley. Supporting this, there was a clear tendency of selection against the dominant *Btr1/Btr2* allele on chromosome 3H (Pourkheirandish et al., 2015) conferring a brittle rachis and shattering of the ear of wild barley at maturity. In $BC_1S_{3:8}$ lines, the genomic region of *btr1/btr2* was predominantly fixed toward the cultivated *Hv* allele. Plausibly, predominately intact ears without brittle rachis were harvested, leading to a fixation of non-brittle ears in future generations. This is rather a source of artificial selection due to harvesting than a natural advantage of non-brittle lines. Likewise, another potentially artificially selected genomic region is the threshability locus *thresh-1* on 1H (Schmalenbach et al., 2011). Here, the wild allele causes awns and rachis remaining attached to the grain after mechanical threshing. In contrast, the domesticated dominant allele confers a "clean" grain that can directly be used for sowing in the next season. Although for all grains remaining awns and rachis were manually removed, there might have been the tendency that already "clean" seeds were preferred for next season sowing. Another prominent domestication locus, *vrs1*, leading to the six-rowed ear phenotype in cultivated barley (Komatsuda et al., 2007), could not reliably be captured in our study. We would expect a clear shift toward the six-rowed phenotypes, since the probability that their grains contribute to the next generation is increased due to the higher number of grains per ear. For the sake of completeness, however, it should be mentioned that this effect could be partly compensated by the increased seedling vigor of larger seeds produced in two-rowed ears (Boyd et al., 1971). However, only a single HEB family (F24) shows six-rowed ears, derived from its *H. v.* ssp. *agriocrithon* parent HID380. Only one out of 56 lines of F24 was heterozygous at the *vrs1* locus (BOPA2_12_30897) in BC_1S_3. Therefore, a clear statement for this locus is not possible. However, in general, we observed that loci affecting the grain number per ear in HEB-25 (Herzig et al., 2019) were often associated with genomic regions with an increased AFR. This may indicate that more harvested grains from genotypes carrying a grain number-increasing locus lead to a higher probability of those grains being selected for next season sowing. This tendency is supported by the slightly positive correlation of 0.28 between *Hsp* allele SNP effects for grain number and RFD (**Figure 3**; **Supplementary Figure S3**).
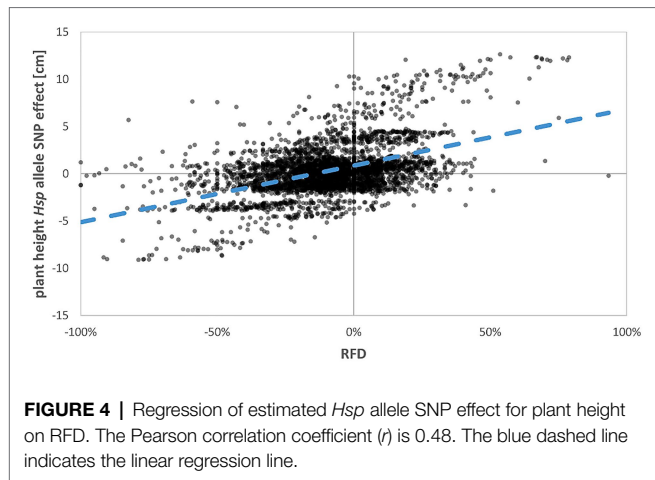
Phenology-Related Effects

Strikingly, many of the selection-affected regions correspond to flowering time and plant development genes (**Figure 2**; **Supplementary Table S1**). As indicated in Maurer et al. (2015) and Maurer et al. (2016), eight major flowering time loci could be identified in HEB-25 that explained large proportions of the variance for flowering time and other developmental traits. Interestingly, most of them showed a clear fixation tendency. However, the direction of fixation

FIGURE 3 | Relative fixation direction (RFD) compared to *Hsp* allele effects of different agronomic traits. Barley chromosomes are indicated as inner circle of colored bars; centromeres are highlighted as transparent boxes. Grey connector lines represent the genetic position (in cM) of SNPs on chromosomes based on Maurer et al. (2015). The inner heat map **(I)** represents the RFD, indicating the tendency which allele was fixed in originally heterozygous HEB-25 lines, summarized over 10 consecutive SNPs. The other heat maps **(A-H)** indicate *Hsp* allele effect size and direction as compared to the *Hv* allele, based on linear regression. The extreme values of the heat maps were standardized to the most extreme observed absolute *Hsp* allele effect for each trait. Candidate genes of barley domestication and flowering time as well as further striking genomic regions are indicated outside.

differed and was not correlated with the accelerating or delaying effects of the *Hsp* alleles (**Figure 3; Supplementary Figure S4**). At these loci predominantly, the *Hv* allele was fixed, except at the *sdw1* locus, where the *Hsp* allele was fixed more often. This finding indicates that the reason for the fixation might be a selection for higher plants rather than flowering time, since *sdw1* is the main determinant of plant height in HEB-25 and Barke carries the semi-dwarf allele. Either higher plants were favored at manual harvesting avoiding strenuous stooping or higher plants impeded growth of semi-dwarf plants early on. The latter is supported by Raggi et al. (2016), who also detected

several plant height associated genetic loci under natural selection in machine-harvested trials with barley composite crosses, and by Chen et al. (2018), who substantiated that taller plants were more competitive in terms of light interception. Most of the obtained fixed loci in the present study co-localize with QTL for plant height in HEB-25 (Maurer et al., 2016), which was underlined by a correlation of $r = 0.48$ between *Hsp* allele effects for plant height at maturity and RFD (**Figures 3, 4**), indicating that alleles increasing plant height were preferentially fixed. Most likely, those alleles increase competitiveness already early on during plant development. In spring wheat, cultivars with increased

FIGURE 4 | Regression of estimated *Hsp* allele SNP effect for plant height on RFD. The Pearson correlation coefficient (*r*) is 0.48. The blue dashed line indicates the linear regression line.

plant height showed an increased weed suppression ability (Murphy et al., 2008). Conscious selection on early competitiveness might be useful to breed new cultivars with increased weed suppression ability enabling a sustainable herbicide-reduced cropping system.

One of the prerequisites for the expansion of barley cultivation during domestication was the development of a spring growth habit. In contrast to wild barley, spring barley has no or reduced vernalization requirement and is insensitive to day length allowing for an extended vegetative growth period under long-day conditions (Pourkheirandish and Komatsuda, 2007). The optimum exploitation of the growth period results in higher grain yield and is, therefore, most likely the reason why major flowering time genes were fixed toward the *Hv* allele, which also indicates that Barke (*Hv*) seems to be phenologically optimally adapted to the investigated environment Halle. The most extreme example in this context is *Ppd-H1*, the main determinant of photoperiod response in barley (Turner et al., 2005). Out of the 78 heterozygous lines in BC_1S_3 74 (95%) were fixed in $BC_1S_{3:8}$ (**Supplementary Table S1**, BK_12). Thereof, 67 lines were fixed toward the *Hv* allele conferring photoperiod insensitivity. This is among the highest AFR observed throughout the whole genome and outlines the importance of *Ppd-H1* for adaptation to Central and Northern European climates. However, interestingly, in 7 lines, the *Hsp* allele of *Ppd-H1* was fixed in $BC_1S_{3:8}$. We assume that this indicates random selection events or is due to simultaneous co-selection for another locus, as we do not see any indication of a specific *Ppd-H1* haplotype being preferred (Maurer et al., 2015). Note that the observed fixation of *Ppd-H1* is environment-specific and would probably be less pronounced or even opposite in other latitudes. For instance, in the Mediterranean region, earliness is key to escape early season terminal drought, which would mean a selection advantage for the *Hsp* allele (Andrés and Coupland, 2012; Wiegmann et al., 2019a; Fernández-Calleja et al., 2021).

Further Interesting Insights

Besides many known domestication loci, also other genomic regions showed specific allele preferences, although the AFR was not higher than the background noise in the rest of the genome. However, the clearly directed fixation at these loci might point to the fixation of alleles in specific families of the NAM population. Examples are the centromere regions of chromosomes 1H and 3H, where a significant tendency toward the *Hsp* allele was observed. This was also true for distal parts of chromosome 3HL and 6HS. All these regions might harbor wild barley alleles that create a positive selection pressure, hinting to promising sources of new allelic variation for future breeding. These wild barley alleles may confer (a)biotic stress tolerance, for instance against drought, which frequently occurs in the studied environment Halle. In 2011, the first year of field propagation, the lowest total precipitation sum until maturity was observed (**Supplementary Figure S5**). However, one has to admit that these loci might also be a source of artificial selection.

In contrast to looking for an increased AFR, also the opposite approach is interesting. Loci with a low fixation rate might hint to alleles conferring hybrid vigor and might represent potential candidates for hybrid barley breeding. In this context, the peri-centromeric region of chromosome 4H and a region on the long arm of chromosome 2H might be promising targets.

CONCLUSION

By screening heterozygosity patterns in a wild × elite barley NAM population, we were able to determine loci affected by selective fixation of either the exotic or the elite barley allele during five cycles of reproduction and field cultivation. The factors causing allele frequency changes through competition in composite crosses are manifold (Blijenburg and Sneep, 1975; Enjalbert et al., 1999; Steele et al., 2004; Phillips and Wolfe, 2005; Raggi et al., 2016). The selection factors may be grouped into (i) natural selection caused by plant–plant competitiveness, phenological advantages as well as superior abiotic and biotic stress resilience and (ii) artificial selection through cultivation and harvesting practices. With our approach, we could unveil the genetic basis of those selection events and define alleles which are superior during reproduction and plant cultivation in the investigated environment. The use of a large segregating mapping population for the temporal screening of heterozygosity enabled to define alleles in distinct genome regions as drivers of adaptive evolution. Defining such superior alleles could help to select chromosomal regions covering potentially beneficial wild alleles conferring, for example, stress tolerance, which is of special importance to cope with drastic upcoming challenges in times of climate change.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can

be directed to the corresponding author. Original genotype matrices that were used for the analysis are available in Maurer et al., (2015) and Maurer and Pillen (2019). The resulting fixation matrix is available in **Supplementary Table S1**.

AUTHOR CONTRIBUTIONS

AM analyzed data and wrote the manuscript. KP coordinated the project and co-wrote the paper. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.764537/full#supplementary-material

REFERENCES

Andrés, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nat. Rev. Genet.* 13, 627–639. doi: 10.1038/nrg3291

Bayer, M. M., Rapazote-Flores, P., Ganal, M., Hedley, P. E., Macaulay, M., Plieske, J., et al. (2017). Development and evaluation of a barley 50k iSelect SNP Array. *Front. Plant Sci.* 8:1792. doi: 10.3389/fpls.2017.01792

Blijenburg, J. G., and Sneep, J. (1975). Natural selection in a mixture of eight barley varieties, grown in six successive years. 1. Competition between the varieties. *Euphytica* 24, 305–315. doi: 10.1007/BF00028195

Boyd, W., Gordon, A., and LaCroix, L. (1971). Seed size, germination resistance and seedling vigor in barley. *Can. J. Plant Sci.* 51, 93–99. doi: 10.4141/cjps71-021

Büttner, B., Draba, V., Pillen, K., Schweizer, G., and Maurer, A. (2020). Identification of QTLs conferring resistance to scald (Rhynchosporium commune) in the barley nested association mapping population HEB-25. *BMC Genomics* 21:837. doi: 10.1186/s12864-020-07258-7

Chen, T. W., Cabrera-Bosquet, L., Alvarez Prado, S., Perez, R., Artzet, S., Pradal, C., et al. (2018). Genetic and environmental dissection of biomass accumulation in multi-genotype maize canopies. *J. Exp. Bot.* 70, 2523–2534. doi: 10.1093/jxb/ery309

Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., et al. (2012). Natural variation in a homolog of antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44, 1388–1392. doi: 10.1038/ng.2447

Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., and Guarino, L. (2017). Past and future use of wild relatives in crop breeding. *Crop Sci.* 57, 1070–1082. doi: 10.2135/cropsci2016.10.0885

Döring, T. F., Knapp, S., Kovacs, G., Murphy, K., and Wolfe, M. S. (2011). Evolutionary plant breeding in cereals—into a new era. *Sustain. For.* 3, 1944–1971. doi: 10.3390/su3101944

Dreissig, S., Maurer, A., Sharma, R., Milne, L., Flavell, A. J., Schmutzer, T., et al. (2020). Natural variation in meiotic recombination rate shapes introgression patterns in intraspecific hybrids between wild and domesticated barley. *New Phytol.* 228, 1852–1863. doi: 10.1111/nph.16810

Enjalbert, J., Goldringer, I., Paillard, S., and Brabant, P. (1999). Molecular markers to study genetic drift and selection in wheat populations. *J. Exp. Bot.* 50, 283–290. doi: 10.1093/jxb/50.332.283

Fernández-Calleja, M., Casas, A. M., and Igartua, E. (2021). Major flowering time genes of barley: allelic diversity, effects, and comparison with wheat. *Theor. Appl. Genet.* 134, 1867–1897. doi: 10.1007/s00122-021-03824-z

Guo, T., Mu, Q., Wang, J., Vanous, A. E., Onogi, A., Iwata, H., et al. (2020). Dynamic effects of interacting genes underlying rice flowering-time phenotypic plasticity and global adaptation. *Genome Res.* 30, 673–683. doi: 10.1101/gr.255703.119

Herzig, P., Backhaus, A., Seiffert, U., von Wirén, N., Pillen, K., and Maurer, A. (2019). Genetic dissection of grain elements predicted by hyperspectral imaging associated with yield-related traits in a wild barley NAM population. *Plant Sci.* 285, 151–164. doi: 10.1016/j.plantsci.2019.05.008

Herzig, P., Maurer, A., Draba, V., Sharma, R., Draicchio, F., Bull, H., et al. (2018). Contrasting genetic regulation of plant development in wild barley grown in two European environments revealed by nested association mapping. *J. Exp. Bot.* 69, 1517–1531. doi: 10.1093/jxb/ery002

Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvel, P., Kanamori, H., Perovic, D., et al. (2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci.* 104, 1424–1429. doi: 10.1073/pnas.0608580104

Li, X., Guo, T., Wang, J., Bekele, W. A., Sukumaran, S., Vanous, A. E., et al. (2021). An integrated framework reinstating the environmental dimension for GWAS and genomic selection in crops. *Mol. Plant* 14, 874–887. doi: 10.1016/j.molp.2021.03.010

Maurer, A., Draba, V., Jiang, Y., Schnaithmann, F., Sharma, R., Schumann, E., et al. (2015). Modelling the genetic architecture of flowering time control in barley through nested association mapping. *BMC Genomics* 16:290. doi: 10.1186/s12864-015-1459-7

Maurer, A., Draba, V., and Pillen, K. (2016). Genomic dissection of plant development and its impact on thousand grain weight in barley through nested association mapping. *J. Exp. Bot.* 67, 2507–2518. doi: 10.1093/jxb/erw070

Maurer, A., and Pillen, K. (2019). "50k Illumina Infinium iSelect SNP Array data for the wild barley NAM population HEB-25". e!DAL - Plant Genomics and Phenomics Research Data Repository (PGP), IPK Gatersleben, Seeland OT Gatersleben, Corrensstraße 3, 06466, Germany).

McCouch, S., Baute, G. J., Bradeen, J., Bramel, P., Bretting, P. K., Buckler, E., et al. (2013). Agriculture: feeding the future. *Nature* 499, 23–24. doi: 10.1038/499023a

Mehnaz, M., Dracatos, P., Pham, A., March, T., Maurer, A., Pillen, K., et al. (2021). Discovery and fine mapping of Rph28: a new gene conferring resistance to Puccinia hordei from wild barley. *Theor. Appl. Genet.* 134, 2167–2179. doi: 10.1007/s00122-021-03814-1

Merchuk-Ovnat, L., Silberman, R., Laiba, E., Maurer, A., Pillen, K., Faigenboim, A., et al. (2018). Genome scan identifies flowering-independent effects of barley HsDry2.2 locus on yield traits under water deficit. *J. Exp. Bot.* 69, 1765–1779. doi: 10.1093/jxb/ery016

Murphy, K. M., Dawson, J. C., and Jones, S. S. (2008). Relationship among phenotypic growth traits, yield and weed suppression in spring wheat landraces and modern cultivars. *Field Crop Res.* 105, 107–115. doi: 10.1016/j.fcr.2007.08.004

Pham, A.-T., Maurer, A., Pillen, K., Brien, C., Dowling, K., Berger, B., et al. (2019). Genome-wide association of barley plant growth under drought stress using a nested association mapping population. *BMC Plant Biol.* 19:134. doi: 10.1186/s12870-019-1723-0

Pham, A.-T., Maurer, A., Pillen, K., Taylor, J., Coventry, S., Eglinton, J. K., et al. (2020). Identification of wild barley derived alleles associated with plant development in an Australian environment. *Euphytica* 216:148. doi: 10.1007/s10681-020-02686-8

Phillips, S., and Wolfe, M. (2005). Evolutionary plant breeding for low input systems. *J. Agric. Sci.* 143, 245–254. doi: 10.1017/S0021859605005009

Pourkheirandish, M., Hensel, G., Kilian, B., Senthil, N., Chen, G., Sameri, M., et al. (2015). Evolution of the grain dispersal system in barley. *Cell* 162, 527–539. doi: 10.1016/j.cell.2015.07.002

Pourkheirandish, M., and Komatsuda, T. (2007). The importance of barley genetics and domestication in a global perspective. *Ann. Bot.* 100, 999–1008. doi: 10.1093/aob/mcm139

Raggi, L., Ceccarelli, S., and Negri, V. (2016). Evolution of a barley composite cross-derived population: an insight gained by molecular markers. *J. Agric. Sci.* 154, 23–39. doi: 10.1017/S0021859614001269

Saade, S., Maurer, A., Shahid, M., Oakey, H., Schmöckel, S. M., Negrão, S., et al. (2016). Yield-related salinity tolerance traits identified in a nested association mapping (NAM) population of wild barley. *Sci. Rep.* 6:32586. doi: 10.1038/srep32586

Schmalenbach, I., March, T. J., Bringezu, T., Waugh, R., and Pillen, K. (2011). High-resolution genotyping of wild barley introgression lines and fine-mapping of the Threshability locus thresh-1 using the Illumina GoldenGate assay. *G3: Genes Genomes, Genetics* 1, 187–196. doi: 10.1534/g3.111.000182

Sharma, R., Draicchio, F., Bull, H., Herzig, P., Maurer, A., Pillen, K., et al. (2018). Genome-wide association of yield traits in a nested association mapping population of barley reveals new gene diversity for future breeding. *J. Exp. Bot.* 69, 3811–3822. doi: 10.1093/jxb/ery178

Steele, K. A., Edwards, G., Zhu, J., and Witcombe, J. R. (2004). Marker-evaluated selection in rice: shifts in allele frequency among bulks selected in contrasting agricultural environments identify genomic regions of importance to rice adaptation and breeding. *Theor. Appl. Genet.* 109, 1247–1260. doi: 10.1007/s00122-004-1732-7

Suneson, C. A. (1956). An evolutionary plant breeding method. *Agron. J.* 48, 188–191. doi: 10.2134/agronj1956.00021962004800040012x

Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063

Turner, A., Beales, J., Faure, S., Dunford, R. P., and Laurie, D. A. (2005). The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science* 310, 1031–1034. doi: 10.1126/science.1117619

Vatter, T., Maurer, A., Kopahnke, D., Perovic, D., Ordon, F., and Pillen, K. (2017). A nested association mapping population identifies multiple small effect QTL conferring resistance against net blotch (Pyrenophora teres f. teres) in wild barley. *PLoS One* 12:e0186803. doi: 10.1371/journal.pone.0186803

Vatter, T., Maurer, A., Perovic, D., Kopahnke, D., Pillen, K., and Ordon, F. (2018). Identification of QTL conferring resistance to stripe rust (Puccinia striiformis f. sp. hordei) and leaf rust (Puccinia hordei) in barley using nested association mapping (NAM). *Plos One* 13:e0191666. doi: 10.1371/journal.pone.0191666

Wiegmann, M., Maurer, A., Pham, A., March, T. J., Al-Abdallat, A., Thomas, W. T. B., et al. (2019a). Barley yield formation under abiotic stress depends on the interplay between flowering time genes and environmental cues. *Sci. Rep.* 9:6397. doi: 10.1038/s41598-019-42673-1

Wiegmann, M., Thomas, W. T. B., Bull, H. J., Flavell, A. J., Zeyner, A., Peiter, E., et al. (2019b). Wild barley serves as a source for biofortification of barley grains. *Plant Sci.* 283, 83–94. doi: 10.1016/j.plantsci.2018.12.030

Zamir, D. (2008). Plant breeders go back to nature. *Nat. Genet.* 40, 269–270. doi: 10.1038/ng0308-269

Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., and Song, B.-H. (2017). Back into the wild—apply untapped genetic diversity of wild relatives for crop improvement. *Evol. Appl.* 10, 5–24. doi: 10.1111/eva.12434

Zhao, S., Jing, W., Samuels, D. C., Sheng, Q., Shyr, Y., and Guo, Y. (2018). Strategies for processing and quality control of Illumina genotyping arrays. *Brief. Bioinform.* 19, 765–775. doi: 10.1093/bib/bbx012