Analysis of historical selection in winter wheat

Yang, Chin Jian; Ladejobi, Olufunmilayo; Mott, Richard; Powell, Wayne; Mackay, Ian

Published in:
Theoretical and Applied Genetics

DOI:
10.1007/s00122-022-04163-3

Print publication: 01/09/2022

Document Version
Publisher's PDF, also known as Version of record

Theoretical and Applied Genetics
https://doi.org/10.1007/s00122-022-04163-3

ORIGINAL ARTICLE

Analysis of historical selection in winter wheat

Chin Jian Yang[1] · Olufunmilayo Ladejobi[2] · Richard Mott[2] · Wayne Powell[1] · Ian Mackay[1,3]

Received: 23 April 2022 / Accepted: 22 June 2022

Abstract

Key Message **Modeling of the distribution of allele frequency over year of variety release identifies major loci involved in historical breeding of winter wheat.**

Abstract Winter wheat is a major crop with a rich selection history in the modern era of crop breeding. Genetic gains across economically important traits like yield have been well characterized and are the major force driving its production. Winter wheat is also an excellent model for analyzing historical genetic selection. As a proof of concept, we analyze two major collections of winter wheat varieties that were bred in Western Europe from 1916 to 2010, namely the Triticeae Genome (TG) and WAGTAIL panels, which include 333 and 403 varieties, respectively. We develop and apply a selection mapping approach, Regression of Alleles on Years (RALLY), in these panels, as well as in simulated populations. RALLY maps loci under sustained historical selection by using a simple logistic model to regress allele counts on years of variety release. To control for drift-induced allele frequency change, we develop a hybrid approach of genomic control and delta control. Within the TG panel, we identify 22 significant RALLY quantitative selection loci (QSLs) and estimate the local heritabilities for 12 traits across these QSLs. By correlating predicted marker effects with RALLY regression estimates, we show that alleles whose frequencies have increased over time are heavily biased toward conferring positive yield effect, but negative effects in flowering time, lodging, plant height and grain protein content. Altogether, our results (1) demonstrate the use of RALLY to identify selected genomic regions while controlling for drift, and (2) reveal key patterns in the historical selection in winter wheat and guide its future breeding.

Introduction

Modern agriculture benefits from long-standing breeding effort in creating new and improved crop varieties over time. Genetic gain is often used as a measure of the success in breeding for trait improvement. For example, in wheat, the genetic gains in yield and other agriculturally valuable traits have been well quantified (Mackay et al. 2011, Tadesse et al. 2019 and Shorinola et al. 2022). The introduction of genomic selection (GS) (Meuwissen et al. 2001) in breeding programs further shortens breeding cycles, improves selection accuracy and intensity, and accelerates genetic gain (Voss-Fels et al. 2018). Lastly, genetic gain is further increased by the rise of knowledge exchange between plant and animal breeding through GS (Hickey et al. 2017).

In recent years, there has been a growing interest in mapping quantitative selection loci (QSLs) that are associated with genetic gain independently of any phenotype. This trait-free mapping approach typically involves correlating continuous variables, such as year of variety release and geographical parameters, to genomic markers in a historical variety dataset. Conceptually, it is similar to selection mapping which tests for selection signatures among genomic markers using population genetic models (Johnsson 2018). This approach has been variously named as Birth Date Selection Mapping (Decker et al. 2012), Generation Proxy Selection Mapping (Rowan et al. 2021) and environmental GWAS (EnvGWAS) (Li et al. 2020; Sharma et al. 2021). Here, we will refer to it as EnvGWAS because the underlying mixed linear model is no different from a conventional genome-wide association study (GWAS).

Communicated by Antonio Augusto Franco Garcia.

✉ Ian Mackay
ian.mackay@sruc.ac.uk

[1] Scotland's Rural College (SRUC), Kings Buildings, West Mains Road, Edinburgh EH9 3JG, UK

[2] Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK

[3] IMplant Consultancy Ltd, Chelmsford, UK

Published online: 21 July 2022

Related to EnvGWAS, EigenGWAS uses eigenvectors (principal components) of the genomic relationship matrix as the dependent variable in a mixed linear model (Li et al. 2020; Sharma et al. 2021). The EigenGWAS approach may yield similar results to EnvGWAS if the dependent variables in EnvGWAS are correlated strongly with any eigenvector. Otherwise, EigenGWAS may identify additional QSLs where it incorporates variables that have not been quantified directly. A key confounding factor for determining whether a locus has been under sustained historical selection or drift is that varieties are linked by a complex historical pedigree and unequal relatedness. By correcting for population structure using a mixed linear model (Yu et al. 2006), year effects and principal components that are associated with drift can be controlled in EnvGWAS and EigenGWAS, respectively.

Here, we introduce a new application of an old method by modeling allele frequency change over years in a historical variety dataset. This trait-free method, termed Regression of Alleles on Years (RALLY), fits a logistic regression to model the allele count as a dependent variable and the year of variety release as an independent variable. A logistic model is commonly used in case–control studies where the dependent variables are binary traits of whether an individual is diseased and the independent variables are test factors (Prentice and Pyke 1979). A logistic model is appropriate because changes in allele frequencies are small when the starting frequencies are near the extrema, and large when they are intermediate. Therefore, the power of detection in RALLY is highest when the allele introduction and fixation are fully captured over the years. In addition, the model response is bounded asymptotically by 0 and 1. The dependent and independent variables are switched between RALLY and EnvGWAS. EnvGWAS estimates the rate of increase in year of variety release for a unit change in allele frequency, *i.e.,* from 0 to 1. In contrast, RALLY estimates the mean of allele counts for each year, which is equivalent to the allele frequency for a given year. Recently, Looseley et al. (2020) applied a similar approach to RALLY on significant GWAS markers in a historical barley variety dataset. RALLY is a genome-wide approach that employs parametric control (PC) as a correction to drift-induced allele frequency change. PC is a novel hybrid approach of genomic control (GC) (Devlin and Roeder 1999) and delta control (DC) (Gorroochurn et al. 2006), which are two common control approaches against population structure in human GWAS studies without the need for a mixed linear model.

Our analyses in the simulated and historical variety datasets demonstrate the usefulness of RALLY in mapping QSLs. We begin the evaluation of RALLY in simulated populations where the truth is known, both with and without selection, to quantify RALLY detection power and limit. We use the simulations to calibrate PC, which is then applied to the two historical winter wheat datasets, namely the panels of Triticeae Genome (TG) (Bentley et al. 2014) and Wheat Association Genetics for Trait Advancement and Improvement of Lineages (WAGTAIL) (Fradgley et al. 2019). The WAGTAIL panel is used only as a replicate RALLY analysis. Within the TG panel, we identify 22 RALLY QSLs and compare them to the GWAS QTLs from Ladejobi et al. (2019). Some notable QSLs include one in 2B which coincides with *Ppd-B1* (Mohler et al. 2004), *Yr7*/*Yr5*/*YrSP* (Marchal et al. 2018) and alien introgression from *Triticum timopheevii* (Tsilo et al. 2008; Martynov et al. 2018), as well as another in 6A that coincides with *TaGW2* (Su et al. 2011), *Rht24* (Würschum et al. 2017) and *Rht25* (Mo et al. 2018). To further support the RALLY QSLs, we show that all 22 QSLs have nonzero local heritabilities for at least one trait. Next, we find clear directional selection in traits like flowering time, lodging, yield, plant height and grain protein content by comparing the signs of predicted marker allele effects with their directions of allele frequency change as given by RALLY. By extending the results to pairs of traits, we identify the selection priorities. For example, more ears with lighter grains have been preferred over fewer ears with heavier grains. Finally, we employ the multivariate breeder's equation (Lande and Arnold 1983) to estimate selection parameters, although our results suggest a limited use in modern crops, in contrast to its original application in evolutionary studies. Overall, we have shown that many major genomic regions have been extensively used in winter wheat breeding and we suggest that future selection should emphasize on improving other unexplored genomic regions.

Materials and methods

Population simulation with and without selection

We initiated our population simulation (Fig. 1) in a fictitious species with 10 chromosomes. The genetic lengths of the chromosomes were set from 100 to 280 centiMorgans (cM) with an increment of 20 cM in subsequent chromosomes. The populations spanned over 50 generations (years) with and without selection. All the simulations were performed using the "AlphaSimR" package (Gaynor et al. 2021) in R (R Core Team 2021). We first created 32 inbred founders using the *runMacs* function and we placed one marker (segregating site) at every 0.1 cM, which totals to 19,000 markers. Two causal quantitative trait loci (QTLs) were chosen randomly from the markers at each frequency ranging from 1/32 to 16/32, which resulted in 32 QTLs. The QTL effects were drawn such that rarer QTL alleles have larger effects than more common QTL alleles, which ensure that individual QTLs have similar variances in the founders regardless of their allele frequencies (Coster et al. 2010). This QTL effect model is consistent with

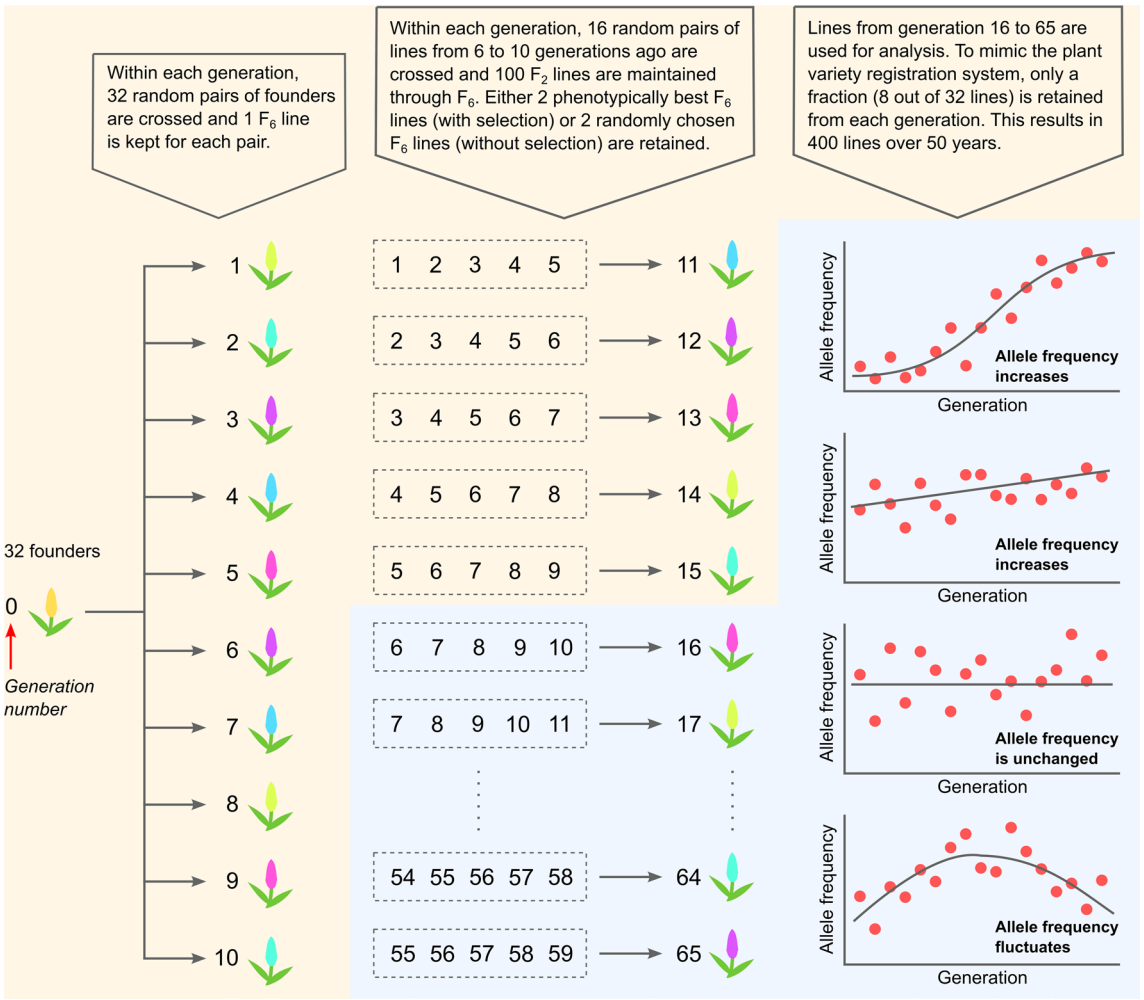

Fig. 1 Population simulation and changes in allele frequency over time. The simulated populations with and without selection are described in detail here. The first 15 generations were used as burn-ins and discarded. Eight varieties from each generation starting at 16 and ending at 65 were randomly chosen to create a population of 400 varieties that span over 50 generations. Examples of how allele frequency changes over time are shown: (1) low to high with a fit of S-shaped logistic curve, (2) moderate to slightly higher with a fit of linear part of the logistic curve, (3) unchanged, and (4) fluctuating. The first two examples are more likely to be significant in RALLY than the last two examples

the combined infinitesimal and rare allele models of QTL effects (Gibson 2012). Specifically, QTLs with frequencies of 1/32, 2/32–3/32, 4/32–7/32 and 8/32–16/32 have the effects of 4, 3, 2 and 1, respectively. We standardized the QTL effects such that the total variance of additive genetic effects is $\sigma_g^2 = 1$, which is calculated using Eq. 1 using a similar approach as Garin et al. (2021). The notations $\sigma_{QTL,i}^2$, $X_{QTL,i}$, β_i and p_i refer to the additive genetic variance, marker allele genotype, marker allele effect and marker allele frequency at the ith QTL. Phenotypic values for each variety were set as a sum of QTL effects and residual effects drawn from a normal distribution of mean 0 and variance 1, which is equivalent to a heritability of 0.5.

$$\sigma_g^2 = \sum_{i=1}^{32} \sigma_{QTL,i}^2 = \sum_{i=1}^{32} Var(X_{QTL,i}\beta_i) = \sum_{i=1}^{32} p_i(1-p_i)\beta_i^2 \quad (1)$$

We created selected (S) and unselected (U) populations from the 32 founders using a simplified model that mimics new variety breeding of major crops in Europe. All varieties were derived as F_6 recombinant inbred lines (RILs) from bi-parental crosses. This is equivalent to 4 generations of single seed descent (SSD) from an F_2 individual. The first 10 generations were created by crossing the initial 32 founders at random. In the subsequent generations, we randomly sampled 32 parents from 6 to 10 generations ago and created 16 bi-parental populations with each having 100 F_6 RILs. By keeping 2 RILs per bi-parental population, we maintained

32 lines at each generation. The 2 RILs were chosen either from the two highest phenotypic values (selected) or randomly (unselected). This step was repeated until the population underwent 55 generations of phenotypic selection. The first 15 generations were discarded as burn-in because none of the parents of the individuals from these 15 generations have been selected, and hence there is no selection-induced allele frequency change. In our simplistic modeling of the plant variety rights (PVRs) system where only a fraction of new lines passing the PVR test, we randomly sampled and retained 8 lines per generation for a total of 400 varieties that spanned over 50 generations. This additional step completed the simulation of a historical variety dataset (selected population) and a control dataset (unselected population), which were used in subsequent analyses.

RALLY and GWAS in simulated populations

We compared the performances of Regression of Alleles on Years (RALLY) and Genome Wide Association Study (GWAS) when applied to the selected and unselected populations. The primary goal of the comparison is to demonstrate the advantages of trait-free mapping approach in populations under selection. Therefore, the expectation here is that RALLY outperforms standard GWAS when selection is involved but not when selection is absent. The model for RALLY was fitted in a logistic regression using the *glm* function in R (R Core Team 2021). The model for GWAS was fitted in a mixed linear model using the *GWASpoly* function in the "GWASpoly" R package (Rosyara et al. 2016).

Briefly, the logistic regression model for RALLY can be shown as below:

$$p_i = \frac{1}{1 + e^{-(\mu_{R,i} + \beta_{R,i} X_R)}} \tag{2}$$

Or, alternatively, by applying the logit link function to Eq. 2:

$$logit(p_i) = \ln\left(\frac{p_i}{1 - p_i}\right) = \mu_{R,i} + \beta_{R,i} X_R \tag{3}$$

where the model terms are described as below:

p_i is the probability of $z_i = 1$, and z_i is a binary variable indicating the absence (0) or presence (1) of an allele at marker i in n varieties.

$\mu_{R,i}$ is the log-odds of p_i in year 0.

$\beta_{R,i}$ is the fixed year effect, or regression coefficient of the year variable.

X_R is the year variable.

The GWAS model is written as below:

$$y = X\beta + Zg + \varepsilon \tag{4}$$

where the model terms are described as below:

y is a vector of trait values in n varieties.

X is a design matrix for the fixed effects such that the first column is a vector of 1's for the mean, the second column is a vector of years, and the third column is a vector of 0's and 2's indicating the count of reference alleles at marker i in n varieties.

β is a vector of fixed effects including the mean effect, year effect and additive genetic effect (β_i) at marker i.

Z is an incidence matrix relating n varieties to observations y.

g is a vector of random genetic background effect with a distribution of $N(0, G)$ where $G = \sigma_g^2 K$.

K is the additive genetic relationship matrix in which the elements $K_{jk} = \frac{\sum_i^m (x_{ij} - p_i)(x_{ik} - p_i)}{\sum_i^m 2p_i(1-p_i)}$; x_{ij} is the marker score for variety j at marker i, x_{ik} is the marker score for variety k at marker i, p_i is the allele frequency at marker i, and m is the total number of markers.

ε is the residual effect with a distribution of $N(0, \sigma_\varepsilon^2 I)$ and I is the identity matrix.

The fixed marker effects are estimated as:

$$\begin{bmatrix} \hat{\beta} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \tag{5}$$

In the given models, the terms of interest are $\beta_{R,i}$ and β_i in RALLY and GWAS, respectively. The term significances are determined by their corresponding standard normal Z -statistics at a Bonferroni corrected threshold of $P = 0.05$. The Bonferroni correction is meant for multiple testing correction in the m markers, which vary slightly across simulations because different markers are excluded in each simulation. Due to how the populations are simulated, some markers may not segregate or have low minor allele frequencies (maf < 0.01) in all the populations. These markers, along with the QTLs and other markers that are highly linked ($r^2 > 0.99$) to QTLs, were excluded from the RALLY and GWAS analyses. The simulations were repeated for 100 iterations and the models were fitted for each simulated population separately.

Model correction by parametric control (PC)

In the previously described naïve RALLY model (Eqs. 2 and 3), the RALLY test statistics may be inflated by population structure arising from consanguinity and population stratification. These factors can prevent a proper separation of markers under selection or drift if they are not addressed. To control for the inflation, we used a combined approach of genomic control (GC) (Devlin and Roeder 1999) and delta control (DC) (Gorroochurn et al. 2006) which we call parametric control (PC). As neither GC nor DC are commonly

used in plant genetics, we provide a brief explanation in subsequent paragraphs before we describe PC in more detail.

In the description of GC given by Mackay and Powell (2007), the population structure adjustment entails dividing the observed $\chi^2_{df=1}$ for a trait-marker association test by a variance inflation factor ($\lambda > 1$). The significance of association is then determined from the p value of the adjusted $\chi^2_{df=1}$, which is calculated as $\chi^2_{df=1,GC} = \chi^2_{df=1}/\lambda$. The inflation is estimated from either a set of reference markers or all markers as the mean of $\chi^2_{df=1}$ from the test of association between a set of null markers and the trait. Alternatively, the ratio of the observed median across m markers to the expected median of 0.456 is often used (Hinrichs et al. 2009) because it is less influenced by small numbers of significant associations arising directly from marker-trait associations (see Eq. 6). This inflation in χ^2 across the genome is judged to arise from the effects of population structure and has been justified by the consideration of the effect of subpopulation differentiation (F_{st}) in inflating the χ^2 statistics (Devlin and Roeder 1999). GC continues to be used in conjunction with other methods of accounting for population structure in association studies to test the effectiveness of those methods and potentially to "mop-up" any residual inflation (e.g., van den Berg et al 2019).

$$\lambda = \frac{median(\chi^2_1, \chi^2_2, \ldots, \chi^2_m)}{0.456} \tag{6}$$

DC was introduced by Gorroochurn et al. (2006) as an alternative to GC. DC addresses the bias in non-centrality parameter (NCP) ($\delta^2 > 0$) instead of the variance inflation factor λ in the χ^2 test. For an association arising from linkage disequilibrium (LD) between a marker and a QTL, the test statistic no longer follows a χ^2 distribution with an NCP of zero, but with a value determined by the magnitude of the QTL effect and the strength of LD. However, an additional increase in the NCP can arise from population structure (Gorroochurn et al. 2006). The bias in NCP can be estimated as the mean χ^2 for the difference in allele frequencies in case and control groups within a set of null markers (Gorroochurn et al. 2006, 2007). Subsequently, DC adjustment is done as $\chi^2_{df=1,DC} = \chi^2_{df=1} - \delta^2$.

However, the original publications on DC (Gorroochurn et al. 2006, 2007) were found to be flawed as it assumes allele frequency changes at the null markers are all in the same direction as the observed change at the test marker (Dadd et al. 2010). A correction for this was proposed in which only null markers with allele frequency changes that approximately matched to the sign of the change at the test marker are used (Gorroochurn et al. 2011). Under drift, a change in allele frequency of either sign is equally likely provided sample sizes and allele frequencies are not extreme. Therefore, we can further improve the DC method by treating positive and negative changes in allele frequency as equally likely and estimate δ using a composite likelihood approach described later.

Working in Z statistics, in the absence of population structure, we expect the null test statistics (Z-scores) to be distributed as $N(0, 1)$. However, in its presence the distribution of the test statistic at a null marker becomes $N(\delta, \sigma)$. As the terms imply, DC controls the inflation in mean δ and GC controls the inflation in standard deviation σ. If we can estimate δ and σ, we can adjust the test statistics as the following:

$$Z_{adj} = \frac{Z - \delta}{\sigma} \tag{7}$$

However, at any individual null allele, Z is equally likely to be positive or negative, since (1) the regression coefficient of one allele has the same magnitude but opposite sign of the other allele, (2) the coding of alleles is arbitrary, and (3) the null alleles are regarded as changing in frequency as a result of drift only, and for modest population sizes and non-extreme allele frequencies positive and negative changes are equally likely. We therefore test for association as:

$$Z_{adj} = \frac{|Z| - |\delta|}{\sigma} \tag{8}$$

We used a maximum likelihood (ML) approach to estimate δ and σ. We can construct a composite likelihood function from two standard normal probability density functions that accounts for positive and negative Z values. The likelihood function is shown in Eq. 9. To simplify the calculation, we used the log likelihood function as described in Eq. 10. We computed δ and σ for an m_0-vector of Z values by either maximizing the log likelihood function, or equivalently, minimizing the negative log likelihood function using the "nlm" package in R (R Core Team 2021). Note that m_0 is the number of null markers.

$$f(Z|\delta, \sigma) = \prod_{i=1}^{m_0} \frac{1}{\sigma\sqrt{2\pi}} \bullet \frac{1}{2} \bullet \left(e^{\frac{-(Z_i-\delta)^2}{2\sigma^2}} + e^{\frac{-(-Z_i-\delta)^2}{2\sigma^2}} \right) \tag{9}$$

$$\begin{aligned} l(\delta, \sigma) =& ln(f(\delta, \sigma)) = m_0 \bullet ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) \\ &+ \sum_{i=1}^{m_0} ln\left[\frac{1}{2}\left(e^{\frac{-(Z_i-\delta)^2}{2\sigma^2}} + e^{\frac{-(-Z_i-\delta)^2}{2\sigma^2}} \right)\right] \end{aligned} \tag{10}$$

We emphasize that the summation of the log likelihood in Eq. 10 is over the set of null markers to estimate δ and σ, which are then used to estimate Z_{adj} in the test set of markers according to Eq. 8.

An important factor in PC is the selection of null marker sets for calculating the inflation factors for adjusting the

test statistics. The most conservative approach is to use all markers as the null, but this approach is unrealistic as it results in over-correction when the selection is strong and prevalent across the whole genome. Therefore, PC is best estimated from markers that have not undergone selection, although it is paradoxical given that such markers are unknown at this stage. As a compromise, we may assume that the allele frequency differences between first and last years are larger for markers under selection than drift. This assumption is reasonable for a modern breeding population that has undergone intensive selection. We first predicted the allele frequency change for each marker using the RALLY model and then identified the null marker set from markers that fall below various thresholds of allele frequency change. We tested the thresholds ranging from 0.05 to 0.50 at an increment of 0.01, in which the thresholds of 0.05 and 0.50 correspond to 40% and 99% of the total markers, respectively. An alternative, not explored here, would be to select a threshold from a target percentile of markers based on the changes in allele frequencies. Unfortunately, because the variance of allele frequency change, $\sigma^2_{\Delta p} = \frac{p(1-p)}{2N}$, is largest when the initial allele frequencies are intermediate (Falconer and Mackay 1996), a loss in RALLY's power to distinguish between weak selection signal and drift at those markers is unavoidable.

Detection limits of RALLY

We estimated the detection limits of RALLY using a simple example that is based on the simulated populations as described previously. We considered a QTL marker and 10 other proximal markers that are 1–10 cM away. The initial QTL frequencies were set to 1/32 to 16/32 with an increment of 1/32, and all possible marker-QTL haplotype frequencies were considered. We modeled selection on the QTL by increasing QTL initial frequency to the final frequency of 31/32 over 50 generations according to either a logistic or linear distribution. Consequently, the proximal markers experienced hitch-hiking effect due to the selection on QTL. Assuming an infinite population size, recombination is the sole factor that is responsible for the hitch-hiking effect, which allowed us to model the change in allele frequencies of the proximal markers. Non-recombinants are inherited at a probability of $1 - \theta$ and recombinants are inherited at a probability of θ. From this, we derived the expected allele frequencies for the proximal markers at each generation. Next, we randomly sampled 8 individuals per generation using a binomial distribution with the expected frequencies as the sampling probabilities. This step was repeated for 100 times for each tested marker-QTL haplotype frequency. A more detailed description of this is provided in Figure S1.

RALLY in two wheat panels

We first applied the RALLY approach in the Triticeae Genome (TG) panel (Bentley et al. 2014; Ladejobi et al. 2019) as a proof of concept. The TG panel has 344 winter wheat varieties from the UK, France and Germany that were released between 1948 and 2007 (Figure S2), which is ideal for analyzing selection over time in modern wheat breeding. We retained 333 varieties that were in common between the TG panel data derived from DArT markers (Bentley et al. 2014) and genotype-by-sequencing (GBS) markers (Ladejobi et al. 2019). The DArT marker data was only used in a later analysis for estimating multivariate selection parameters. From the initial 41,861 GBS markers, we removed 3,009 markers that are in high linkage disequilibrium (LD) ($r^2 > 0.2$) with markers from other chromosomes which left us with 38,852 markers. These markers were positioned according to the IWGSC RefSeq v1.0 genome assembly. Here, we applied a similar model to Eq. 2 with an additional fixed effect to account for the country of origin. We identified the year regression coefficients, applied the same level of PC as identified from the simulation to adjust the test statistics, and determined the significance at a Bonferroni-corrected threshold of 0.05 which is meant for multiple testing correction in m markers.

Next, we replicated the analysis in the WAGTAIL panel (Fradgley et al. 2019) to test RALLY performance in a different sampling panel of modern wheat varieties. The WAGTAIL panel has 403 winter wheat varieties of mostly UK origin that were released between 1916 and 2010. Of the 403 varieties, 283 originated from the UK, 51 from France, 34 from Germany and 35 from other countries including Australia, Belgium, Canada, Denmark, the Netherlands, Sweden, Switzerland, and the United States. There were 99 overlapping varieties between the TG and WAGTAIL panels. Since the WAGTAIL panel was genotyped using the wheat 90 k array (Wang et al. 2014) and did not immediately have physical map positions for direct comparison with the TG panel, we identified the physical map positions from the IWGSC RefSeq v1.0 annotation file. We retained 5,592 out of 26,015 markers that had matching chromosomes between the original WAGTAIL genetic map and the physical map. We also removed 319 markers that are in high LD ($r^2 > 0.2$) with markers from other chromosomes which left us with 5,273 markers. We applied Eq. 2 with an additional fixed country of origin effect to the WAGTAIL panel and computed the year regression coefficients with the same PC and multiple testing correction to the test significances.

Estimating local heritabilities from RALLY QSLs

We clustered the significant markers identified from RALLY into groups based on the extent of LD surrounding the

markers. Because genomic markers are not completely independent, some significant markers may be tagging the same QSLs. Starting with the most significant (focal) markers within each chromosome, we assigned markers that have $r^2 > 0.2$ with the focal marker to the same group. To avoid incorrectly mapped markers, we require the groups to have a minimum of 10 markers in the TG panel and 5 markers in the WAGTAIL panel due to lower marker density. As a trade-off, there may be bias against genomic regions with sparse marker density such as the D-genome. We repeated the process for the next significant marker that has not been assigned to any group until all significant markers had been assigned. Lastly, we merged all overlapping groups.

We estimated the local heritabilities (h_l^2) for each QSL in the TG panel using the genomic heritabilities partitioning method that was introduced by Schork (2001) and Visscher et al. (2007). QSLs with nonzero h_l^2 would support the hypothesis of selection over drift for the observed change in allele frequency. The TG panel includes 12 traits: flowering time (FT), lodging (LODG), yield (YLD), plant height (HT), grain protein content (PROT), winter kill (WK), awns (AWNS), specific weight (SPWT), total grain weight (TGW), ears per m^2 (EM2), tiller number (TILL) and maturity (MAT) (Bentley et al. 2014; Ladejobi et al. 2019). We were not able to estimate the h_l^2 in the WAGTAIL panel since we did not have multi-trait data for the WAGTAIL panel. For each trait and QSL combination, we estimated the h_l^2 from the following mixed model fitted using the *mmer* function from the "sommer" package (Covarrubias-Pazaran et al. 2016) in R (R Core Team 2021):

$$y = X\beta + Zg_a + Zg_b + \varepsilon \tag{11}$$

where the model terms are described as below.

y is a vector of phenotypic trait values for n varieties.

X is a design matrix for the fixed effects such that the first column is a vector of 1's for the mean, the second column is a vector of years, and the remaining $c-1$ columns are vectors of 0's and 1's indicating the country of origin in n varieties.

β is a vector of fixed effects including the mean effect, year effect and country of origin effects. There are $c-1$ country of origin effects, where c is the number of countries.

Z is an incidence matrix relating n varieties to observations y.

g_a is a vector of random genetic background effect due to relationship among varieties calculated from m_a markers which are in $a = 1, 2, ..., n_a$ marker group, and it follows a distribution of $N(0, \sigma^2_{g,a}K_a)$.

K_a is the additive genetic relationship matrix in which the elements $K_{a,jk} = \frac{\sum_i^{m_a}(x_{ij}-p_i)(x_{ik}-p_i)}{\sum_i^{m_a} 2p_i(1-p_i)}$; x_{ij} is the marker score for variety j in marker i, x_{ik} is the marker score for variety k in marker i, p_i is the allele frequency at marker i, and m_a is the total number of markers which are in group a.

g_b is a vector of random genetic background effect due to relationship among varieties calculated from m_b markers which are not in group a, and it follows a distribution of $N(0, \sigma^2_{g,b}K_b)$.

K_b is the additive genetic relationship matrix in which the elements $K_{b,jk} = \frac{\sum_i^{m_b}(x_{ij}-p_i)(x_{ik}-p_i)}{\sum_i^{m_b} 2p_i(1-p_i)}$; x_{ij} is the marker score for variety j in marker i, x_{ik} is the marker score for variety k in marker i, p_i is the allele frequency at marker i, and m_b is the total number of markers which are not in group a.

ε is the residual effect with a distribution of $N(0, \sigma^2_\varepsilon I)$ and I is the identity matrix.

After each model was fitted, we calculated the h_l^2 as $\frac{\sigma^2_{g,a}}{\sigma^2_{g,a}+\sigma^2_{g,b}+\sigma^2_\varepsilon}$. For any trait, we identified the nonzero h_l^2 groups ($h_l^2 > 0.001$) and refitted all n_a groups of markers in a new mixed model. The model is shown below with the similar terms as explained in Eq. 11.

$$y = X\beta + \sum_1^{n_a} Zg_a + Zg_b + \varepsilon \tag{12}$$

From Eq. 12, we estimated the new h_l^2 as $\frac{\sigma^2_{g,a}}{(\sum_1^{n_a}\sigma^2_{g,a})+\sigma^2_{g,b}+\sigma^2_\varepsilon}$ and used these as the final estimated h_l^2 for each trait and group combination.

Associating marker effects with alleles that are increasing over time

We estimated the marker allele effects for each trait in the TG panel using Ridge regression (RR) (Hoerl and Kennard 1970) and Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani 1996) approaches. For the RR approach, we used the *mixed.solve* function from the "rrBLUP" package (Endelman 2011) in R (R Core Team 2021). For the LASSO approach, we used the *cv.glmnet* function from the "glmnet" package (Friedman et al. 2010) in R (R Core Team 2021). In both approaches, we fitted a multiple linear regression model as shown below:

$$y = X\beta + ZMu + \varepsilon \tag{13}$$

where the model terms are described as below.

y is a vector of phenotypic trait values for n lines.

X is a design matrix for the fixed mean effect such that it is a vector of 1's.

β is a vector of fixed mean effect.

Z is an incidence matrix relating n varieties to observations y.

M is a $n \times m$ matrix of numerical marker genotypes coded as -1, 0 and 1 for homozygous first allele, heterozygous

and homozygous second allele, respectively. The number of markers is m.

u is a vector of marker allele effects. In RR, u is estimated from minimizing the loss function of $L_{RR}(u) = \|y - X\beta - ZMu\|^2 + \lambda\|u\|^2$ where $\lambda = \frac{\sigma_\epsilon^2}{\sigma_u^2}$ and $u \sim N(0, \sigma_u^2 I)$ (Endelman 2011). In LASSO, u is estimated from minimizing the loss function of $L_{LASSO}(u) = \|y - X\beta - ZMu\|^2 + \lambda\|u\|$ where λ is determined from the default tenfold cross-validations in *cv.glmnet* (Friedman et al. 2010). In addition, the multivariate LASSO model in "glmnet" was used to ensure that the effects for all traits are estimated from the same set of chosen markers.

ϵ is a vector of residual effects that follows a distribution of $N(0, \sigma_\epsilon^2 I)$.

For each trait j and marker k, we identified $\widetilde{d}_{j,k}$ which is the effect direction for the allele that is increasing in frequency over time, as follows: first, we determined $\widetilde{u}_{j,k}$ which is the direction of marker allele effect estimated from either RR or LASSO using the *sign* function in R (R Core Team 2021). This resulted in $\widetilde{u}_{j,k} = -1$ for negative effect, $\widetilde{u}_{j,k} = 0$ for no effect and $\widetilde{u}_{j,k} = 1$ for positive effect. Next, we determined $\widetilde{\beta}_{j,k}$ which is the direction of year regression coefficient estimated from RALLY. This resulted in $\widetilde{\beta}_{j,k} = -1$ for decreasing allele and $\widetilde{\beta}_{j,k} = 1$ for increasing allele. Because the marker alleles were coded similarly in the RALLY and marker BLUP models, we could calculate $\widetilde{d}_{j,k}$ as $\widetilde{u}_{j,k} \times \widetilde{\beta}_{j,k}$ directly. $\widetilde{d}_{j,k} = 1$ suggests that the increasing allele has a positive effect and $\widetilde{d}_{j,k} = -1$ suggests that the increasing allele has a negative effect. $\widetilde{d}_{j,k} = 0$ is only possible in LASSO due to variable selection, which simply implies that there is no effect. For any trait, an excess of either $\widetilde{d}_{j,k} = -1$ or $\widetilde{d}_{j,k} = 1$ across all markers indicates a possible directional selection.

For a pair of traits $j1$ and $j2$, we calculated $\widetilde{d}_{j1,j2,k} = \left[\widetilde{d}_{j1,k}, \widetilde{d}_{j2,k}\right]$ which is the pairwise effect direction for the increasing allele. $\widetilde{d}_{j1,j2,k} = [1, 1]$ implies that the increasing allele has positive effects on both traits, $\widetilde{d}_{j1,j2,k} = [1, -1]$ or $\widetilde{d}_{j1,j2,k} = [-1, 1]$ implies that the increasing allele has a positive and a negative effect on either trait, and $\widetilde{d}_{j1,j2,k} = [-1, -1]$ implies that the increasing allele has negative effects on both traits. By forming a contingency table from the counts of all four possible $\widetilde{d}_{j1,j2,k}$ combinations, we tested for selection-related interaction between the pairs of traits using a $\chi^2_{df=1}$ test in the results involving LASSO. We did not test the results involving RR because the marker effects are not independent.

Estimating multivariate selection parameters

We estimated the multivariate selection parameters in the TG panel using the multivariate breeder's equation of $\Delta Z = G\beta_{sel}$ (Lande and Arnold 1983). We obtained the selection response (ΔZ), genetic variance–covariance matrix (G) and phenotypic variance–covariance matrix (P) from the trait and marker data. Next, we solved the multivariate breeder's equation for the selection gradient β_{sel} and the equations of $S = P\beta_{sel}$ and $i = \frac{S}{\sqrt{diag(P)}}$ for the selection differential (S) and selection intensity (i) (Falconer and Mackay 1996). Lastly, we decomposed the multivariate selection parameters into direct and indirect partitions as a method to quantify the direct and indirect historical selection in the TG panel. As a check, we repeated the same process in a simulated example. Complete details on the methods on estimating multivariate selection parameters are provided in the Supplementary Methods.

Results

RALLY and GWAS in simulated populations

We tested RALLY's ability in identifying selection- or drift-induced marker allele frequency changes in simulated populations with (S) and without (U) selection (Fig. 1) by varying the degree of parametric control (PC). Briefly, PC combines genomic control (GC) (Devlin and Roeder 1999) and delta control (DC) (Gorroochurn et al. 2006) to correct for inflation in test statistics due to population structure. Details on the PC approach and simulations are described in the Materials and Methods section. Across all tested allele frequency change thresholds (t) for null marker set, setting $t > 0.11$ produced better control of test statistics (%sig-S < 1.867%, %sig-U < 0.109%, %sig = percentage of total markers that are significant) than without correction (%sig-S = 1.942%, %sig-U = 0.089%) (Fig. 2a, Table S1). At $t = 0.15$, we found little significance in the unselected population across all 100 simulations with some inevitable loss of significance in the selected population (%sig-S = 0.994%, %sig-U = 0.012%) (Table S1). This result suggests that PC at this threshold can reasonably separate out the true selection signals from drift in our simulation. To err on the cautious side, we used a higher threshold of $t = 0.20$ in the simulation, TG and WAGTAIL panels.

We evaluated the QSL/QTL mapping performances of RALLY and GWAS in the simulated populations with selection (Fig. 1) and found a higher mapping power in RALLY over GWAS (Fig. 2). Across the 100 simulations, we found that the individual significant markers are rarely shared between RALLY and GWAS (Fig. 2b), and even less likely to be found in GWAS but not RALLY (Fig. 2d). Most of the significant markers are found in RALLY but not GWAS (Fig. 2c). The low number of significances in GWAS is likely because the simulated QTLs have small effects and low heritabilities, which is common for quantitative traits.

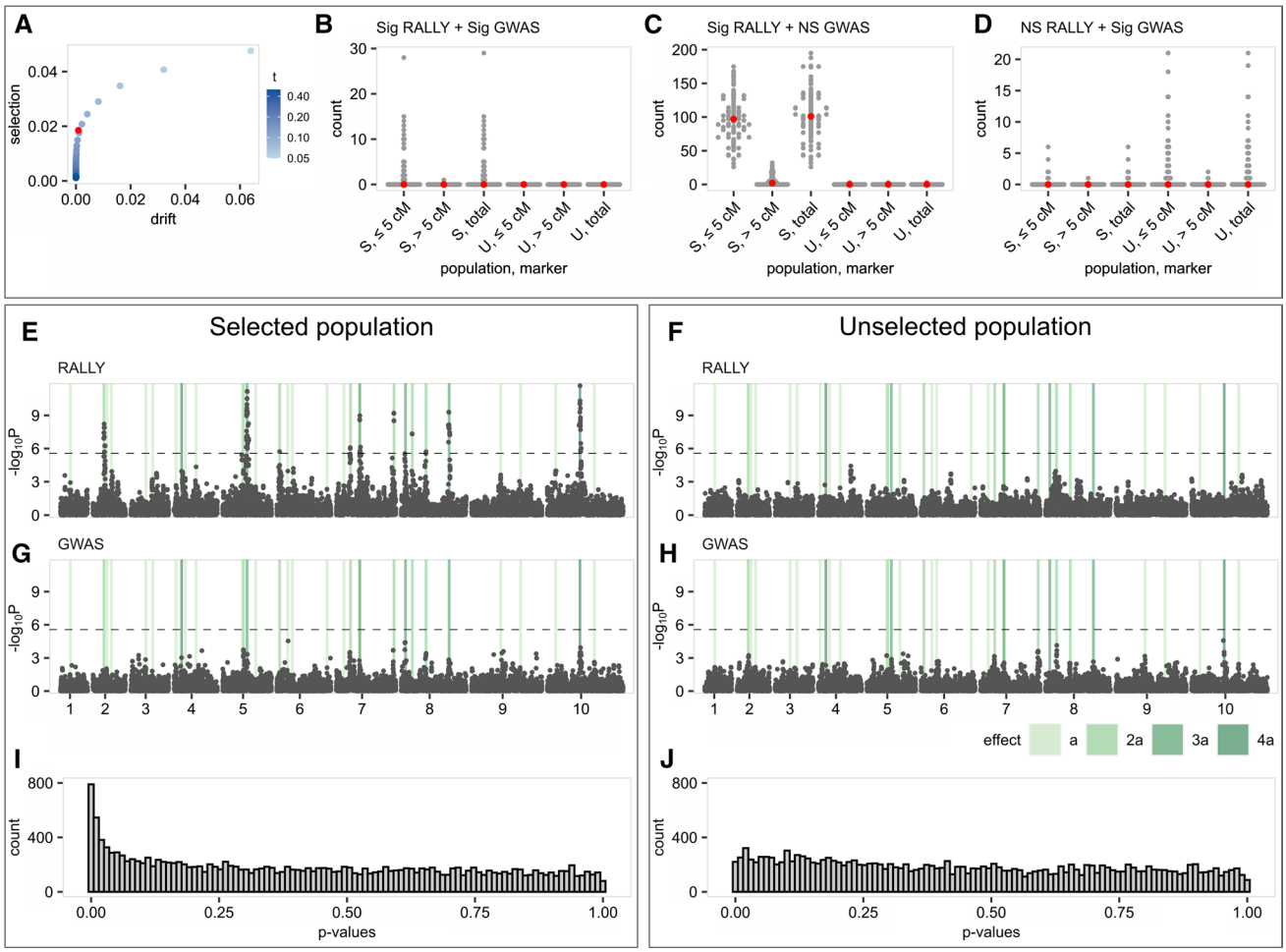

Fig. 2 RALLY and GWAS in simulated populations. Selected (S) and unselected (U) populations are simulated for 100 iterations and mapped for QSLs/QTLs using RALLY and GWAS. **a** Significant proportions of total markers identified from RALLY in S and U populations are calculated under various thresholds used in choosing null markers for delta control (DC) and genomic control (GC). The red point is estimated without DC and GC (uncorrected). Under the assumption that significant markers in U are due to drift alone and in S are due to both drift and selection, the *X*-axis is shown as the proportions in U while the *Y*-axis is shown as the differences in proportions between S and U. **b–d** Counts of significant markers identified from RALLY and GWAS are shown according to their distance from QTLs in both S and U populations. Medians are shown in red points. **e** Manhattan plot for RALLY in one simulated S population. QTLs are highlighted in vertical bars according to their effect sizes. **f** Manhattan plot for RALLY in one simulated U population. **g** Manhattan plot for GWAS in one simulated S population. **h** Manhattan plot for GWAS in one simulated U population. **i** Histogram of RALLY *p* values for the same simulated S population. **j** Histogram of RALLY *p* values in the same simulated U population

The heritabilities for the largest QTLs are approximately 0.030 and the smallest QTLs are approximately 0.002. An additional consequence of having low heritabilities is to reduce the fixation rate of QTL due to selection and prevent premature fixation of QTL in the simulated population. Premature fixation of QTL is more likely to increase the power of RALLY over GWAS, which may result in an unfair comparison between the two methods.

We repeated the RALLY and GWAS analyses in the unselected populations as a control for the same analyses in the selected populations (Fig. 2). On average across all 100 simulations, RALLY identifies 0.1 significant markers out of 19,000 total markers in the unselected population compared to 104.5 significant markers in the selected population. This result suggests that less than 0.1% of the significant markers in the selected population are likely caused by drift instead of selection. In the selected population, there are more significant markers (means of 99.4 versus 5.1) that are close to the QTLs (≤ 5 cM) than far (> 5 cM) (Fig. 2b–c). Assuming that all 32 QTLs are selected and all markers within 5 cM of the QTLs experience hitch-hiking effect, there should be a maximum of 3,200 significant markers in the selected population. However, the number of significant markers is much lower in reality because: (1) the selection force is proportional to the QTL effects (Figure S3), (2) the hitch-hiking effect depends on the initial marker-QTL haplotype

distribution (Figure S1), and (3) the hitch-hiking effect decreases as marker distance increases. On the other hand, GWAS performance remains similar between the selected and unselected populations (Fig. 2).

Detection limits of RALLY

Following from the previous simulation, we investigated the relationship between QTL under selection and its proximal markers and the results suggested a detection limit of approximately 5 cM (Figure S4). Here, we considered 10 markers that are evenly spaced between 1 and 10 cM away from a QTL and evaluated how these marker allele frequencies change as a result of increasing QTL frequency. Because the markers are linked to the QTL, we expect their frequencies to follow the QTL frequency in an inversely proportional way according to their genetic distances from the QTL. This process is commonly known as hitch-hiking, and it is an important consideration for RALLY because hitch-hiking markers are more likely to be genotyped than the true QTLs. As expected, our results suggest that the ability of RALLY in identifying significant hitch-hiking markers depends on the QTL-marker haplotypes, QTL initial frequency, and genetic distance between QTL and marker (Figure S4). With all factors considered, RALLY rarely detects significance beyond 5 cM although our previous results showed that some long-range significances may still be present (Fig. 2b). A possible explanation for this is when multiple QTLs co-localize into one major QTL haplotype, which may amplify the significances of surrounding markers.

RALLY in two wheat panels

We mapped 22 significant QSLs (Bonferroni corrected $p < 0.05$) across 14 chromosomes in the Triticeae Genome (TG) panel using RALLY (Table 1, Fig. 3, Figure S5, File S1). Because the distances between significant markers and true QTLs are unknown, we used a linkage disequilibrium (LD) measure of $r^2 > 0.2$ as a method to identify the genomic boundaries that the significant markers tag. This method resulted in QSL intervals ranging from 1.46 Mb to 774.73 Mb with a mean of 148.74 Mb. Given the large blocks of genomic regions and a previously approximated RALLY detection limit of 5 cM, many of the QSLs are likely to fall within low recombination regions. QSLs in high recombination regions are harder to map due to the lack of markers tagging the causative QTLs. Besides, sustained selection is more likely to be observed on multiple weakly favorable alleles in low than high recombination regions.

Table 1 Genomic positions of 22 RALLY QSLs in the TG panel. The IWGSC RefSeq v1.0 physical positions of the peaks and LD boundaries are shown, along with the $-\log_{10}$P scores associated with the peaks and overlapping QSLs/QTLs from RALLY in WAGTAIL panel and GWAS in TG panel (Ladejobi et al. 2019). The numbers in the overlapping QSLs/QTLs columns correspond to the QSL and QTL numbers in the first columns of Table S2 and S4

QSL	Chr	Position (bp)			$-\log_{10}$P	Overlapping QSLs/QTLs	
		Peak	Start	End		WAGTAIL	GWAS
1	1A	40,535,368	36,797,100	42,151,448	6.954	2	1
2	1A	138,028,803	105,888,613	395,485,807	6.409	2	0
3	1B	274,240,580	52,771,404	572,972,254	8.726	3	0
4	2A	19,049,803	502,328	36,134,675	7.012	5	7
5	2A	56,159,824	56,159,824	115,431,692	7.951	6	0
6	2B	230,348,363	10,888,962	785,614,875	10.621	9,10	10
7	3A	544,972,180	488,305,885	574,586,196	6.965	13	19
8	3B	20,035,143	19,086,765	31,366,713	6.011	0	0
9	3B	829,382,536	813,333,316	829,954,621	9.159	0	0
10	4A	690,425,855	507,739,498	695,893,542	7.189	14	22
11	4B	570,537,081	507,170,910	593,797,914	5.924	0	26,27
12	5A	59,666,472	31,088,127	449,788,941	7.816	16	28
13	5B	703,651,326	681,349,598	703,858,824	5.914	0	0
14	5D	69,776,655	43,408,942	233,674,405	6.148	0	0
15	6A	2,160,664	684,328	5,113,555	7.259	0	32
16	6A	89,355,276	61,817,777	545,399,189	6.538	18	33–37
17	6A	609,106,971	596,590,923	617,255,792	6.459	0	38
18	7A	612,599,663	610,209,166	612,599,663	6.263	0	0
19	7A	681,696,004	669,820,116	695,003,193	6.632	0	0
20	7B	3,693,110	3,366,069	4,826,131	6.438	0	0
21	7B	43,221,041	40,293,564	58,886,832	6.494	0	48
22	7B	704,838,082	698,229,993	707,941,517	7.941	0	49

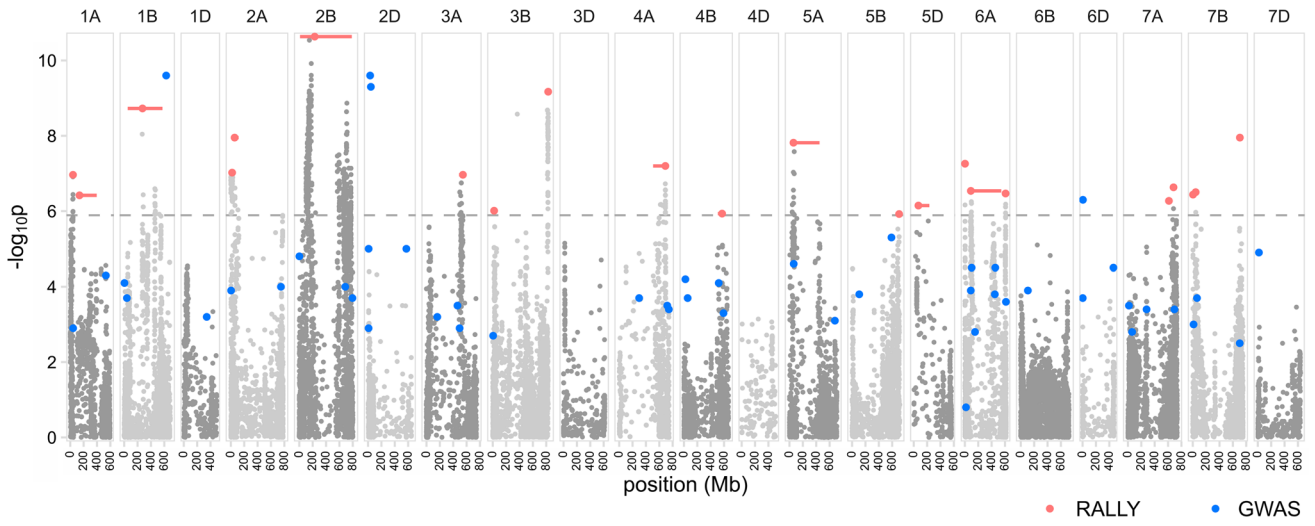

Fig. 3 Manhattan plot for RALLY results in the TG panel. RALLY peaks and their extents of LD are shown in red points and horizontal bars, respectively. GWAS peaks from Ladejobi et al. (2019) are shown in blue points. The dashed horizontal line represents the Bonferroni threshold of 0.05

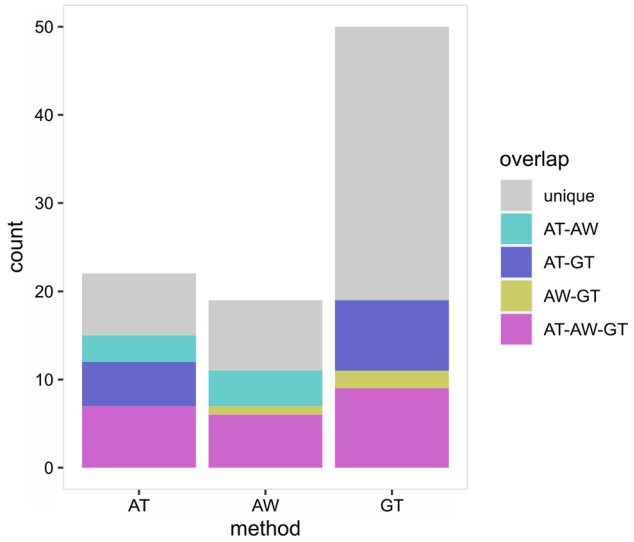

Fig. 4 QSL/QTL overlaps across different results. The number of overlapping QTLs among RALLY in TG panel (AT), RALLY in WAGTAIL panel (AW) and GWAS in TG panel (GT) are shown

Of the 22 QSLs, 12 co-localize with previously mapped QTLs using GWAS (Ladejobi et al. 2019) in the TG panel (Table 1, Figs. 3, 4, Table S2). The overlap between RALLY QSLs and GWAS QTLs is not statistically significant (two-tailed $P = 0.1488$) according to a randomization test (Figure S6), which is likely due to the large QSL intervals. QSLs/QTLs found in both RALLY and GWAS indicate that their effects are likely beneficial and have been selected during the breeding process. QSLs unique to RALLY suggest that their effects might be too small for GWAS to detect or the specific traits have not been analyzed for GWAS. QTLs unique to GWAS suggest that they are still segregating in the population, which could be due to various reasons like recent introduction into the breeding population and linkage drag.

A literature search showed that RALLY QSLs occur in both well-characterized and novel genomic regions in winter wheat (Table S3). The most significant RALLY QSL-6 mapped to a large region in chromosome 2B: 11–230 Mb, which includes *Ppd-B1* (Mohler et al. 2004) and multiple resistance loci of *Yr5*, *Yr7* and *YrSP* (Marchal et al. 2018). Another major QSL-16 mapped to a large region in chromosome 6A: 62–545 Mb, which contains *TaGW2* (Su et al. 2011) and the GA-responsive dwarfing genes of *Rht24* (Würschum et al. 2017) and *Rht25* (Mo et al. 2018). Interestingly, the durum wheat dwarfing gene *Rht14/16/18* resides in the same genomic region, although it remains to be tested whether it is allelic to *Rht24* (Haque et al. 2011). A recent EnvGWAS in winter wheat by Sharma et al. (2021) also mapped to the same genomic region (6A: 396 Mb) but without mention of any *Rht* candidate gene. On a broader scale, 16 RALLY QSLs co-localize with the recently identified meta-QTLs on yield and yield-related traits in wheat (Yang et al. 2021). 9 RALLY QSLs overlap with the QTLs identified from a Multi-parental Advanced Generation Inter-Cross (MAGIC) population of 16 diverse UK winter wheat varieties (Scott et al. 2021).

In addition, we found 11 RALLY QSLs that overlap with known alien and non-alien introgressions in wheat (Cheng et al. 2019). These include major introgressions like the 2A: 0–11 Mb from *Aegilops ventricosa* (Robert et al. 1999; Rhoné et al. 2007) and 2B: 90–749 Mb from *Triticum timopheevii* (Tsilo et al. 2008; Martynov

et al. 2018). These two introgressions were shown to segregate among the UK winter wheat varieties by Scott et al. (2021). Because alien introgressions tend to suppress recombination (Gill et al. 2011), they can be easily mapped using RALLY. Considering all overlaps in results between RALLY and the studies described thus far, we found 19 RALLY QSLs that can be traced to at least one study.

In the WAGTAIL panel, we mapped 19 significant QSLs across 13 chromosomes using RALLY (Table S4, Figure S7, File S1). We used the same approach as we did with the TG panel to identify the boundaries of these significant QSLs. With 99 varieties in common between the TG and WAGTAIL panels, we expect a high number of overlapping QSLs. 10 out of 19 QSLs in the WAGTAIL panel matched with 10 out of 22 QSLs in the TG panel (Fig. 4), which is approximately one-half overlap between them. The overlap between QSLs in TG and WAGTAIL panels is weakly significant (two-tailed $P = 0.0462$) according to a randomization test (Figure S6). Given that the TG panel was genotyped using GBS (Elshire et al. 2011) while the WAGTAIL panel was genotyped using the 90 k array (Wang et al. 2014), the genotyping and mapping quality of these two panels are likely different. This may partially explain why the results from the TG and WAGTAIL panels did not fully overlap. Another possible reason is that the distributions of countries of origin differ in the two panels in which the TG panel is more homogeneous than the WAGTAIL panel.

Local heritabilities in the RALLY QSLs

We calculated local heritabilities for the 22 RALLY QSLs as a support for possible selection over drift at these QSLs (Table 2, Fig. 5). We found that all 22 QSLs have nonzero local heritabilities for at least one trait. We tested for nonzero in the local heritabilities using a likelihood ratio test to compare between the mixed models with and without QSLs (Santantonio et al. 2019). However, most of the tests were non-significant due to low power (Table S5). The tests for QSLs collectively showed significance in 5 out of 12 traits, which comes at a cost of losing the test on individual QSL in exchange for a slightly higher power. In an extreme example with a total heritability of 0.379, QSL-16 at 6A: 89,355,276 is associated with 8 traits and found to co-localized with all other previously mentioned results. While it is possible that the underlying candidate genes *TaGW2* (Su et al. 2011), *Rht24* (Würschum et al. 2017) and *Rht25* (Mo et al. 2018) have pleiotropic effects that are beneficial for wheat breeding, we cannot exclude the possibility of additional genes that provide breeding

Table 2 Local heritabilities associated with 22 RALLY QSLs. Local heritabilities less than 0.001 are not shown

QSL	FT	LODG	YLD	HT	PROT	WK	AWNS	SPWT	TGW	EM2	TILL	MAT
1	–	–	–	0.054	0.048	0.051	–	–	–	0.005	–	–
2	–	–	0.007	0.068	0.105	–	–	–	0.003	0.043	–	–
3	0.084	–	0.006	–	0.014	–	–	–	–	0.002	0.000	0.060
4	–	–	–	–	–	–	0.001	–	–	–	0.016	–
5	–	0.009	–	0.002	–	–	–	–	–	–	–	–
6	–	–	0.009	–	–	0.003	–	–	0.009	–	–	0.037
7	–	–	–	–	–	–	0.014	–	0.055	–	0.009	–
8	0.005	–	–	–	–	–	–	0.020	–	–	–	–
9	–	–	–	–	0.002	–	0.004	–	–	–	0.006	–
10	0.010	0.000	–	–	–	0.002	0.003	0.008	0.090	–	0.012	0.019
11	–	0.040	0.036	–	0.012	0.002	0.011	0.005	0.005	–	–	–
12	–	–	0.003	–	–	0.005	0.011	0.075	–	0.003	–	–
13	–	–	0.006	–	–	0.049	–	–	0.002	–	–	–
14	0.069	–	0.019	–	0.018	0.000	–	–	–	–	–	0.006
15	0.011	–	–	–	–	0.004	0.018	–	0.005	0.014	–	–
16	–	0.086	0.045	0.127	0.071	–	–	0.013	0.012	0.021	–	0.006
17	0.072	–	0.001	–	0.001	0.016	–	–	0.002	–	0.013	0.069
18	–	–	0.006	–	–	–	–	–	0.003	–	–	–
19	–	0.020	0.005	–	–	–	0.011	–	0.028	0.001	–	–
20	–	–	0.008	0.089	0.004	0.000	–	–	–	0.065	–	–
21	–	0.039	–	0.041	–	–	–	0.009	0.025	–	–	–
22	0.010	–	0.001	0.021	–	–	–	–	0.016	–	0.013	–
Total	0.260	0.194	0.153	0.402	0.275	0.133	0.074	0.129	0.254	0.154	0.068	0.197
Others	0.284	0.123	0.160	0.223	0.205	0.220	0.583	0.156	0.050	0.216	0.051	0.179

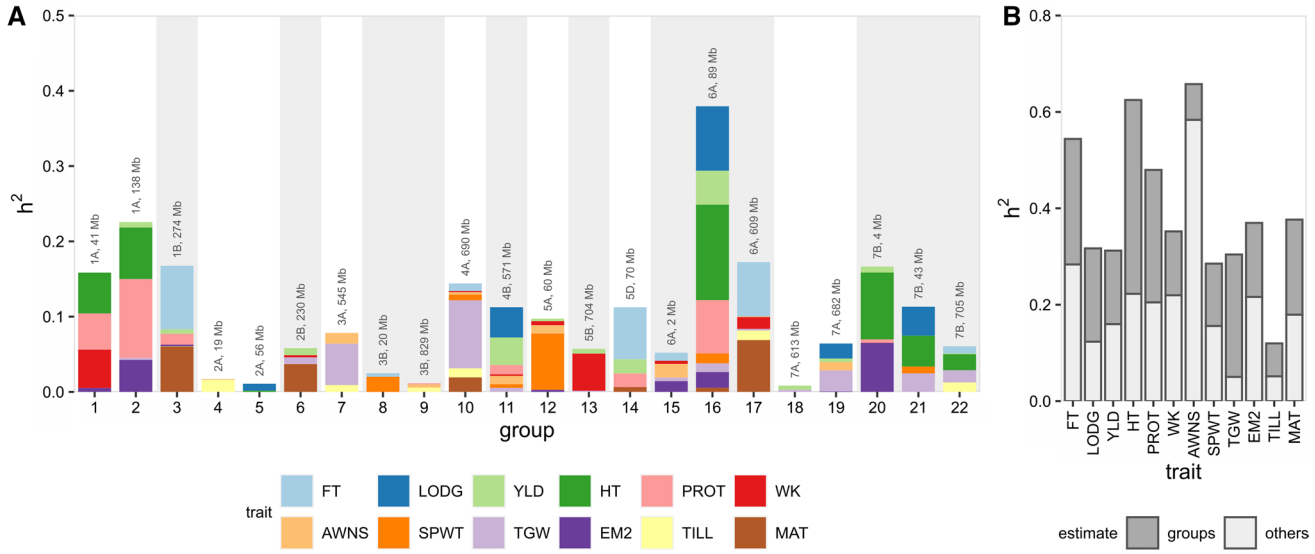

Fig. 5 Local heritabilities in RALLY QSL groups. **a** Local heritabilities for all 12 traits are shown as stacked bars for each RALLY QSL (defined in Table 1). **b** Local heritabilities from 22 RALLY groups are summed and compared against the heritabilities from other genomic markers

advantages in the same haplotype block. Nonetheless, given that QSL-16 has already played a major role in wheat breeding, it is unlikely to be useful for future breeding. The genomic region with the next largest total heritability of 0.226 is located in QSL-2 at 1A: 138,028,803. While no known gene has been mapped around QSL-2, results from our analysis and others (Cadalen et al. 1998, Griffiths et al. 2010, Tiwari et al. 2016) suggest that it may contain loci responsible for plant height and grain protein content.

Between the cumulative heritabilities explained by these 22 QSLs and the remaining genomic regions, HT and TGW are higher in the QSLs, AWNS is lower in the QSLs and the other 9 traits are about equal (Table 2, Fig. 5). This result highlights the narrow genetic diversity that is often seen in modern varieties (Reif et al. 2005) due to the repeated use of identical favorable haplotypes in wheat breeding. Fortunately, the remaining "unselected" genomic regions for important traits like yield, grain protein content and plant height are not fully devoid of heritabilities. There is still room for varietal improvement without the introduction of favorable exotic alleles in the short term, which suggests that it might be better to devote some of the resources in pre-breeding on these genomic regions instead. However, it is important to note that some selection may have already occurred in these genomic regions but failed to be detected as QSLs due to lack of power. For traits like TGW and TILL, breeders may need to look for alternative genetic resources to compensate for the lack of diversity.

Marker effects of alleles that are increasing over time

We evaluated the marker allele effects using the prediction models from Ridge regression (RR) and Least Absolute Shrinkage and Selection Operator (LASSO). Across all 12 traits, RR resulted in higher prediction accuracy than LASSO although the differences were comparable in some traits (Figure S8 and S9). Despite that, we retained the results from both approaches because the variable selection step in LASSO is important for a follow-up test involving trait pairs.

We examined the marker allele effect directions for increasing alleles and found excesses in one over another direction across each of the 12 traits (Fig. 6, Figure S10). We first partitioned the markers based on their RALLY significance into three groups: (1) markers with p values lower than the Bonferroni corrected threshold of 0.05, (2) markers with p values between 0.05 and the Bonferroni corrected threshold of 0.05, and (3) markers with p values higher than 0.05. The results from using either RR (Fig. 6) or LASSO (Figure S10) are similar although the differences across the significance groups in LASSO are less pronounced, i.e., there are more differences between group 1 and 2 in RR than LASSO results. This might be due to LASSO selected markers having weak but small,

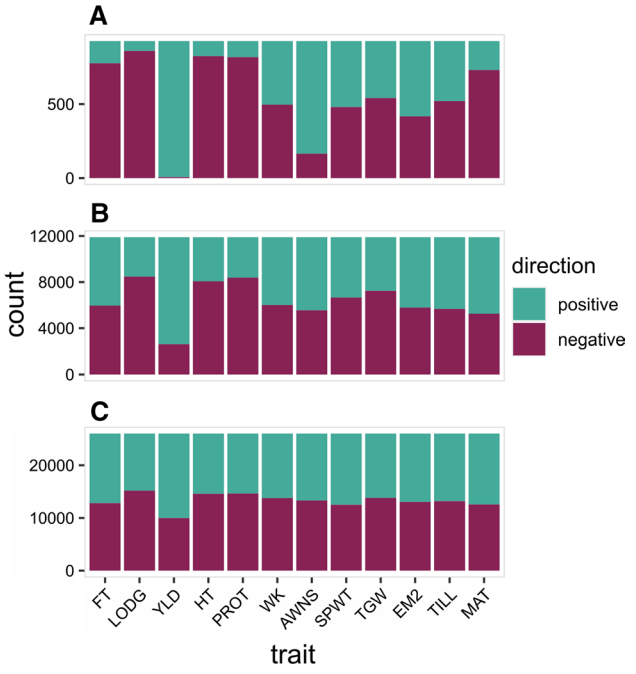

Fig. 6 Distributions of positive and negative RR effects in the increasing alleles. **a** Markers with RALLY *P* values of lower than the Bonferroni corrected threshold. **b** Markers with RALLY *P* values between 0.05 and the Bonferroni corrected threshold. **c** Markers with RALLY *P* values of higher than 0.05

Table 3 Counts of pairwise LASSO effects for alleles that are increasing over time. For each allele with increasing frequency over time, it is classified into pairs of traits for which the allele has an increasing effect on both traits (+/+), a decreasing effect on both traits (−/−) or antagonistic effects (+/− and −/+). The distribution for each pair of traits is tested with a $\chi^2_{df=1}$ contingency table where the significant threshold is set to $-\log_{10}p = 3.121$ (equivalent to a Bonferroni-corrected threshold of $p = 0.05$ for 66 possible pairs of traits). Only the significant trait pairs are shown here, and the full results are available in Table S7

Trait pair	+/+	+/−	−/+	−/−	$-\log_{10}p$
FT/YLD	270	43	145	59	4.435
FT/HT	112	201	22	182	9.352
FT/WK	135	178	135	69	6.326
FT/AWNS	139	174	140	64	6.959
FT/MAT	268	45	57	147	38.908
LODG/HT	75	56	59	327	20.083
YLD/PROT	74	341	62	40	17.486
YLD/MAT	280	135	45	57	4.687
HT/PROT	56	78	80	303	5.408
HT/WK	34	100	236	147	11.996
HT/AWNS	49	85	230	153	5.361
HT/MAT	102	32	223	160	3.474
PROT/SPWT	82	54	138	243	5.742
AWNS/MAT	145	134	180	58	7.317
SPWT/EM2	148	72	122	175	8.195
SPWT/MAT	119	101	206	91	3.269
TGW/EM2	85	131	185	116	5.962
EM2/MAT	149	121	176	71	3.643

non-significant changes in allele frequencies over time. Within the RR results, the excesses in effect directions are strongest in the significance group 1 and weakest in the significant group 3, which suggest that the excesses can be related to the favored direction of selection. The lack of excesses in significance group 3 implies that favorable and unfavorable alleles are still segregating about equally in the unselected genomic regions.

Across all 12 traits, the excesses agree with our expectation of traits that are important in wheat breeding. The most extreme example is yield (YLD) where both the RR and LASSO results show a near complete excess of positive effects in the increasing alleles in significance group 1. As shown previously by Mackay et al. (2011), the genetic gain in the UK winter wheat yield has been rising steadily over time. The next four traits with strong excesses are flowering time (FT), lodging score (LODG), plant height (HT) and grain protein content (PROT). FT, LODG and HT are favored for lower trait values, and thus the increasing alleles have excesses in negative effects. On the contrary, higher PROT is valuable for bread making quality, which is unfortunately going in the opposite direction due to a strong negative genetic correlation with yield (Scott et al. 2021). This result suggests that the selection for higher yield is a lot stronger than the selection for higher grain protein content. In the remaining traits, the excesses are smaller and less obvious given the variations seen from RR and LASSO results, which suggests that directional selection is likely weak for these traits.

By comparing the effect directions for increasing alleles in pairs of traits, we identified the priorities of traits under selection (Table 3, Table S6 and S7, Fig. 7). Taking YLD and PROT for example, there is a strong excess for alleles with positive YLD but negative PROT. This result reiterates the priority of YLD over PROT in wheat breeding. Between TGW and EM2, there is an excess for alleles with positive EM2 and negative TGW which suggests that more ears with lighter grains are preferred over fewer ears with heavier grains. In a different perspective, the results here also highlight the constraints imposed by genetic correlations across traits. For example, there is a small proportion of alleles with the same effect directions for YLD and PROT. These alleles could be used in breeding high YLD and PROT varieties, although it is still important to consider the possibility that these alleles could be unfavorably associated with other traits.

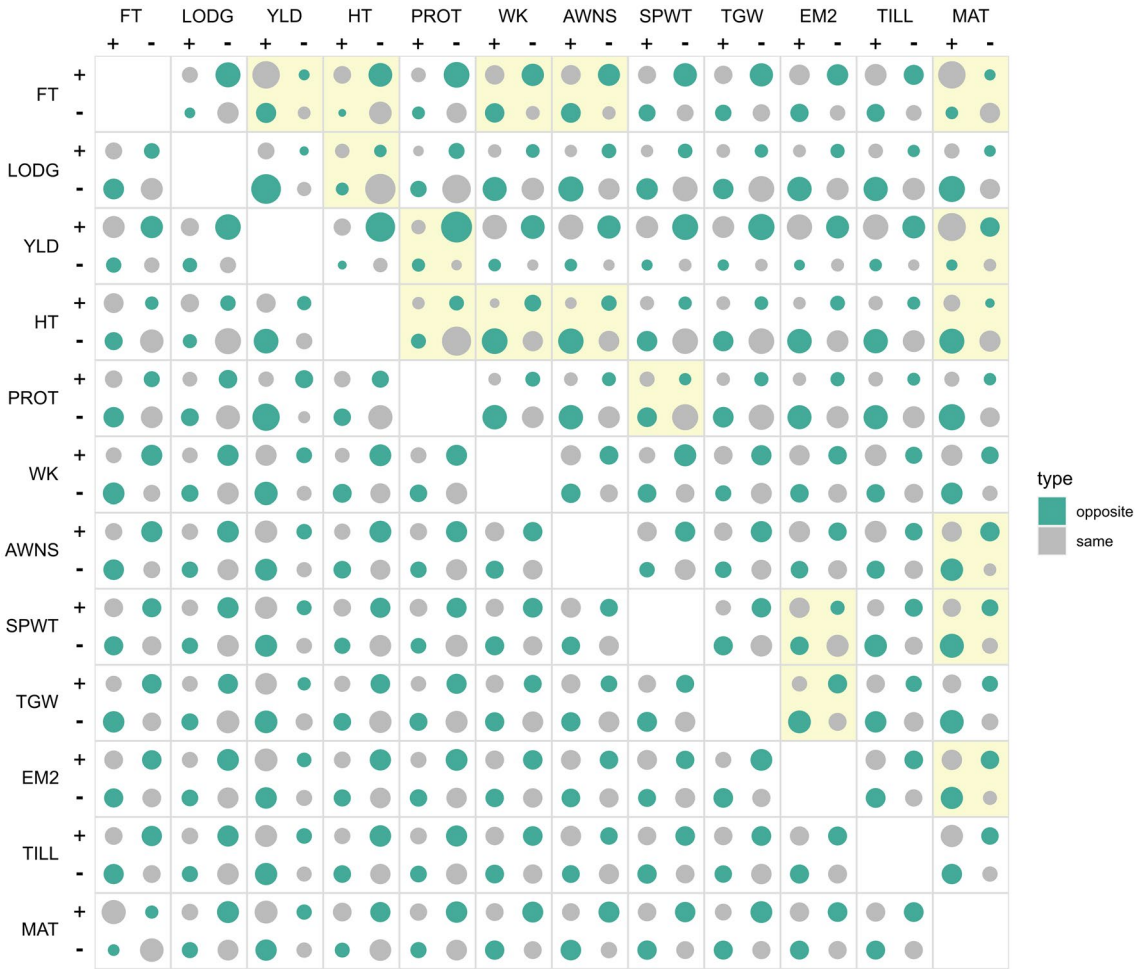

Fig. 7 Distribution of pairwise effects for alleles with increasing frequency over time. For each allele with increasing frequency over time, it is classified into pairs of traits for which the allele has an increasing effect on both traits (+/+), a decreasing effect on both traits (−/−) or antagonistic/opposite effects (+/- and −/+). The circle areas are scaled according to the marker counts. The bottom left triangle represents the RR effects and the top right triangle represents the LASSO effects. The distributions of the effect classes in each trait pair are tested using $\chi^2_{df=1}$ in the LASSO effects and significant results (Bonferroni-corrected threshold of $p=0.05$) are highlighted in yellow. No test is performed in the RR effects because the marker effects in RR are not independent

Multivariate selection parameters

In contrast to a genomic-centric approach that has been described thus far, the multivariate selection parameters may provide an alternative, trait-focused perspective on the historical selection of winter wheat represented by the TG panel. We found a strong misalignment between the selection response (ΔZ) and gradient (β_{sel}) where the directions of the vectors' elements are the opposite in 5 out of 12 traits (Table S8). If the selection parameters are estimated accurately, such divergence may imply an inefficient selection process. In addition, we partitioned the selection response (ΔZ), differential (S) and intensity (i) into direct and indirect components to quantify the amount of each selection parameter that is directly due to the available variation within a trait or indirectly due to the covariation with other traits. In an example with HT, we found positive direct effects in ΔZ, S and i, which contradicts the known selection on dwarfing genes like *Rht1*, *Rht2* and *Rht24* (Pearce et al. 2011; Würschum et al. 2017). Given the uncertainties in the multivariate selection parameters, we have provided the full results in the supplementary results and we advise to treat these estimates with caution.

Following the results, we investigated the possible causes of issues in estimating multivariate selection parameters using a simulated example with a single generation of selection involving three genetically correlated traits. First, we found that the genetic variances and covariances (G) estimated from mixed linear model were close to the true simulated values but with low precision (Table S9, Figure S11). Next, we computed the selection parameters (ΔZ, β_{sel}, S, i) from the simulation directly, true G and estimated G, which

are referred to as true, realized and estimated values, respectively. Given the imprecise estimates of *G*, we observed lower correlations between the estimated and true values than between the realized and true values (Table S9, Figure S12-S17). Despite using the true *G*, the realized values still failed to match the true values perfectly, which indicates that the deviations in realized ΔZ are carried over into the other selection parameters that are estimated downstream.

Discussion

Advantages and disadvantages of RALLY

RALLY has a major feature of being a trait-free method for mapping QSLs; however, this feature is a double-edged sword. For any population, RALLY involves only a single, relatively simple logistic regression analysis. In contrast, GWAS requires either multiple, simple mixed model analyses for each trait or a single, yet computationally intensive multi-trait analysis. Unlike any other trait-based mapping methods, the QSLs identified through RALLY are not restricted to only traits that are scored. While this makes RALLY a convenient method, the results do not inform us which traits the QSLs are associated with. In this regard, we will need to rely on other trait-based analyses like GWAS or genomic variance partitioning (Schork 2001; Visscher et al. 2007) to relate QSLs to traits. This additional step is not restricted to the same population as the QSL-trait information can be drawn from other studies such as GWAS on 47 traits in the wheat MAGIC diverse population (Scott et al. 2021). Therefore, RALLY can function as a replication of results from other studies.

As a kinship-free method, RALLY avoids any potential issues that may arise from the use of genomic relationship matrix (GRM) in mixed linear models. Recently, kinship estimates have been shown to be biased under complex population structure (Ochoa and Storey 2021), which can arise due to selection and migration of materials across breeders and countries. Besides, kinship estimates depend on the assumption that the alleles frequencies observed in the study population are representative of the reference or base population. For a population that has only experienced weak to no selection, the mean of genome-wide marker variance might be a reasonable approximation to the reference population. But, in populations under strong selection like those of modern crop varieties, the deviation between observed and true (reference) distribution of allele frequencies may not be trivial. Jiang et al. (2021) showed that the kinship estimates are biased when the observed distribution of allele frequencies fails to match the true distribution. In addition, a similar study on populations of modern wheat and barley varieties suggested that their kinship estimates may be biased due to long period of intensive selection (Sharma et al. 2021). However, the bias impacts on mapping power in GWAS and accuracy of variance component estimates remain to be evaluated.

Given that RALLY is designed specifically for mapping QSLs that have been selected over a time period, there may be limited utilities outside of its target scope. Our RALLY analyses model the change in allele frequency under a logistic distribution, which requires both genomic marker and year of variety release information. So, RALLY cannot be immediately applied to typical artificial mapping populations like bi-parental, nested association mapping (NAM) or MAGIC populations. However, we can extend the use of RALLY by conceptualizing it in its simplest form, which is a regression of marker allele on a variable of interest. For example, we can regress the marker allele on a continuous geographical origin variable such as latitudes and altitudes. The outcomes would directly define alleles that are relevant to local adaptation. Furthermore, the hybrid approach of parametric control (PC) is independent of RALLY and can be used in any genome-wide mapping analyses as a replacement for GRM and mixed linear model.

Selection history and future direction in winter wheat breeding

Given the largely incomplete overlap between RALLY and GWAS QSLs/QTLs in the TG panel, GWAS-specific QTLs may not have been directly useful in breeding. Several possible reasons include linkage drag between the QTLs, recent introduction of QTL alleles into the breeding pool, and ineffective selection at those QTLs. In the absence of genome editing to remove unfavorable alleles (Johnsson et al. 2019), linkage drag is unavoidable due the low probability of creating favorable recombinant haplotypes. New QTL alleles are hard to map under RALLY due to low power issue, but it can be improved by including more recent varieties. Ineffective selection is a direct consequence of the selection tendency toward low-hanging fruits. In an extreme example involving a cross between an elite variety and an exotic wild relative, selection is bound to reconstitute the elite genome because of the higher probabilities of favorable alleles in the elite over exotic genomes (Gorjanc et al. 2016). This phenomenon is observed in a large-scale crossing program involving groups of one exotic and two elite parents, in which the resulting lines lost approximately two-thirds of the expected exotic genome (Singh et al. 2021). In this regard, the approach of Origin Specific Genomic Selection (OSGS) (Yang et al. 2020) can be used to specifically target genomic regions outside of RALLY QSLs for selection.

The association between directions of allele frequency change and predicted marker effects provides us with an

overview of selection priorities (Figs. 6 and 7). High yield, short plants, early flowering, reduced lodging and reduced grain protein content are clearly preferred under directional selection. However, there is no obvious directional selection on spikes and grain-related traits, which suggests that there is no specific morphology that provides advantage in the breeding practice. The pairwise analysis further demonstrates the selection priorities and genetic correlations between traits. The results can be used to formulate a future breeding direction, for example, breeding for varieties with high yield and grain protein content by focusing on increasing the frequencies of the favorable alleles on both traits. In line with the global interest in shifting toward more sustainable agricultural practice (Hoad 2010), this approach can be extended to include traits relevant to sustainability and climate resilience to better guide the breeding direction.

Limited practical use of multivariate breeder's equation

As shown in the results involving the TG panel, the multivariate breeder's equation has limited practical use in estimating selection parameters (Table S8). An important component of the equation is the genetic variance–covariance matrix (G). The assumption that G is constant is likely violated because G should have been calculated from the base population (Walsh and Lynch 2018) rather than a population under selection over a time period. While this violation likely contributes to the poor estimates of the selection parameters, it is not the only source of issue. Variations across two tested genotyping methods (GBS and DArT) resulted in severely different selection parameters (Table S8) even when the G were similar across the two methods (Table S10).

Despite fulfilling the assumption of constant G and eliminating the genotyping discrepancy in our simulated example, additional issues remain in estimating selection parameters from the multivariate breeder's equation. We found that the poor estimation of multivariate selection parameters is caused by imprecise G estimated from mixed linear model. However, the estimation of multivariate selection parameters cannot be completely recovered even when the true G is used. This is probably because the multivariate breeder's equation can only capture the means but not the variances of the selection parameters (ΔZ, β_{sel}, S, i). Since the selection parameters are derived sequentially, repeated deviations from the means result in poor estimates of the selection parameters. This issue can be remedied by increasing the sample size, although there is a limit to the sample size due to practicality in breeding practice. Furthermore, the deviation is amplified across multiple generations of selection. Given the multi-layered issues with estimating selection parameters using the multivariate breeder's equation, it is best to limit its use to predict forward for a single generation as a rough guide to selection experiments involving crop varieties.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-022-04163-3.

Acknowledgements We thank Rajiv Sharma, Ian Dawson and David Marshall for helpful discussion on the work. We also thank the anonymous referees for valuable suggestions which have substantially improved this manuscript.

Author contributions statement CJY and IM conceived the work, performed the analyses and wrote the manuscript. OL provided data for the Triticeae Genome (TG) panel. RM and WP provided critical comments. All authors read, revised and approved the manuscript. No external funding was received for the work in this manuscript.

Data availability The GBS and phenotypic trait data for the TG panel (Ladejobi et al. 2019) were downloaded from https://doi.org/10.6084/m9.figshare.7350284. The TG panel DArT data (Bentley et al. 2014) and WAGTAIL panel data (Fradgley et al. 2019) were downloaded from https://www.niab.com/ research/agricultural-crop-research/resources. The IWGSC RefSeq v1.0 annotation file containing the physical map positions for the 90 k wheat array was downloaded from https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.0/iwgsc_refseqv1.0_Marker_mapping_summary_2017Mar13.zip. Computational analyses were performed using R version 4.1.0. All R scripts used in the analyses can be found at https://cjyang-work.github.io/RALLY.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Bentley AR, Scutari M, Gosman N, Faure S, Bedford F, Howell P, Cockram J, Rose GA, Barber T, Irigoyen J, Horsnell R, Pumfrey C, Winnie E, Schacht J, Beauchêne K, Praud S, Greenland A, Balding D, Mackay IJ (2014) Applying association mapping and genomic selection to the dissection of key traits in elite European wheat. Theor Appl Genet 127:2619–2633

Cadalen T, Sourdille P, Charmet G, Tixier MH, Gay G, Boeuf C, Bernard S, Leroy P, Bernard M (1998) Molecular markers linked to genes affecting plant height in wheat using a doubled-haploid population. Theor Appl Genet 96:933–940

Cheng H, Liu J, Wen J, Nie X, Xu L, Chen N, Li Z, Wang Q, Zheng Z, Li M, Cui L, Liu Z, Bian J, Wang Z, Xu S, Yang Q, Appels R, Han D, Song W, Sun Q, Jiang Y (2019) Frequent intra- and inter-species introgression shapes the landscape of genetic variation in bread wheat. Genome Biol 20:136

Coster A, Bastiaansen JWM, Calus MPL, van Arendonk JAM, Bovenhuis H (2010) Sensitivity of methods for estimating breeding values using genetic markers to the number of QTL and distribution of QTL variance. Genet Sel Evol 42:9

Covarrubias-Pazaran G (2016) Genome-assisted prediction of quantitative traits using the R package *sommer*. PLoS ONE 11:e0156744

Dadd T, Lewis CM, Weale ME (2010) Delta-centralization fails to control for population stratification in genetic association studies. Hum Hered 69:285–294

Decker JE, Vasco DA, McKay SD, McClure MC, Rolf MM, Kim J, Northcutt SL, Bauck S, Woodward BW, Schnabel RD, Taylor JF (2012) A novel analytical method, Birth Date Selection Mapping, detects response of the Angus (*Bos Taurus*) genome to selection on complex traits. BMC Genomics 13:606

Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 55:997–1004

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome. https://doi.org/10.3835/plantgenome2011.08.0024

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Prentice Hall, Essex

Fradgley N, Gardner KA, Cockram J, Elderfield J, Hickey JM, Howell P, Jackson R, Mackay IJ (2019) A large-scale pedigree resource of wheat reveals evidence for adaptation and selection by breeders. PLoS Biol 17:e3000071

Friedman J, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. J Stat Softw 33:1–22

Garin V, Wimmer V, Borchardt D, Malosetti M, van Eeuwijk F (2021) The influence of QTL allelic diversity on QTL detection in multi-parent populations: a simulation study in sugar beet. BMC Genom Data 22:4

Gaynor RC, Gorjanc G, Hickey JM (2021) AlphaSimR: an R package for breeding program simulations. G3: Genes, Genomes, Genetics 11:jkaa017

Gibson G (2012) Rare and common variants: twenty arguments. Nat Rev Genet 13:135–145

Gill BS, Friebe BR, White FF (2011) Alien introgressions represent a rich source of genes for crop improvement. Proc Natl Acad Sci USA 108:7657–7658

Gorjanc G, Jenko J, Hearne SJ, Hickey JM (2016) Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. BMC Genom 17:30

Gorroochurn P, Heiman GA, Hodge SE, Greenberg DA (2006) Centralizing the non-central chi-square: a new method to correct for population stratification in genetic case-control association studies. Genet Epidemiol 30:277–289

Gorroochurn P, Hodge SE, Heiman GA, Greenberg DA (2007) A unified approach for quantifying, testing and correcting population stratification in case-control association studies. Hum Hered 64:149–159

Gorroochurn P, Hodge SE, Heiman GA, Greenberg DA (2011) An improved delta-centralization method for population stratification. Hum Hered 71:180–185

Griffiths S, Simmonds J, Leverington M, Wang Y, Fish L, Sayers L, Alibert L, Orford S, Wingen L, Herry L, Faure S, Laurie D, Bilham L, Snape J (2010) Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm. Mol Breed 29:159–171

Haque MA, Martinek P, Watanabe N, Kuboyama T (2011) Genetic mapping of gibberellic acid-sensitive genes for semi-dwarfism in durum wheat. Cereal Res Commun 39:171–178

Hickey JM, Chiurugwi T, Mackay I, Powell W, Selection IG, in CGIAR Breeding Programs Workshop Participants, (2017) Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. Nat Genet 49:1297–1303

Hinrichs AL, Larkin EK, Suarez BK (2009) Population stratification and patterns of linkage disequilibrium. Genet Epidemiol 33:S88–S92

Hoad SP (2010) Evaluation of new varieties for sustainable cereal production in Europe. Farmers Club Charitable Trust. https://pure.sruc.ac.uk/en/publications/e0ac4808-c755-4f70-a224-4bc536106813. Accessed 25 November 2021.

Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12:55–67

Jiang W, Zhang X, Li S, Song S, Zhao H (2021) Correcting statistical bias in correlation-based kinship estimators. Biorxiv. https://doi.org/10.1101/2021.01.13.426515

Johnsson M (2018) Integrating selection mapping with genetic mapping and functional genomics. Front Genet 9:603

Johnsson M, Gaynor RC, Jenko J, Gorjanc G, de Koning DJ, Hickey JM (2019) Removal of alleles by genome editing (RAGE) against deleterious load. Genet Sel Evol 51:14

Ladejobi O, Mackay IJ, Poland J, Praud S, Hibberd JM, Bentley AR (2019) Reference genome anchoring of high-density markers for association mapping and genomic prediction in European winter wheat. Front Plant Sci 10:1278

Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. Evolution 37:1210–1226

Li J, Chen G-B, Rasheed A, Li D, Sonder K, Espinosas CZ, Wang JK, Costich DE, Schnable PS, Hearne SJ, Li H (2020) Identifying loci with breeding potential across temperate and tropical adaptation via EigenGWAS and EnvGWAS. Mol Ecol 28:3544–3560

Looseley ME, Ramsay L, Bull H, Swanston JS, Shaw PD, Macaulay M, Booth A, Russell JR, Waugh R, Thomas WTB (2020) Association mapping of malting quality traits in UK spring and winter barley cultivar collections. Theor Appl Genet 133:2567–2582

Mackay I, Powell W (2007) Methods for linkage disequilibrium mapping in crops. Trends Plant Sci 12:57–63

Mackay I, Horwell A, Garner J, White J, McKee J, Philpott H (2011) Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time. Theor Appl Genet 122:225–238

Marchal C, Zhang J, Zhang P, Fenwick P, Steuernagel B, Adamski NM, Boyd L, McIntosh R, Wulff BBH, Berry S, Lagudah E, Uauy C (2018) BED-domain-containing immune receptors confer diverse resistance spectra to yellow rust. Nat Plants 4:662–668

Martynov SP, Dobrotvorskaya TV, Krupnov VA (2018) Analysis of the distribution of *Triticum timopheevii* Zhuk. Genetic material in common wheat varieties (*Triticum aestivum* L.). Russ J Genet 54:166–175

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

Mo Y, Vanzetti LS, Hale I, Spagnolo EJ, Guidobaldi F, Al-Oboudi J, Odle N, Pearce S, Helguera M, Dubcovsky J (2018) Identification and characterization of *Rht25*, a locus on chromosome arm 6AS affecting wheat plant height, heading time, and spike development. Theor Appl Genet 131:2021–2035

Mohler V, Lukman R, Ortiz-Islas S, William M, Worland AJ, van Beem J, Wenzel G (2004) Genetic and physical mapping of photoperiod insensitive gene *Ppd-B1* in common wheat. Euphytica 138:33–40

Ochoa A, Storey JD (2021) Estimating F_{ST} and kinship for arbitrary population structures. PLoS Genet 17:e1009241

Pearce S, Saville R, Vaughan SP, Chandler PM, Wilhelm EP, Sparks CA, Al-Kaff N, Korolev A, Boulton MI, Phillips AL, Hedden P, Nicholson P, Thomas SG (2011) Molecular characterization of *Rht-1* dwarfing genes in hexaploid wheat. Plant Physiol 157:1820–1831

Prentice RL, Pyke R (1979) Logistic disease incidence models and case-control studies. Biometrika 66:403–411

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Reif JC, Zhang P, Dreisigacker S, Warburton ML, van Ginkel M, Hoisington D, Bohn M, Melchinger AE (2005) Wheat genetic diversity trends during domestication and breeding. Theor Appl Genet 110:859–864

Rhoné B, Raquin AL, Goldringer I (2007) Strong linkage disequilibrium near the selected *Yr17* resistance gene in a wheat experimental population. Theor Appl Genet 114:787–802

Robert O, Abelard C, Dedryver F (1999) Identification of molecular markers for the detection of the yellow rust resistance gene *Yr17* in wheat. Mol Breed 5:167–175

Rosyara UR, de Jong WS, Douches DS, Endelman JB (2016) Software for genome-wide association studies in autopolyploids and its application to potato. Plant Genome 9:1–10

Rowan TN, Durbin HJ, Seabury CM, Schnabel RD, Decker JE (2021) Powerful detection of polygenic selection and evidence of environmental adaptation in US beef cattle. PLoS Genet 17:e1009652

Santantonio N, Jannink J-L, Sorrells M (2019) A low resolution epistasis mapping approach to identify chromosome arm interactions in allohexaploid wheat. G3: Genes Genomes, Genetics 9:675–684

Schork NJ (2001) Genome partitioning and whole-genome analysis. Adv Genet 42:299–322

Scott MF, Fradgley N, Bentley AR, Brabbs T, Corke F, Gardner KA, Horsnell R, Howell P, Ladejobi O, Mackay IJ, Mott R, Cockram J (2021) Limited haplotype diversity underlies polygenic trait architecture across 70 years of wheat breeding. Genome Biol 22:137

Sharma R, Cockram J, Gardner KA, Russell J, Ramsay L, Thomas WTB, O'Sullivan DM, Powell W, Mackay IJ (2021) Trends of genetic changes uncovered by Env- and Eigen-GWAS in wheat and barley. Theor Appl Genet 135:667–678

Shorinola O, Simmonds J, Wingen LU, Uauy C (2022) Trend, population structure and trait mapping from 15 years of national varietal trials of UK winter wheat. G3: Genes, Genomes, Genetics 12:jkab415

Singh S, Jighly A, Sehgal D, Burgueño J, Joukhadar R, Singh SK, Sharma A, Vikram P, Sansaloni CP, Govindan V, Bhavani S, Randhawa M, Solis-Moya E, Singh S, Pardo N, Arif MAR, Laghari KA, Basandrai D, Shokat S, Chaudhary HK, Saeed NA, Basandrai AK, Ledesma-Ramírez L, Sohu VS, Imtiaz M, Sial MA, Wenzl P, Singh GP, Bains NS (2021) Direct introgression of untapped diversity into elite wheat lines. Nat Food 2:819–827

Su Z, Hao C, Wang L, Dong Y, Zhang X (2011) Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 122:211–223

Tadesse W, Sanchez-Garcia M, Assefa SG, Amri A, Bishaw Z, Ogbonnaya FC, Baum M (2019) Genetic gains in wheat breeding and its role in feeding the world. Crop Breed Genet Genom 1:e190005

Tibshirani R (1996) Regression shrinkage and selection via the lasso. J R Stat Soc 58:267–288

Tiwari C, Wallwork H, Arun B, Mishra VK, Velu G, Stangoulis J, Kumar U, Joshi AK (2016) Molecular mapping of quantitative trait loci for zinc, iron and protein content in the grains of hexaploidy wheat. Euphytica 207:563–570

Tsilo TJ, Jin Y, Anderson JA (2008) Diagnostic microsatellite markers for the detection of stem rust resistance gene *Sr36* in diverse genetic backgrounds of wheat. Crop Sci 48:253–261

van der Berg S, Vandenplas J, van Eeuwijk FA, Lopes MS, Veerkamp RF (2019) SIgnificance testing and genomic inflation factor using high-density genotypes or whole-genome sequencing data. J Anim Breed Genet 136:418–429

Visscher PM, Macgregor S, Benyamin B, Zhu G, Gordon S, Medland S, Hill WG, Hottenga J-J, Willemsen G, Boomsma DI, Liu Y-Z, Deng H-W, Montgomery GW, Martin NG (2007) Genome partitioning of genetic variation for height from 11,214 sibling pairs. Am J Hum Genet 81:1104–1110

Voss-Fels KP, Cooper M, Hayes BJ (2018) Accelerating crop genetic gains with genomic selection. Theor Appl Genet 132:669–686

Walsh B, Lynch M (2018) Evolution and selection of quantitative traits. Oxford University Press, Oxford

Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L, Mastrangelo AM, Whan A, Stephen S, Barker G, Wieseke R, Plieske J, IWGSC, Lillemo M, Mather D, Appels R, Dolferus R, Brown-Guedira G, Korol A, Akhunova AR, Feuillet C, Salse J, Morgante M, Pozniak C, Luo M-C, Dvorak J, Morrell M, Dubcovsky J, Ganal M, Tuberosa R, Lawley C, Mikoulitch I, Cavanagh C, Edwards KJ, Hayden M, Akhunov E (2014) Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. Plant Biotechnol J 12: 787–796

Würschum T, Langer SM, Longin CFH, Tucker MR, Leiser WL (2017) A modern Green Revolution gene for reduced height in wheat. Plant J 92:892–903

Yang CJ, Sharma R, Gorjanc G, Hearne S, Powell W, Mackay I (2020) Origin specific genomic selection: a simple process to optimize the favorable contribution of parents to progeny. G3: Genes Genomes, Genetics 10:2445–2455

Yang Y, Aduragbemi A, Wei D, Chai Y, Zheng J, Qiao P, Cui C, Lu S, Chen L, Hu Y-G (2021) Large-scale integration of meta-QTL and genome-wide association study discovers the genomic regions and candidate genes for yield and yield-related traits in bread wheat. Theor Appl Genet 134:3083–3109

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208