Plant Biotechnology Journal Open Access

Plant Biotechnology Journal (2024) **22**, pp. 2788–2807 doi: 10.1111/pbi.14405

Review Article

Epistasis and pleiotropy-induced variation for plant breeding

Sangam L. Dwivedi[1], Pat Heslop-Harrison[2,3], Junrey Amas[4], Rodomiro Ortiz[5] and David Edwards[4,*]

[1]*Independent Researcher, Hyderabad, India*
[2]*Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China*
[3]*Department of Genetics and Genome Biology, Institute for Environmental Futures, University of Leicester, Leicester, UK*
[4]*Centre for Applied Bioinformatics, School of Biological Sciences, University of Western Australia, Perth, WA, Australia*
[5]*Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden*

Received 18 November 2023;
revised 7 May 2024;
accepted 24 May 2024.
**Correspondence* (Tel +61 8 6488 6000;
email dave.edwards@uwa.edu.au)

Keywords: genetic correlation, genomic selection and heterosis, multi-role pleiotropy genes, machine learning algorithms, trade-off.

Summary

Epistasis refers to nonallelic interaction between genes that cause bias in estimates of genetic parameters for a phenotype with interactions of two or more genes affecting the same trait. Partitioning of epistatic effects allows true estimation of the genetic parameters affecting phenotypes. Multigenic variation plays a central role in the evolution of complex characteristics, among which pleiotropy, where a single gene affects several phenotypic characters, has a large influence. While pleiotropic interactions provide functional specificity, they increase the challenge of gene discovery and functional analysis. Overcoming pleiotropy-based phenotypic trade-offs offers potential for assisting breeding for complex traits. Modelling higher order nonallelic epistatic interaction, pleiotropy and non-pleiotropy-induced variation, and genotype × environment interaction in genomic selection may provide new paths to increase the productivity and stress tolerance for next generation of crop cultivars. Advances in statistical models, software and algorithm developments, and genomic research have facilitated dissecting the nature and extent of pleiotropy and epistasis. We overview emerging approaches to exploit positive (and avoid negative) epistatic and pleiotropic interactions in a plant breeding context, including developing avenues of artificial intelligence, novel exploitation of large-scale genomics and phenomics data, and involvement of genes with minor effects to analyse epistatic interactions and pleiotropic quantitative trait loci, including missing heritability.

Epistasis and pleiotropy in the evolution of complex traits

Complex multigenic traits, those not following Mendelian inheritance, are often affected by intra- or inter-locus interactions (dominance or epistasis, respectively) and pleiotropy (gene interactions). They often exhibit large phenotypic variation and are highly influenced by genotype × environment interactions (GEI; Cooper *et al*., 2009), while gene identification and inheritance patterns do not explain transmitability (the missing heritability problem). These traits vary in number or locus effects; i.e. from a few to many, with small effects. Epistasis has been recognized as fundamentally important to unlock and exploit the structure and function of genetic pathways and evolutionary dynamics of complex genetic systems (Phillips, 2008).

Bateson (1909) first coined the term epistasis to describe the deviation between the prediction of segregation ratios based on the action of individual genes and the phenotypes of a segregating dihybrid population. Geneticists over the years expanded the term epistasis into the functional relationship between genes, genetic ordering of regulatory pathways, and quantitative differences of allele-specific effects (Phillips, 1998). Epistasis, in addition to additive and dominance models, is an important genetic component of the variation of phenotypic traits in natural and breeding populations. A classical dihybrid ratio assuming two different genes controls two different traits and if each gene locus had an independent effect on a single phenotype result in a typical 9 : 3 : 3 : 1 ratio in a F_2 population. A deviation from this well-established ratio results in a different phenotypic ratio when two or more genes contribute to the same phenotype. The most detected form of two gene epistasis interactions is 12 : 3 : 1 (dominant epistasis), 9 : 3 : 4 (recessive epistasis), 9 : 6 : 1 (duplicate gene with cumulative effect), 15 : 1 (duplicate dominant genes), 9 : 7 (duplicate recessive genes), and 13 : 3 (dominant and recessive interaction). The digenic epistasis interactions for multigenic traits are termed as additive × additive, additive × dominance, and dominance × dominance interactions for the inheritance of quantitative traits (Kearsey and Pooni, 1996; Mather and Jinks, 1982).

Epistatic interactions involving three or more loci contributing to complex traits (high-order epistasis) are challenging to detect, typically requiring very large mapping populations with robust phenotype data, with the traits themselves often being sensitive to very local environmental conditions. The evidence from model organisms, however, indicates that high-order genic interactions frequently influence genetics and contribute to the evolution of complex traits (Taylor and Ehrenreich, 2015). Epistasis interactions shape the genotype–phenotype map of every species, and beneficial variants are more likely to exhibit genetic interactions (Ang *et al*., 2023). The evidence suggests that epistasis significantly influences the genetics of complex traits and heterosis (Sang *et al*., 2022; Yu *et al*., 1997). The interaction of several minor-effect genes in hybrids could activate the transcription activators of epistatic genes, thereby resulting in a cascade of amplified heterosis (Sang *et al*., 2022).

The term pleiotropy, introduced for the first time by a German geneticist Ludwig Plate in 1910, refers to the phenomenon in which a single locus affects two or more distinct phenotypic traits. Pleiotropy has significant influence in development, physiological, and medicinal genetics and on molecular and evolutionary biology in all living organism (Paaby and Rockman, 2013; Stearns, 2010), which could be a gene pleiotropy, developmental pleiotropy, or selectional pleiotropy (Paaby and Rockman, 2013). Both forward- and reverse-genetic approaches may be used to unearth the pleiotropic nature of trait expression. There are conceptual and operational constraints in detecting and quantifying pleiotropy. A review of empirical data on pleiotropy shows an L-shaped distribution of the degree of pleiotropy (i.e. the number of traits affected), with most genes having low pleiotropy. In crop breeding, pleiotropy could either be antagonistic, which may constrain adaptation or synergistic, which may produce phenotypic novelties to favour adaptation (Zhang, 2023). The genetic correlation between traits remains inherent in pleiotropy, while in linkage, it is breakable. Both sources of correlations could be modified, the former by mutation or segregation modifiers, the latter by recombination (Paaby and Rockman, 2013). Linkage refers to the increased likelihood of two traits being inherited together if they are located relatively close to each other on the same chromosome. The available literature indicates wide occurrence of both epistasis and pleiotropy, though to varying degrees, in crops. All these interactions are difficult to detect, conventionally requiring large mapping populations. The combination of the better understanding of the nature of the interactions, dense genotype and marker data from multiple lineages, and developing analytical tools, is enabling selection to advance with strong genetic support.

Form, function and extent of epistasis and pleiotropy affecting trait variation

Epistasis involves interactions between multiple genes, a phenomenon contributing to the complexity of understanding and exploiting genetic variation in crop plants. Much quantitative variation (traits showing continuous variation) in crops involves epistatic or pleiotropic interactions, while distinct character states (Mendelian characters) often involve single genes (the mantra of 'one-gene, one-trait'). Additional quantitative variation may arise from, or involve, epigenetic and gene silencing mechanisms, so analysis must also account for these sources of variation.

Epistatic interactions, where the effect of one gene (genetic locus) on a trait is influenced by one or more other genes, can mask or enhance the expression of certain traits. Thus, epistasis leads to non-additive genetic effects, where the combined effect of two (or more) genes is greater or less than the sum of their individual effects. The interaction leads to 'missing heritability' when using a genome-wide association study (GWAS) approach, in which the fraction of trait variation accounted for by inheritance of various genes, is other than one (Zhang *et al*., 2015). This requires more complex analysis of regulatory networks (Kim *et al*., 2008; Zhou *et al*., 2022). Epistatic interactions can be either positive (synergistic) or negative (antagonistic). Positive epistasis occurs when the combined effect of genes is greater than expected, thereby resulting in a larger effect on trait variation. Negative epistasis occurs when the combined effect is less than expected, which leads to a smaller effect on trait variation.

In the case of pleiotropy, a single gene affects multiple traits with various phenotypic effects. The traits may be clearly related (e.g. biomass yield and heading date for a forage grass; Anhalt *et al*., 2009), or under study, such as multiple biotic and abiotic stress resistances from single genes (Wisser *et al*., 2011). Pleiotropy can also occur from the simultaneous expression of unrelated traits under the control of shared transcription factors. Because genes affecting multiple traits can have cascading effects on aspects of plant development, physiology, or metabolism, the extent of pleiotropy may vary because it depends on the gene and traits involved. Some genes may have a relatively narrow pleiotropic effect and there may be alternative pathways, so they affect only a few related traits, while others may have a broader pleiotropic effect, thereby influencing a wide range of agronomic traits.

Both epistasis and pleiotropy contribute to the complexity of trait variation in crop plants. Understanding these genetic phenomena, including the development of appropriate analytical tools, is crucial for plant breeders. Manipulation of both epistatic interactions and pleiotropic effects helps breeders better predict and control trait expression, leading to a more efficient crop selection strategy.

Epistasis bias in estimating genetic parameters and in response to selection

Epistasis bias affects the estimation of genetic parameters in plants due to the presence of gene interactions. These interactions are difficult to capture and are often ignored in genetic analysis. This leads to bias in predicting traits from genomic information and reduces the speed at which crop plants can be improved using molecular technology. Epistasis bias has been previously noticed, e.g. in the resistance of pepper (*Capsicum annuum* L.) to phytophthora stem blight (*Phytophthora capsici* L., Bartual *et al*., 1993). In a study of pod yield and related traits in peanut (*Arachis hypogaea* L.), epistasis affected the expression of eight out of 11 traits analysed across environments (Upadhyaya and Nigam, 1998). A follow-up study noted epistasis in the prediction of protein and oil content in peanut (Upadhyaya and Nigam, 1999). Bias due to linkage disequilibrium (LD) and epistasis was observed in a study of seed yield and 100-seed weight in common bean (*Phaseolus vulgaris* L.), which observed that epistasis was widespread, and bias was more frequent in inter-gene pool crosses than in intra-gene pool crosses (Borel *et al*., 2016). While epistasis bias is difficult to avoid due to the complexity of gene interactions and the large number of markers that have small effects, Bocianowski (2014) suggests that using all loci, not only the loci with significant main effects, for estimation of the epistatic effects may help capture this missing heritability. This was supported by extensive analysis of several barley (*Hordeum vulgare* L.) populations (Bocianowski, 2014).

As the ability to generate large quantities of both DNA marker and complex phenotype data continues to grow, together with more advanced methods for trait association, including pangenomics and machine learning (Bayer *et al*., 2021), breeders, and researchers will gain an increased capacity to assess epistatic effects and reduce epistasis bias in crop trait prediction.

Linkage and pleiotropy affect genetic correlations

Quantitative traits are often controlled by many genes with small effects. A significant part of this variation is attributed to the environment and its interaction with the genotype. The nature and magnitude of genetic correlation measures the strength and direction of trait association, which may arise from either pleiotropy or linkage disequilibrium (LD), confounded at the level of quantitative trait loci (QTL) or genes (Chen and Lübberstedt, 2010; Schulthess *et al*., 2017). Selection and mutation affect genetic correlations (Chantepie and Chevin, 2020; Chebib and Guillaume, 2021). Differentiating between the two competing phenomena remains a significant challenge in biology. Advances in genomics provide researchers the opportunity to differentiate intragenic linkage from true pleiotropy.

Disadvantageous genetic correlations may limit breeders' ability to tailor resource use efficient cultivars to produce high yields and desired produce quality. A genome wide study involving a rice (*Oryza sativa* L.) core collection phenotyped under long and short days and genotyped with 4.6 million single nucleotide polymorphisms (SNPs) revealed a positive correlation between heading date and grains per panicle, and negative correlation between grains per panicle and panicles per plant, as well as varying correlations among other traits. In total, 47 pleiotropic genes in 15 pleiotropic quantitative trait loci (pQTLs), 18 pleiotropic genes containing 37 pleiotropic SNPs in eight pQTLs, and 39 pairs of interactive genes from eight metabolic pathways contributed to phenotypic correlations, of which 27 pQTLs had r^2 of LD above 0.2. Thus, deploying pQTLs and interactive genes and associated SNPs may open new avenues to overcome disadvantageous correlations and to utilize the advantageous correlations through genomic-assisted breeding (Li *et al*., 2018a). Plant height and seed weight in domesticated barley remain positively correlated irrespective of growth type and habit. A recent study involving plant height and seed weight on a diverse barley panel and genome-wide SNPs demonstrated that 17 SNPs had a pleiotropic effect on both traits and genes with function in diverse traits related to plant growth and development (He *et al*., 2023a).

Differentiating the effect of close genetic linkage from pleiotropy is difficult. A universal, rapid, and precise breeding system integrating knowledge from genetic mapping and functional analyses, as evidenced in rice, may provide an efficient platform to bridge ever-increasing genomic knowledge and diverse improvement needs (Wei *et al*., 2021). A comprehensive and versatile genome database with functional needs has also been developed to navigate and dissect wheat (*Triticum aestivum* L.) germplasm resources for functional characterization of genes associated with complex traits (Chen *et al*., 2024).

Models, software, and algorithms to analyse quantitative trait loci (QTL) related to epistasis and pleiotropy

Discovering and understanding the causes of pleiotropy may improve our understanding of how a gene is involved in trait expression. A number of valuable models have been developed, although their complexity often means the algorithms are computationally demanding. Software, some overviewed here (Figure 1; Table 1), is often freely available. The null hypothesis in multivariate methods assumes that none of the traits are associated with a genetic variant, and deviations from the null hypothesis indicate the presence of pleiotropy affecting two or more traits. A novel likelihood ratio test (LRT) for pleiotropy is based on linear regression methods for quantitative traits. LRT closely follows a χ^2 distribution when only one trait is associated with a genetic variant. In a sequential test, the null hypothesis first tests a 'no association' model and, if significant, it is to be

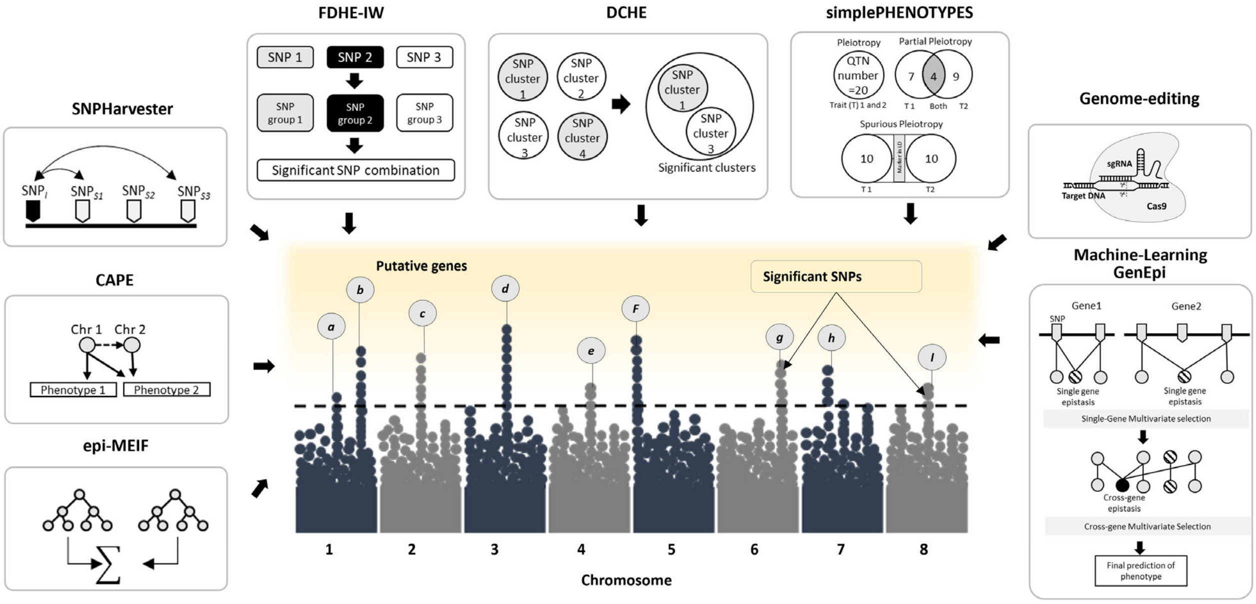

Figure 1 Overview of the methods for detecting and validating epistatic and pleiotropic interactions among genetic variants, including SNPs and genes.

Table 1 Summary of software package, access code, and important features for detecting pleiotropy and/or high-order epistasis interactions

Software package	Access	Remarks
epiMEIF	https://github.com/TAGC-NetworkBiology/epiMEIF	A high-performance flexible method for detecting higher order epistasis interactions
BADTrees	https://github.com/guyrt/WFUBMC/	Can determine high order interactions in the presence of marginal effects
FDHE-IW	http://www.mdpi.com/2073-4425/9/9/435/s1	Detects high-order epistasis based on an interaction weight method
SingleMI	https://github.com/sleeepyjack/singlemi/	Ultra-fast GPU based method for detecting higher order epistatic interactions
SNPHarvester	https://doi.org/10.1093/bioinformatics/btn652	Method to detect SNP–SNP interactions in GWA studies and harvests significant SNP groups that pass the statistical tests
RIL-StEp	https://doi.org/10.1093/g3journal/jkab130	Can detect epistasis in a pair of genetic variations of RILs based on the comparison of simple linear models
CAPE	https://cran.r-project.org/web/packages/cape/vignettes/cape.html	An R package for the combined analysis of epistasis and pleiotropy that infers directed interaction networks between genetic variants for predicting the influence of genetic perturbations on quantitative traits
qtl2pleio	https://doi.org/10.1534/g3.119.400098	Tests pleiotropy in the case of more than two alleles, incorporating polygenic random effects to account for population structure and uses a parametric bootstrap to determine statistical significance
CLIP	https://doi.org/10.5061/dryad.m0584	A method to distinguish whether correlated traits are due to pleiotropy and/or close linkage
simplePHENOTYPES	https://github.com/samuelbfernandes/simplePHENOTYPES	Used for the simulation of pleiotropy and linkage disequilibrium under additive, dominance, and epistatic models. The simulation currently takes a marker dataset as an input and then uses it for simulating multiple traits
PolarMorphism	https://github.com/UMCUGenetics/PolarMorphism	The software uses genome-wide summary statistics from GWAS of multiple related phenotypes to output statistics per SNP that describe its degree of 'sharedness' across the phenotypes of interest and its overall (multivariate) effect size, as well as *P*-values indicating significance for each statistic
Computational Framework for Statistical Epistasis Detection	https://github.com/EpistasisLab/epistasis_detection	Presents several efficient algorithms to compute linear regression models for statistical epistasis detection in Python and R

followed by a test whether only one or multiple traits are associated. If the test of only one associated trait rejects, then follow sequential testing of the null of 'j' associated traits, until the sequential test fails to reject the null hypothesis. Following this, one could assess the number of traits associated with a genetic variant, accounting for the correlation among traits. However, this test including other methods has a limited power if the allele is rare in the population. Furthermore, this test assumes that all traits are quantitatively inherited traits (Schaid *et al*., 2016). They further extended this approach to include traits that can be modelled by generalized linear models, such as binary traits, ordinal, or quantitative traits, or a mixture of these traits. It provides a testing framework to determine the number of traits associated with a genetic variant, as well as which traits, while accounting for correlations among the traits (Schaid *et al*., 2019). Sample size, number of traits, and trait correlations however may limit the detection of pleiotropy.

Complex traits are controlled by many genes with small effects. Detecting high-order epistasic interactions is a significant challenge. A novel method, epiMEIF, detects higher order epistatic interactions based on a group of SNPs potentially associated with the phenotype. The tree structure facilitates the identification of n-way interactions between the SNPs via extensive simulations. It can handle a wide range of complex genome-wide association studies. The effectiveness of epiMEIF is further validated in large simulation experiments reflecting a wide spectrum of complex models and with real datasets (Saha *et al*., 2022).

Ponte-Fernandez *et al*. (2022) compared a variety of statistical methods to detect high-order interactions in terms of runtime, detection power, and type I error rate. The exhaustive methods perform well across all experiments, although their computational cost may be prohibitive. The consistency of epistasis interactions, in the absence of marginal effects, is a drawback in non-exhaustive methods. Some of these (BADTrees, FDHE-IW, SingleMI, or SNPHarvester) are however effective in determining high-order interactions in the presence of marginal effects. For controlling false-positives, SNPHarvester, FDHE-IW, and DCHE perform well. No single epistasis detection method applies in all situations, so researchers should balance analysis between exhaustive methods if sufficient computational resources are available, and non-exhaustive methods when the analysis time is prohibitive.

Sakai *et al*. (2021) developed a simple and easy-to-perform method to detect epistasis. Recombinant inbred lines stepwise epistasis detection (RIL-StEp) uses genome-wide SNPs to detect epistasis, as evidenced for seed hull colour and chlorophyll content in rice, with a potential to identify epistatic interactions in other crops as well. The combined analysis of pleiotropy and epistasis (CAPE) integrates data across multiple quantitative traits to determine epistatic interactions. By doing this, it not only increases power to detect genetic interactions but also interprets these interactions across traits to identify interactions consistent across all datasets. An informative and interpretable interaction network, it explains how variants interact with each other to influence related traits. A platform-independent R package can be applied to data from both genetic screens and a variety of segregating populations (Tyler *et al*., 2021).

The R package qtl2pleio detects whether causal variation is due to pleiotropy or separate independent QTL. This software incorporates polygenetic random effects to account for population structure and uses parametric bootstrap to determine statistical significance (Boehm *et al*., 2019a,b). The close linkage versus pleiotropy (CLIP) test differentiates whether correlation in genetically correlated traits is due to pleiotropy or tight linkages. CLIP is a fast, simple, and powerful method that is based on the comparison of the square of the observed correlation between a combination of apparent effects at the marker level to the minimum value it can take under pleiotropic assumption. Simulation results show a higher CLIP detection power (68%) to detect close-linked QTLs than confidence interval (43%) or alpha risk (4%) tests (David *et al*., 2013).

The R/CRAN package simplePHENOTYPES simulates pleiotropy, partial pleiotropy, and spurious pleiotropy in a wide range of statistical-genetics models. Its ability to interface both marker data and quantitative genetics software and packages facilitates rigorous assessment of existing and emerging statistical GWAS and genomic prediction of breeding values for selection (GS) (Fernandes and Lipka, 2020). The authors of this package further modified the use of both multivariate and univariate GWAS to infer whether causal mutations underlying peak GWAS are pleiotropic as it narrows down the candidate loci with potential pleiotropic effects (Fernandes *et al*., 2021). Furthermore, the use of multi-trait multi-locus GWAS is recommended for traits controlled by both pleiotropic and non-pleiotropic loci, as evidenced for tocochromanol-related traits in maize (*Zea mays* L.) grain (Fernandes *et al*., 2022).

Another approach is taken by the algorithm implemented in PolarMorphism, which detects pleiotropic SNPs from GWAS summary statistics using two or more trait domains. It facilitates construction of a pleiotropy network, showing the extent to which traits share SNPs to gain insight into relationships between traits and trait domains, and contrast it with genetic correlation (von Berg *et al*., 2022).

Modelling epistasis and pleiotropy in association genetics

A systematic analysis of several hundreds of publicly available GWAS in humans suggests that pleiotropic genes are often involved in transcription regulation (Watanabe *et al*., 2019), which is a key biological function. In plants, flower characteristics show correlated variation because they are likely under the same gene control (Smith, 2016); i.e. a pleiotropic genetic architecture. Integrating biochemistry and dense metabolomics data gives further insights into pleiotropy affecting complex traits (Smith *et al*., 2022). Chung *et al*. (2014) also proposed a statistical approach facilitated by a computational algorithm that performs integrative analysis of datasets from multiple GWAS incorporating pleiotropy and functional annotations. They used GWAS' SNP-wise *p* as input from the summary statistics, thus overcoming the lack of information when phenotypic and genotypic data are unavailable, to identify pleiotropic effects. Mendelian randomization (Hu *et al*., 2022) is another approach that accounts for both pleiotropy and sample structure due to population stratification using GWAS summary statistics. It avoids detecting false positives and improves the statistical power for finding causal effects.

The relative contribution of pleiotropy and linkage to correlation among quantitative characteristics was estimated by Chebib and Guillaume (2021), who indicated that trait architecture affects finding causal loci in multi-trait association analysis, while pleiotropic variants underly multiple associations. According to their research, tight linkage between non-pleiotropic causal loci keeps high trait genetic correlations, thus leading to false positives in GWAS. The pleiotropic locus exploration and interpretation using optimal test (PLEIO) provides a summary-statistic-based framework that maps and interprets pleiotropic loci considering correlations among either binary or quantitative traits and their heritability (Lee *et al*., 2021). When having sample overlaps in association genetics, it will be possible to quantify the expected spurious correlations through a simple linear correction, adjusting the joint distribution of test statistics from the two GWAS (LeBlanc *et al*., 2018). Recently, Fernandes *et al*. (2021) demonstrated with publicly available maize and soybean databases that combining multivariate and univariate GWAS may be used to infer if associations result from true or spurious pleiotropy.

Epistasis and pleiotropy affect the adaptive value of fitness-related genes in plants as shown by Scarcelli *et al*. (2007) in the model species *Arabidopsis thaliana* L. (thale cress). Multi-locus models in GWAS revealed that epistasis has a main role for host plant resistance to blast in rice (Rosas *et al*., 2020), flowering time in soybean (*Glycine max* (L.) Merr.) (Kim *et al*., 2020) or grain morphology and yield traits in wheat (Fradgley *et al*., 2023; Malik *et al*., 2021, 2022) and affect seed yield in rapeseed (*Brassica napus* L.) (Luo *et al*., 2017a). Epistasis accounts partially for the missing or phantom heritability in GWAS (Slim *et al*., 2020). However, both the sheer magnitude of the search and related computational complexity remain the main challenges when dealing with epistasis in GWAS. There are several methods for finding epistasis in GWAS. For example, epi-GWAS (Slim *et al*., 2020), encompassing several methods, recognizes interactions between a target SNP and the rest of the genome. Likewise, epiSNP (Weeks *et al*., 2018) allows finding pairwise SNP interactions in GWAS in a realistic amount of time. Moreover, a LASSO-penalized-model search algorithm has proven to be effective for detecting epistasis in genome-wide association analysis (Zhou *et al*., 2014). Furthermore, GenEpi is a computational package that uncovers epistasis associated with phenotypes using a machine learning approach (Chang *et al*., 2020).

Detecting epistatic gene interactions using SNPs

Epistatic gene interactions are challenging to identify during plant breeding, not least due to complex environmental interactions between the traits and multiple genes involved. Hence, experimental design for phenotyping, appropriate marker systems for genotyping, statistical approaches for analysis, and biological validation are all topics of research. With plant breeding in the 21st century targeting yield improvements as low as 1%, and phenotypes related to quality and sustainability of production becoming increasingly important, it is essential to consider genetic selection of traits showing epistasis. Robustness with stable and predictable productivity over multiple years may be difficult to measure, requiring at the least trials in multiple locations, but is important, not least because of increasing climatic variation. However, with a complex polygenic trait or phenotype, detection of epistatic groups of genes (including in polyploid crops) from genotype information, typically SNPs from

PCR, array, genotyping by sequencing, or other approaches, is computationally intensive in a typical approach based on a GWAS with additional interaction terms between markers to detect epistatic effects. New approaches, including use of artificial intelligence, have huge prospects for discovering epistatic interactions and allowing selection of genes (Bhat *et al*., 2023). Compounding the multi-genic nature of the traits (not confined to epistasis), rare genetic variants may be the most important to identify for crop improvement, enabling co-adaptation to be selected (Dwivedi *et al*., 2023).

Targeted approaches to analyse epistasis based on knowledge of genes identified in a genome-wide scan (association or segregation studies) can be powerful. Not requiring an all-by-all analysis of genes, but testing for interactions between candidate genes and others (statistical or data mining) directly captures pairwise or higher order interactions between genetic markers. However, a favourable allele in one genetic background may be deleterious in another, and such studies require larger datasets to identify potential non-linear relationships. Some favourable alleles may show opposite effects in different backgrounds due to context dependence (Eguchi *et al*., 2019). There are several reasons why this background dependency may occur. Genetic interactions could be additive, synergistic, or antagonistic, thus resulting in different outcomes in different genetic contexts. Likewise, when positive epistasis occurs, the effect of an allele may be enhanced in the presence of certain alleles at other loci, while in case of negative epistasis, the effect may be suppressed, which leads to opposite effects of the same allele in different genetic backgrounds. Genetic diversity may also explain such an outcome because alleles that are beneficial in one population may have different effects in another population due to differences in genetic composition and evolutionary history. Moreover, regulatory elements and epigenic modifications affect the expression of genes, leading to differential gene expression and phenotypic results. Furthermore, environmental factors can modulate the effects of genetic variation because alleles that are advantageous in one environment may have contrasting effects in another environment; i.e. the GEI influencing phenotypic expression across distinct genetic backgrounds.

Detecting epistatic gene interactions in plant breeding populations is challenging, with a large number of genetic markers and complex environmental interactions, thus research outcomes may not reach significance. However, once epistatic interactions are identified, the underlying biological mechanisms and network analysis involving gene expression analysis, pathway analysis, or functional characterization (Borrill *et al*., 2019) of the interacting genes may suggest further selection targets.

Integrated networks of genes involved in pleiotropic and epistatic interactions

Both epistasis and pleiotropy have the potential to cause deviation from the Mendelian genetics of one gene-one phenotype paradigm and may limit progress towards a deeper understanding of biological systems. Both constitute ubiquitous and inherent properties of biomolecular networks. Understanding and harnessing epistasis and pleiotropy genetic variants at a systems level is therefore important to unlock the functioning of complex traits (Tyler *et al*., 2009). Biomass recalcitrance is a significant challenge to harness plant sugar for the biofuels industry. Using high-resolution GWAS data combined with co-expression, co-methylation, and SNP correlation networks in a multi-omic data layering approach, Weighill *et al*. (2018) identified new target genes involved in lignin biosynthesis in poplar (*Populus trichocarpa* Torr. & A. Grey ex. Hook). Poplars are of particular interest for bioenergy and feed production (Sangster *et al*., 2004; Tsarev *et al*., 2021).

SNP-based multi-phenotype associations (MPAs) in GWAS provide useful information about the impact of a gene on closely related traits. MPA decomposition is a new network-based approach that decomposes the outcomes of a multi-phenotype GWAS to unlock multi-phenotype signatures of genes. The decomposition involves the construction of a phenotype powerset space that maps genes into this new space. Clustering of genes in this powerset space groups genes based on their detailed MPA signatures. MPA decomposition tested in a large set of poplar genotypes (882 accessions) detects multiple different MPA and pleiotropic signatures within individual genes. It classifies and clusters genes based on these SNP–phenotype association topologies. Thus, they are helpful in interpreting large GWAS datasets that could aid in future synthetic biology efforts in designing and optimizing phenotypes of interest (Weighill *et al*., 2019).

Plants being sessile in nature use vast regulatory networks, involving epistasis between genes within and across networks to adapt in diverse environments. The accumulation of aliphatic glucosinolate can regulate plants fitness. An investigation involving a large network of TFs regulating aliphatic glucosinolate biosynthesis in *Arabidopsis* single and pairwise mutants uncovered extensive pairwise epistasis between TFs, regardless of subnetwork membership. The accumulation of metabolites shows antagonistic epistasis, indicating a buffering mechanism. Epistasis affecting enzymatic activity is highly conditional on the tissue and environment that shift between both antagonistic and synergistic forms. The epistasis shifts, however, depend on how the network's phenotype is quantified, suggesting epistasis may be a common feature of large regulatory networks influencing adaptive traits in plants (Li *et al*., 2020).

Correlations among traits and epistatic interactions impacting single traits pose a significant challenge in cultivar development. GWAS involving 809 soybean accessions and 84 trait data collected over two years from three environments detected 245 significant genetic loci, of which 95 interacted with other loci. Network analyses identified 51 traits linked to 115 associated loci reflecting phenotypic correlations. Twenty-three loci, including known *Dt1*, *E2*, *E1*, *Ln*, *Dt2*, *Fan*, and *Fap*, and 16, undefined associated loci, have pleiotropic effects on different traits (Fang *et al*., 2017). A rice GWAS involving 21 traits on 113 cultivars and 565 test cross hybrids identified 381 primary significant associated loci (SAL) and 1759 secondary SALs showing epistatic interaction with primary SALs, which could contribute to trait enhancement by pyramiding superior haplotypes with desirable epistatic alleles. Furthermore, incorporating SALs as covariates into a genomic selection model could enhance the prediction accuracy of the parental lines by incorporating epistatic effect SALs to predict trait values more accurately (He *et al*., 2023b).

Pleiotropic and non-pleiotropic QTL affecting multiple traits

Mapping QTL and detection of overlapping genomic regions is a powerful method to identify pleiotropy. High resolution mapping delineates pleiotropic QTL from tightly linked QTL (Knight *et al*., 2001), while joint mapping of QTL affecting multiple traits assesses whether genetic correlations arise because of pleiotropic QTL or unlinked QTL affecting different traits (Jiang and Zeng, 1995). Previous research indicated the existence of pleiotropic variation affecting multiple traits in animals and plants (Flint and Mackay, 2009; Hall *et al*., 2006). Pleiotropic variation in plants has been shown to impact genes with effects on plant growth and development, stress tolerance, and total biomass including their nutritional quality (Table 2). Rice among grain crops is the most extensively studied for detecting QTLs/genes with pleiotropic effect controlling complex traits. Traditionally early generation breeding populations (such as F_2:F_3 progenies), advanced backcross lines, recombinant inbred lines (RILs) derived from biparental crosses or using more complex genetic mating systems such as multiparent advanced generation intercross (MAGIC) or nested association mapping (NAM), doubled haploid lines (DHLs), genome wide association mapping panel (GWAMP), or near isogenic lines (NILs) are employed to identify both independent and pleiotropic QTLs/genes. All forms of mating designs including gene-based mutation ('forward genetics') and transgene were used to detect pleiotropic QTLs/genes in diverse crops (Table 2). GWAMP, RILs, DHLs detect major QTLs because of their limitation in collecting large-scale data sets, whereas use of gene-based mutation, transgene, or NILs provide more robust approach to detect pleiotropic effect.

Investigating the extent to which pleiotropy impacts phenotypes using GWAS summary statistics reanalyzed from previously published phenotyping data (metabolite, field, and expression phenotypes) and over 480 million significant quantitative trait nucleotides across the NAM population and Goodman diversity panel, for example in the maize genome, revealed 1.56%–32.3% of intervals with some degree of pleiotropy. A relationship between pleiotropy and various features (gene expression, chromatin accessibility, sequence conservation, etc.) reveals that pleiotropy of common alleles is not widespread in maize and is impacted by population structure and linkage disequilibrium (Khaipho-Burch *et al*., 2023). Meta-analysis involving an association panel (234 separate datasets on 304 lines) and large SNP data sets (107 751 in 2013 and 257 882 in 2020) captures only 35%–43% of potentially detectable loci controlling variation for traits, and limited evidence for pleiotropy, detected both known (i.e. dwarfing genes on root architecture) and new pleiotropic loci controlling phenotypic trade-offs in sorghum (*Sorghum bicolor* (L.) Moench.) (Mural *et al*., 2021). Both the reports indicate that a substantial proportion of detectable phenotypic variation is governed by non-pleiotropic loci.

Designing future crops to enhance productivity, nutritional quality, and abiotic stress adaptation of staple food crops

Domestication genes contributing to the pleiotropic variation for multiple traits

With the genetic bottleneck of domestication and strong selection for the characters that make a crop worthwhile and possible to grow, genes contributing to pleiotropic variation for multiple traits play a fundamental role in the selection of plants with traits that differ significantly from their wild ancestors. Mendel (1866) in his pea (*Pisum sativum* L.) experiments showed a now classic example of pleiotropy, with three characters, namely seed-coat colour, flower colour, and axillary pigmentation, being inherited together: plants show either a brown seed coat, violet flowers, and axial spots or a white seed coat, white flowers, and no axillary spots. Allelic variation in the *A* gene, a basic helix–loop–helix (bHLH) transcription factor, was shown to be involved (Hellens *et al*., 2010). A second gene, *A2*, a WD40 regulator, was also shown to be involved in the regulation of the anthocyanin pathway, and potentially epistatic dominance hiding the effect of the *A* gene. In general terms, because of potential effects on expression of multiple structural genes, transcription factors, and cis-regulatory elements are frequently detected as underlying genetic changes during domestication and have been considered as targets for genome editing (Swinnen *et al*., 2016).

With segregation analysis and crossing, it can be difficult to distinguish genetic linkage (with genes for multiple characters in adjacent chromosomal loci and minimal recombination) from pleiotropy (discussed by Vaughan *et al*., 2007) where the same gene controls multiple characters. Even with genome assemblies, several genes may be identified in a GWAS or QTL analysis, and a signature of domestication detected as a selective sweep leading to low diversity in a genomic region, associated with domestication and breeding. Furthermore, contrasting selection pressures during domestication may lead to pleiotropic genes with chimeric positive and purifying selection signals (Tao *et al*., 2017). Systematic knockouts will assist detection of genes involved in epistatic effects with CRISPR-Cas9 gene editing technology. Current developments in full-length mRNA sequencing are revealing extensive variation in post-transcriptional modification (e.g. Wang *et al*., 2022b), and particularly for structural genes it will be interesting to find how frequently different functions relate to transcript diversity.

Pleiotropic effects of genes selected during domestication contribute phenotypic diversity that are often related to diverse reproductive (flower and seed) traits. However, the presence of pleiotropic genes means that the functions cannot be separated. The widespread occurrence of pleiotropy may also be important in adaptation to changes in environmental stress, particularly in the light of predicted climate change (Stearns, 2010).

Identifying high value pleiotropic genes

Overlapping genomic regions provide evidence for the presence of pleiotropic genes impacting multiple traits. High value pleiotropic genes refer to those having major (≥ PVE 10%) effect and favourably (i.e. positive pleiotropy) impacting multiple traits. Multi-role pleiotropic genes impacting two or more traits have been discovered in diverse crops. As examples, the *glutathioneS-transferase* (*GST*) gene provides modest levels of resistance to southern leaf blight (*Cochliobolus heterostrophus* (Drechsler) Drechsler, 1934), northern leaf blight [*Setosphaeria turcicum* (Luttr.) K.J. Leonard & Suggs, (1974)], and grey leaf spot (*Cercospora zeae-maydis* Techon & E.Y. Daniels) in maize (Wisser *et al*., 2011). A pleiotropic stress-responsive gene, *OsSGL* confers multiple stress tolerances, increased grain length, grain weight, grains per panicle, and yield in rice (Wang *et al*., 2016) and was identified by overexpression in transgenic plants (reverse genetics) despite lack of natural variation in the gene. *TaNPF5.34* has a pleiotropic effect on plant height, ears per square meter, grains per ear, and 1000-grain weight across N regimes in wheat (Lisker *et al*., 2022). The *compactum* I locus, in addition to its effect on compact spike, exhibited pleiotropic effects on plant height, spike, and grain-related traits in wheat (Wen *et al*., 2022). A pleiotropic natural variant *NAL1* (*GPS*), selected for high yield in rice, also increases photosynthesis rate (Takai *et al*., 2013). Several pleiotropic loci control accumulation of structural and nonstructural carbohydrate accumulation in sorghum for

Table 2 Pleiotropic QTLs reported in cereal, food legumes, and oil crops from 2001 to 2024

Mating design	Pleiotropic QTL/genes and markers impacting multiple traits	References
Barley (*Hordeum vulgare*)		
GWAS	17 pleiotropic SNPs affected plant height and seed weight	He *et al.* (2023a)
Gene mutation	Row type mutant genes show pleiotropic effects on tillering, seeds spike^{-1}, seed size, and 1000-seed weight	Liller *et al.* (2015)
SSD and DHLs	*sdw1*, affect flowering, grain weight spike^{-1} and 1000-grain weight	Kuczyńska *et al.* (2014)
Brassica species		
RILs	Four pleiotropic QTLs, including qSR.A07, associated with high stem strength in rapeseed	Tian *et al.* (2024)
GWAS	*BnaA03g23490D*, *BnaC09g46370D*, *BnaA07g37150D*, *BnaA01g32590D*, and *BnaC09g37280D*, pleiotropic genes controlling multiple traits in *B. napus*	Xiang *et al.* (2023)
DHLs	Affect developmental traits, seed yield, and yield components or seed quality in rapeseed	Luo *et al.* (2017b)
DHLs	Pleiotropic QTLs affecting flowering, stress tolerance, and yield attributing traits in *B. napus*	Udall *et al.* (2006)
Common bean (*Phaseous vulgaris*)		
RILs	Pleiotropic QTLs control seed weight, root architecture, shoot and root traits, and shoot traits	Singh *et al.* (2019)
GWAS	A pleiotropic locus impacting seed weight, width, and length	Lo *et al.* (2019)
RILs	A major pleiotropic QTL for leaf growth trajectories across environments	Wei *et al.* (2018)
Chickpea (*Cicer arietinum*)		
GWAS	60 SNPs pleiotropic affecting multiple traits	Ilgokina *et al.* (2020)
Cotton (*Gossypium hirsutum*)		
GWAS	Genomic regions harbouring favourable pleiotropic loci and candidate genes to simultaneously improve yield and fibre quality	Wang *et al.* (2021)
Maize (*Zea mays*)		
Biparental population	Pleiotropic QTLs affecting multiple grain quality traits	Ndlovu *et al.* (2024)
GWAS and RILs	Forty common intervals impacted multiple ear traits	Dong *et al.* (2023)
F$_2$/F$_3$ and RILs	A pleiotropic consistent *QTL* for plant height overlapped with ear height	Fei *et al.* (2022)
GWAS and RILs	36 pleiotropic SNPs across 16 pleiotropic QTLs impacting multiple grain yield traits; *Zm00001d016656* associated with five traits across environments	Zhang *et al.* (2022)
RILs	Pleiotropic QTLs affecting stem-related traits	Shang *et al.* (2020)
NILs	*qKNPR6*, *ear length*, kernel number per row, kernel weight, and grain yield	Liu *et al.* (2012)
Pea (*Pisum sativum*)		
RILs	Major pleiotropic loci affecting seed testa thickness and permeability	Williams *et al.* (2024)
RILs	Pleiotropic QTLs controlling stress resistance and plant phenological and architectural traits	Boutet *et al.* (2023)
RILs	Pleiotropic QTLs affecting total seed protein, *in vitro* protein digestibility, methionine + cysteine concentration	Zhou *et al.* (2023)
GWAS	Shoot architecture (*PsLE*) or flowering (*PsTFL1*) genes with pleiotropic effects on root system architecture (RSA), and an SNP mapped to the major QTL *Ae-Ps7.6* pleiotropic to both resistance and RSA traits	Desgroux *et al.* (2018)
NILs	Genes affecting starch biosynthesis show pleiotropic effects on seed protein and composition	Hughes *et al.* (2001)
Peanut (*Arachis hypogaea*)		
GWAS	Twelve pleiotropic SNPs associated with multiple fatty acid composition traits	Otyama *et al.* (2022)
RILs	29 unique pleiotropic impacting multiple yield and yield attributing traits	Chen *et al.* (2017)
Pearl millet (*Pennisetum glaucum*)		
RILs	Affecting flowering and plant height	Kumar *et al.* (2017)
Quinoa (*Chenopodium quinoa*)		
F$_2$/F$_3$ mapping population	Affecting flowering, plant height, and 1000-seed weight	Maldonado-Taipe *et al.* (2022)
Rice (*Oryza sativa*)		
RILs	*Hd1*, *Hd16*, and *Ghd7* impacting culm length, panicle length, and % head rice grains; *Ghd7*, impacting # of panicles and grain size	Lee *et al.* (2024)
NILs	*qSCM4*, lodging resistance QTL, enhances panicle branches, and grains panicle^{-1}	Yang *et al.* (2023a)
BC$_6$F$_2$	EP4.2, DEP7 and DEP8 pleiotropic effect on panicle architecture traits	Yang *et al.* (2023b)
GWAS	33 pleiotropic loci impacting multiple traits associated with lodging resistance	Rashid *et al.* (2022)
NILs	*qph12*, simultaneously affected plant height, panicle length, spikelet number and grain yield	He *et al.* (2022)
Transgene and GWAS	*GSE5*, simultaneously affect grain chalkiness and grain shape	Jiang *et al.* (2022)
CRISPR/Cas9-meidated gene edting	*GW2*, plant and grain architecture, and nutritional quality	Achary and Reddy (2021)
RILs	*RFT1* (heading date genes), with a strong effect on most amino acids	Xie *et al.* (2020)
RILs	Straw N percentage, biomass, seed and seed N across treatments and seasons	Vishnukiran *et al.* (2020)

Table 2 Continued

Mating design	Pleiotropic QTL/genes and markers impacting multiple traits	References
GWAS	Pleiotropically impacting heading date (HD), grains panicle^{-1} (GNP), HD and panicles plant^{-1} (PN), HD and kilo-grain weight (KGW), GNP and PN, GNP and KGW, and PN and KGW	Li *et al.* (2018a)
NILs	*qHD5* affect flowering, flag leaf length and width, panicle branches, and 1000-grain weight	Sun *et al.* (2017)
RILs	A major pleiotropic QTL on chromosome 5 affect flowering, source size, and panicle traits	Zhan *et al.* (2015)
Advanced backcross lines	qHd1, pleiotropic effects on yield attributing traits	Chen *et al.* (2014)
NILs	*Ghd7.1*, flowering, plant height and spikelets per panicle	Liu *et al.* (2013)
NILs	*Ghd8*, a major QTL with pleiotropic effects on grain yield, heading date, and plant height	Yan *et al.* (2011)
NILs	*qSCM2*, enhances culm strength and increases spikelet	Ookawa *et al.* (2010)
RILs	qPLSD-9-1 and qPLSD-9-2, novel pleiotropic loci affecting panicle length and spikelet density	Guo and Hong (2010)
RILs/NILs	*qGL7*, exhibited pleiotropic effects on spikelets panicle^{-1}, grain length, width, and thickness, and 1000-grain weight	Bai *et al.* (2010)
NILs	Two pleiotropic intervals controlling three traits by the same QTL or tightly linked QTL	Zhang *et al.* (2008)
Sesame (Sesamum indicum)		
RILs	*qSmin11-1* and *qSmol*, *respectively*, *regulate* seed sesamin (PVE ~68%) and sesamolin (PVE 46%) lignans	Xu *et al.* (2021)
Sorghum (Sorghum bicolor)		
GWAS	*~0.61% of 122 000 SNPs exhibited significant pleiotropic effects for multiple bioenergy traits across the genome*	Kumar *et al.* (2024)
GWAS	Numerous pleiotropic and/or epistatic interactive effects influencing multiple carbon-partitioning and seed composition traits	Boatwright *et al.* (2022)
Soybean (*Glycine max*)		
NILs	*L1*, pleiotropic control pod colour, pod shattering, and seed pigmentation, l	Lyu *et al.* (2023)
NILs	*POWR1* (domestication gene) pleiotropically regulate seed quality and yield	Goettel *et al.* (2022)
RILs	*qPH18*, controls plant height, node number and internode length	Kou *et al.* (2021)
RILs	Leaf-related traits, chlorophyll content and 100-seed weight	Yu *et al.* (2020)
RILs	A novel and validated QTL cluster (8.8%–16.8% PVE) affecting 10 amino acids in soybean seed	Li *et al.* (2018b)
Wheat (*Triticum aestivum*)		
RILs	Pleiotropic effects on yield attributing traits	Liu *et al.* (2023)
RILs	*QWue.acn-2B* increases water use efficiency and enhance multiple agronomic traits	Hui *et al.* (2023)
RILs	Novel pleiotropic QTL (stable and co-localized QTL) control two or more traits	Rathan *et al.* (2023)
Transgene	Green revolution gene *Rht-B1b* with multiple pleiotropic effect on plant architecture and yield attributing traits	Xu *et al.* (2023)
RILs	*C* exhibited pleiotropic effect on plant height, spike- and grain-related traits	Wen *et al.* (2022)
RILs	*qSn-1A.1*, *qFsn-1B*, and *qFsn-7D* relate to spikelet fertility and kernel numbers under low N conditions	Fan *et al.* (2019)

DHLs, doubled haploid lines; GWAS, genome-wide association studies; RILs, recombinant inbred lines; SSD, single seed descent.

bioenergy traits and shared several marker-trait associations among bioenergy and compositional traits (Kumar *et al.*, 2024), while another pleiotropic variant control assimilate partitioning to grain vs vegetative biomass in sorghum (Maina *et al.*, 2022).

Multi-role pleiotropic genes impacting favourable and unfavourable pleiotropy include *BnUD1*, adversely impacted plant height and 1000-seed weight, but positively increased seeds silique^{-1} and photosynthetic efficiency in rapeseed (Yang *et al.*, 2016); while *POWR1* increased seed oil content, seed weight, and seed yield but reduced seed protein content in soybean (Goettel *et al.*, 2022). A novel pleiotropic QTL, *qSS14* simultaneously controls 100-seed weight, seed length, seed width, and seed thickness and is predicted to contain two candidate genes in soybean (Yuan *et al.*, 2023). Seed coat permeability is an important trait in many legumes. The QTL gene *PG031* regulates seed coat permeability and seed weight in soybean. Three haplotypes and a single SNP explain natural variation in seed permeability, and overexpression of the impermeable allele $PG031^{289H}$ significantly reduces seed coat permeability and 100-seed weight in transgenic soybean (Wang *et al.*, 2022a).

Allelic variation within a pleiotropic locus, e.g. domestication gene (*Q*) in wheat, has also shown differential pleiotropy. The *Q* allele increases grain yield, grains per m^{-2}, and 1000-grain weight, but reduced grains per spike/spikelet, whereas the *q* allele contributed to more grains per spike/spikelet (Xie *et al.*, 2018). Q^{s1}, another variant had no effect on grain protein but adversely affected grain yield by decreasing grain weight and grains per spike, while Q^{c1}-*N8* positively affected grain protein and yield by increasing grain weight and grains per spike (Chen *et al.*, 2022).

Epistasis contributes substantially to hybrid vigour in crops

Hybrid vigour or heterosis refers to the superior performance of F_1 hybrid plants over the average of its parental lines or above the high performing parent and has been extensively exploited towards the development of hybrid cultivars in both cross- and self-pollinated crops. Epistasis, involving non-allelic gene interactions, strongly affects the performance of hybrids. Maize hybrids involving temperate and tropical maize inbred lines with varying levels of interparental divergence show significantly higher mid-parent heterosis in temperate by tropical hybrids than in the temperate-by-temperate hybrids. Epistasis effects relative to dominance effects contribute more to enhanced performance in temperate by tropical hybrids and detected 33 and 420 epistatic QTL for grain weight per plant and days to silking. The epistatic genes involved in protein interactions and minor effect gene interactions in the hybrids could activate the transcription activators of epistatic genes and amplify yield heterosis (Sang *et al*., 2022).

Variation between two parental genomes may generate epistasis interactions. The proteome of each parent may provide novel protein to protein interacting (PPIs) partners specific to the F_1 hybrid. An *in silico* study involving the hybrid interactome between rice *japonica* (cv., Nipponbare) and *indica* (cv., 9311) cultigen pools predicted 4612 hybrid-specific PPIs accounting for 20.5% of total PPIs observed in the hybrid interactome. Genes involved in hybrid PPIs encode metabolic enzymes are localized in genomic regions harbouring metabolic gene clusters. Following genomic selection and assuming additive, dominant, and epistatic effects considered separately in the model, it shows that the removal of SNPs associated with hybrid PPIs reduces the prediction accuracy when epistatic effects are considered in the model, but no such significant changes are observed when additive or dominant effects are considered. This clearly shows that genomic divergence between *japonica* and *indica* rice cultigen pools may generate hybrid-specific PPIs, some of which may cumulatively contribute to hybrid vigour (Li *et al*., 2020). These observations regarding the influence of epistasis on heterosis, however, are based on a simplified multiplicative model that assumes very few genes, only additive × additive epistasis, and LD.

A simulation-based study involving diverse populations (nine populations, the selfed populations, the 36 interpopulation crosses, 180 doubled haploids, and their 16 110 crosses) and assuming a more complex model (additive model, hundreds of genes, LD, dominance, and seven types of digenic epistasis interactions) shows that epistasis may have a negative impact on the heterosis of panmictic populations (i.e. population where all potential parents may contribute equally to the gamete pool, and that these gametes are uniformly distributed within the population), leading to erroneous inferences regarding the identification of superior and most divergent populations. This, however, depends on the type of predominant epistasis, percentage of epistatic genes, and magnitude of the epistatic effects. In general, there is a decrease in the average heterosis by increasing the percentage of epistatic genes and the magnitude of their effects. Duplicate genes with cumulative effects and non-epistatic genic interaction maximize average heterosis, whereas for other type of epistasis, increasing the percentage of epistatic genes and the ratio of epistatic variance/(additive plus dominance variances) decreases the average heterosis. Clearly the negative impact of epistasis in heterosis cannot be avoided if the genetic system involves a high number of epistatic genes with great effects (Viana, 2023).

Modelling epistasis in genomic selection of complex traits

Many loci with small effects impact variation of complex characteristics. Genomic prediction is used for estimating breeding values of an individual based on genome-wide SNPs. However, breeding values often ignore the non-additive genetic factors such as epistasis (McGaugh *et al*., 2021). Hence, epistatic genomic prediction models that incorporate SNP interactions are necessary. In this regard, epistatic random regression best linear unbiased predictor (ERRBLUP) and selective epistatic random regression BLUP (sERRBLUP) are useful for modelling epistasis in genomic prediction (Vojgani *et al*., 2021b). EERBLUP uses a full epistatic model that considers all pairwise SNP interactions, while sERRBLUP selects a subset of the pairwise SNPs according to their absolute effect size or the effect variances. Both methods are implemented in the R Package EpiGP that efficiently handles large genomic data. sERRBLUP gives a significant prediction accuracy increase when keeping the optimal proportion of SNP interactions in the model, particularly if they are selected based on their effect sizes.

Genomic prediction accuracy increased in a maize doubled-haploid population derived from two European landraces after including top-ranked SNP interactions in the model (Vojgani *et al*., 2021a), thus demonstrating that sERRBLUP is better than genomic BLUP. Bivariate models considering selected subsets of pairwise SNP interactions further improved prediction accuracy when using multi-environment trial data, assuming a high genomic correlation across years (Vojgani *et al*., 2023). Roth *et al*. (2022) have shown further that a genomic prediction model incorporating additivity, dominance, epistasis, and inbreeding effects was the most robust in two admixed maize hybrid populations, which derived from recombination between structural groups. Similarly, additivity × additivity (a × a) epistasis improved predicting ability in inbred wheat lines (Raffo *et al*., 2022), which was not surprising since a × a is the most important non-additive genetic effect in such a population and useful for developing new cultivars when selecting based on total genetic merit. Epistatic interactions affect flowering time, maturity, and seed size in cowpea (Olatoye *et al*., 2019), which should be taken into account for improving genomic prediction accuracy of selection models for such traits. Moreover, Derbyshire *et al*. (2021) indicated that the accuracy of breeding values was improved between 3 and 40% when adding (a × a) epistasis in linear mixed prediction models for sclerotinia stem rot resistance in rapeseed. Models with SNP interactions also increased prediction accuracy for implementing selection for rhizomania resistance in sugar beet (*Beta vulgaris* L.) (Lange *et al*., 2023). The above results highlight the importance of acknowledging epistasis terms in genomic prediction models for complex characteristics.

Genomic selection models exhibit a spectrum of complexity, encompassing Bayesian methods, ridge regression, random forests, support vector machines, and neural networks. The selection of a model hinges upon factors such as the dataset's size and structure, available computational resources, and the traits earmarked for selection (Alemu *et al*., 2024).

The major limiting factors in genomic selection can vary depending on the context and the specific application. However, both big data and genomic selection model-related factors can present challenges (Bassi *et al*., 2023). Dealing with large-scale genomic data can be computationally intensive and requires robust infrastructure and computational resources. Challenges related to big data in genomic selection include data storage, data management, data preprocessing, and computational efficiency (Bassi *et al*., 2023). Handling and analysing massive datasets efficiently can be a significant bottleneck, especially for organizations with limited computational resources. As noted by Desta and Ortiz (2014), the complexity and accuracy of the genomic selection model are crucial factors in its effectiveness. Developing accurate and reliable prediction models requires a comprehensive understanding of genetic architectures underlying the traits of interest, as well as suitable statistical and machine learning methodologies. Choosing an appropriate model architecture and optimizing its parameters can be challenging, especially when dealing with complex traits or heterogeneous populations (Crossa *et al*., 2017). Additionally, ensuring the scalability and robustness of the model across different populations and environments is essential for its practical utility.

Machine learning to capture epistasis and pleiotropy interactions for breeding

Machine learning has proven to be a valuable tool for capturing epistasis and pleiotropy interactions in plant breeding, with several advantages over more traditional methods. Machine learning algorithms have the ability to capture complex interactions and patterns in large datasets (Costa *et al*., 2022). By considering the interactions between multiple genes, machine learning models can provide more accurate predictions of plant traits compared to traditional statistical models that assume linear relationships.

Machine learning algorithms excel at capturing non-linear relationships. By explicitly modelling these interactions, machine learning models can identify and quantify the nonadditive effects, providing breeders with a better understanding of the underlying genetic architecture. Machine learning can also reveal complex and subtle interactions that may go unnoticed with traditional statistical approaches. By examining a large number of genetic markers simultaneously, machine learning models can identify unexpected relationships and interactions that contribute to the observed trait variation. Machine learning can integrate genotypic data with complex phenotypic data such as drone imagery to identify markers associated with complex traits that cannot easily be tabulated. This can lead to the discovery of new genetic markers or pathways that influence the target traits, providing valuable insights for plant breeders. By identifying specific gene interactions and their effects on multiple traits, machine learning can aid in the design of more targeted and efficient breeding strategies. Machine learning models can also assist in the optimization of a breeding strategy by identifying key genetic markers or regions for marker-assisted selection or genomic selection (GS).

Various machine learning algorithms can be employed to capture epistasis and pleiotropy interactions and some software packages have been developed specifically for this task (Chang *et al*., 2020; Guo *et al*., 2019). These algorithms range from traditional statistical models like linear regression to more complex models such as random forest, support vector machines, or deep learning architectures such as neural networks. The choice of the algorithm depends on the specific characteristics of the dataset and the complexity of the interactions being studied.

The application of machine learning to capture epistasis has predominantly been applied in biomedical fields but has significant potential for crop studies. While offering several advantages over traditional methods, the effectiveness of machine learning in capturing epistasis and pleiotropy interactions relies on high quality and a significant quantity of data, as well as the choice of appropriate features and algorithms. The interpretation of the results should be done with caution, as machine learning models provide correlations rather than causation. Thus, validation and further experimental verification are often necessary to confirm the identified interactions and their practical implications in plant breeding.

Balancing trade-offs as influenced by epistasis and pleiotropy genes

Trade-off refers to 'situations when one trait cannot increase without a decrease in another trait (or *vice versa*)' (Garland Jr., 2014). A variety of trade-offs particularly from genetic and ecological adaptation viewpoints exist, for example, resource acquisition-resource conservation, source-sink, growth defence, and yield nutrition. Pleiotropic effects and tight linkage influence genetic trade-off. Balancing trade-off involving resource acquisition and resource conservation or between multiple stresses, productivity and nutritional quality are significant challenges to adaptation of new species in ecologically different environment or plant breeders to tailor new cultivars (Dwivedi *et al*., 2021).

Sorghum is both a grain and biofuel crop. Bioenergy sorghum rapidly accumulates significant amounts of carbon (C) per unit time per unit area per unit input as above-ground biomass. The interaction between source and sink organs greatly influences C-partitioning in plants. The accumulation and distribution of fixed C is dependent upon the source, pathways, and interactions of the system. A holistic understanding of complex systems demands adoption of multiscale phenotypes including phenotypic measurements (e.g. plant height, above-ground biomass, and dry weight) and compositional traits (ash, lignin, cellulose, hemicellulose, water soluble carbohydrate, water) to deconvolute the C-partitioning pathways on a diverse germplasm panel. Multivariate analyses of these traits in a sorghum C-partitioning panel (Boatwright *et al*., 2021) detected numerous loci associated with several C-partitioning traits, which putatively regulate sugar content, manganese homeostasis, and nitrate transportation. The identification of several loci associated with multiple traits suggests pleiotropic and/or interactive effects, positively influence multiple C-partitioning traits and the overlap indicate molecular switches mediating C-allocation or partitioning networks, and detected C-trade-off wherein reduced lignin content is associated with increased sugar content (Boatwright *et al*., 2022).

C_4 perennial bunch grass (*Panicum hallii* Vasey), adapted to diverse habitats and climates in North America, has two distinct ecotypes, xeric upland and mesic lowland that are classified as distinct varieties, xeric (*P. hallii var. hallii*), and mesic (*P. hallii var. filipes*) types (Gould, 1975). Var. *hallii* is a large seeded upland ecotype adapted to xeric environments, while var. *filipes* is a small-seeded lowland ecotype adapted to mesic environments. Heterogeneity in the environment may drive patterns of functional trait variation and evolution of locally adapted ecotypes. The correlated root and shoot traits that share common genetic and physiological relationships differentiate plant ecotypes. While shoot traits determine plant water loss, it is the root systems that determine water access and constraint shoot water status. Several genomic hotspots, involving mesic and xeric ecotypes of *P. halli*, control suits of correlated root and shoot traits. The genetic coordination between plant organs contributes to ecotype divergence. This study shows that colocalized QTL for most root and shoot growth related traits are independent of colocalized QTL for root and shoot resource acquisition traits. The allelic effects of individual QTL highlight ecological adaptation between ecotypes, which may breakdown due to epistatic interactions (Khasanova *et al*., 2019). Furthermore, traits with ecological significance, for example seed mass, drive ecotype adaptation. Seed mass may impact adult and reproductive traits. Greenhouse and field evaluations of *P. hallii* accessions belonging to both ecotypes reveal that seed mass in the greenhouse varied greatly across ecotype divergence (xeric and mesic ecotypes of *P. hallii*) and correlate with seedling and reproductive traits, whereas in field environments representing xeric and mesic habitats, it has different impacts on seedling and reproduction. Upland habitat favours large seeds, whereas lowland habitat small seeds, consistent with local adaptation, and may provide opportunity to unlock life-history trade-offs across ecotype divergence (Razzaque *et al.*, 2023).

Balancing trade-offs is a significant breeding challenge, for example, pleiotropic *Ideal Plant Architecture 1* (*IPA1*) gene increases grains per panicle but reduces tillers in rice. Deletion of a 54-base pair *cis*-regulatory region in *IPA1* via a tiling-deletion-based CRISPR–Cas9 system mediates the trade-off between grains per panicle and tiller number, which significantly improves seed yield per plant. The deleted fragment is a target site for the transcription factor An-1, which represses *IPA1* expression in panicles and roots (Song *et al.*, 2022). Plants have evolved diverse root system architectures and function across environments. Plant species with inexpensively constructed roots (less biomass per unit root length) explore more nutrient hotspots at the cost of frequent root-tissue turnover, while those with costly constructed roots (more biomass per unit root length) are more conservative in resource acquisition and maintain higher root-tissue persistence. Switchgrass (*Panicum virgatum* L.) adapted to southern lowland displays root features with more costly constructed roots than those of inexpensively constructed roots of northern upland ecotype. The evidence suggests multiple genetic linkages among root morphology, growth, and turnover in two switchgrass ecotypes. Swapping alleles derived from southern ecotypes to northern ecotypes increases root turnover but reduces tissue investment in enhancing root length. Local adaptation based on root systems is facilitated by the trade-off between root construction and turnover strategy along the warm to cold climate gradients of the species range. Improving the bioenergy yield of switchgrass does not clash with enhancing root-derived carbon sequestration (Chen *et al.*, 2021).

In soybean, *Ln* locus pleiotropically regulate leaf shape and seeds per pod. *Ln/Ln* soybean(s) have broad leaflets and low seeds per pod, while *ln/ln* soybean(s) have narrow leaflets and high seeds per pod. QTL analysis revealed that *Ln* is associated with leaf, pod, and seed traits. *ln* allele(s) increases leaflet length, leaflet length/leaflet width ratio, and 3- and 4-seeded pods, but decreases leaflet width, leaflet area, 2-seeded pods, and 100-seed weight. Nevertheless, the effect of *ln* allele(s) towards 3-seeded pods and 100-seed weight is not always detected. Two- and four-seeded pods are negatively correlated. Therefore, reducing in 2-seeded pods cancel the effect of 4-seeded pods in enhancing seed yield in *ln/ln* soybean. Thus, *Ln* pleiotropically and antagonistically regulates 2- and 4-seeded pods (Chanchu *et al.*, 2023).

Variation in inflorescence architecture has direct bearing on grain yield in cereals. It is hypothesized that introducing variation in inflorescence architecture with a branching spike, such as Miracle-Wheat (*Triticum turgidum* convar. *compositum*), may enhance grain yield in wheat. However, the increase in spikelet number in such accessions is generally not translated to greater grain yield advantage because of reduced grains per spikelet and grain weight possibly due to limitation in source-sink strength. Biparental RILs involving spike-branching landraces and elite durum wheat line unlocked a new modifier QTL for spike branching, *branched head 3* (*bht-A3*), epistatic to the known *bht-A1* locus. *bht-A3* also confers more grains per spikelet and a delay in flag leaf senescence rate. More importantly, the favourable alleles that delay senescence (*bht-A3* and *gpc-B1*) are also critical to improve grain number and grain weight in the spike-branching accessions. Thus, achieving a balanced source–sink relationship can minimize grain yield trade-offs in Miracle-wheat (Abbai *et al.*, 2024). Multi-trait principal component analysis-based QTL mapping increases the power to detect novel small effect pleiotropic loci across related traits to optimize trade-offs between nutritional and productivity traits in wheat (Fradgley *et al.*, 2022).

Immunity to disease resistance costs penalty to growth and development in plants. Biomass yield (i.e. grain or stover yields) in field crops is the main breeding target, usually affected by levels of disease resistance. A proper balance between growth and defence is critical for achieving efficient crop improvement. Growth-defence trade-off is mediated by resistance (R) genes, susceptibility (S) genes and pleiotropic genes. Plant breeding is expected to make significant advancement through creative incorporation of artificial intelligence algorithms to maintain a proper balance in achieving appropriate levels of disease resistance at no yield penalty. The integrated disease management, which combines diverse strategies to prevent plant diseases while minimizing the impact on crop yield, has the potential to expedite the breeding of next generation crops, thereby fostering the development of more sustainable and resilient agricultural systems (Gao *et al.*, 2024; Xu *et al.*, 2022).

Selecting for pleiotropic loci with favourable effects on traits and favourable loci with no detectable pleiotropic effects in breeding programs

Pleiotropy promotes the evolution of complexity (Wang *et al.*, 2010). Crops are useful genetic resources to dissect the effects of pleiotropy on multi-trait evolution. Large variation in flowering and maturity (nine major-effect loci, *E1* to *E8*, *J*) is available in soybean germplasm (Watanabe *et al.*, 2012). Pleiotropic variation regulates both flowering and maturity in this crop (Nelson, 1988; Smith and Nelson, 1986). Whether these loci independently control flowering and maturity in soybean was investigated recently. Domestication and diversification of soybean favoured pleiotropic loci (*E* loci) controlling both flowering and maturity. Novel non-pleiotropy loci were isolated following stringent selection in a cross between cultivated and wild (*G. soja* Siebold & Zucc.) soybean, indicating that pleiotropy in flowering and maturity can be genetically separated (Sedivy *et al.*, 2020). Long-term wheat breeding targeted loci with favourable pleiotropy. Analysis of segregating population of a cross between ancestral and modern parents reveals that modern wheat cultivars contain the *Rht-B1b* green revolution semi-dwarfing allele and compensatory alleles that reduce its negative effects. Furthermore, pleiotropic loci with favourable effects on traits and favourable loci with no detectable pleiotropy contributed to the improvement in other traits (Raherison *et al.*, 2020). Thus, combining pleiotropic loci with favourable effects and favourable loci with no detectable pleiotropy is an efficient selection strategy to enhance breeding value of the elite lines.

Combining pleiotropy and regulatory SNPs as novel breeding target for improvement of multiple traits

SNP-based molecular genetic markers have been applied for crop breeding for more than 20 years (Batley *et al.*, 2003). Initial applications were to select for loci that provided single beneficial traits that were not readily amenable to phenotypic assessment (Edwards and Batley, 2004). With the development of genomic selection methods, initially in livestock (Hayes *et al*., 2009) but later adapted for crop breeding, there was a growth in the application of genome wide SNPs in breeding. However, these approaches often failed to capture epistatic interactions, limiting potential gains. Several advances are required to capture and apply epistasis in crop breeding. These include the development of advanced pangenomes representing the diversity of available breeding germplasm (Edwards and Batley, 2022); the continued efficiency and reducing cost of DNA sequencing; the expansion of high throughput phenotyping and sharing of this data (Danilevicz *et al*., 2021b); and the development and application of advanced machine learning tools (Bayer *et al*., 2022). These advances continue to accelerate. Pangenomes are increasingly available for a wide range of crop species (Zanini *et al*., 2021), even those with large and complex genomes (Bayer *et al*., 2022; Montenegro *et al*., 2017). Costs for short read DNA sequencing continue to decline, making DNA sequencing based whole genome genotyping ever more cost effective, while large scale phenomic studies are increasingly being applied for trait prediction (Danilevicz *et al*., 2021a). This growth of technology supports the identification of pleiotropic variants and epistatic SNPs that can be modelled by advanced explainable AI methods to provide an even greater accuracy of traits, supporting both traditional crop breeding as well as genome editing based approaches (Scheben *et al*., 2017).

Applying simultaneous selection for combining nonallelic interactions and high value pleiotropic loci in hybrid breeding

As noticed by Rasmusson and Phillips (1997), epistasis arising from *de novo* non-allelic interactions seems to be very important when selecting in offspring derived from elite breeding pools; i.e., crossing good × good. Likewise, it appears that long-term selection may be accumulating coadapted gene arrangements (Breen *et al*., 2012), which may be disrupted subsequently. Non-allelic interactions may be also affected by a changing environment (Anholt, 2020). The importance of epistasis in plant germplasm enhancement depends on crop breeding systems and genetic improvement methods because they determine the types of epistatic effects to reliably propagate (Holland, 2001). The effect of epistasis has been demonstrated in segregating populations of beans (de Melo *et al*., 2023), cowpea (*Vigna ungiculata* (L.) Walp.) (Olatoye *et al*., 2019), maize (Sofi *et al*., 2007), tomato (*Solanum lycopersicum* L.) (Soyk *et al*., 2017), and wheat (Sannemann *et al*., 2018), among other crops.

Viana (2023) indicated that an increase on epistatic genes and the magnitude of their effects decrease the average heterosis and impact on combining ability analysis in populations of doubled haploids and inbred lines. Moreover, prediction accuracy significantly increased in plant breeding populations after including 2-locus epistatic effects (Wang *et al*., 2012). Nonetheless, modelling and simulation research suggest that phenotypic selection may be superior to marker-aided recurrent selection and genomic selection under epistasis (Ali *et al*., 2020).

According to de Visser *et al*. (2011), pleiotropy is key for epistasis, which could evolve as an adaptive change. Genes involved in pleiotropy affect expression of various traits resulting from the covariance among them (Pavlicev *et al*., 2011). For example, pleiotropic QTLs are important for simultaneously improving multiple amino acid concentrations in soybean (Li *et al*., 2018b). Hence, marker-aided breeding must consider these pleiotropic QTL clusters for multiple amino acid improvement. Moreover, these loci responsible for genetic variation due to pleiotropy are involved in epistatic interactions. The CAPE has been used to model genetic networks affecting multiple characteristics because both are necessary to understand the genetic architecture of complex traits (Vivek *et al*., 2014); the associated R/cape allows the detection of interactions affecting phenotypes simultaneously (Tyler *et al*., 2013).

Precision breeding requires accurate prediction of phenotypes from genotypes. Accurate genomic prediction of complex traits across environments, breeding cycles, and populations remains a challenge in plant breeding. The genomic prediction accuracy of phenotype is compromised in the presence of epistasis, while prediction accuracy can be increased if epistasis is factored into the prediction model (Vojgani *et al*., 2021a). The adoption of multi-trait ensemble genomic prediction models, relative to single trait models, further increases prediction accuracy, as evident in 16-founder wheat multi-parent advanced generation inter-cross population. The non-parametric models (i.e. random forest) for complex traits outperformed simplified additive models (i.e. LASSO). Clearly, factoring both epistasis and pleiotropy in breeding practices may provide greater insights into mechanisms for sustained genetic grain in a limited gene pool and optimization of multiple traits for crop improvement (Fradgley *et al*., 2023).

Non-additive genetic (G × G) and genotype × environment (G × E) interactions generate allelic substitution effects that are non-stationary across different contexts. Such non-stationary effects of alleles are either ignored or presumed to be implicitly captured by most gene-to-phenotype (G2P) maps used in genomic prediction, which necessitates that G2P maps should be re-estimated across different contexts. Hierarchical G2P maps explicitly capture non-stationary effects of alleles to improve short-term prediction accuracy. Long-term prediction remains a challenge. The breeding community should undertake complimentary simulation and empirical experiments to interrogate various hierarchical G2P maps that connect G × G and GEIs simultaneously. To achieve this, the existing genetic correlation matrix can be used to assess the magnitude of non-stationary effects of alleles and the predictive ability of these hierarchical maps in long-term, multi-context genomic prediction of complex traits in crop breeding (Powell *et al*., 2021). Further, a large proportion of the missing heritability hidden in epistatic interactions can be retrieved by developing a comprehensive epistatic 2D maps (termed as next-gen GWAS, NGG) with sufficient SNP density to achieve gene-level resolution, as evidenced in *Arabidopsis*, and can be readily adapted in other biological models (Carré *et al*., 2024).

Identifying conserved (latent genetic variation) pleiotropy genes

Mutation, gene duplication, and gene loss are the primary causes of evolution and divergence of gene function (Holland *et al*., 2017). A distinction between divergence of gene function and conserved function across species along the evolutionary time frame is challenging. Large evolutionary distances often fail to unmask the full breadth of gene function, exposing only part of a gene's pleiotropic functions. *WUSCHEL HOMEOBOX9* (*WOX9*), a plant homeobox gene, has species-specific roles in embryo and inflorescence development. Assessment of functional divergence of *WOX9* in tomato by *cis*-regulatory editing drive system identified hidden pleiotropic roles for *WOX9*, the functions of which are conserved in groundcherry, a relative of

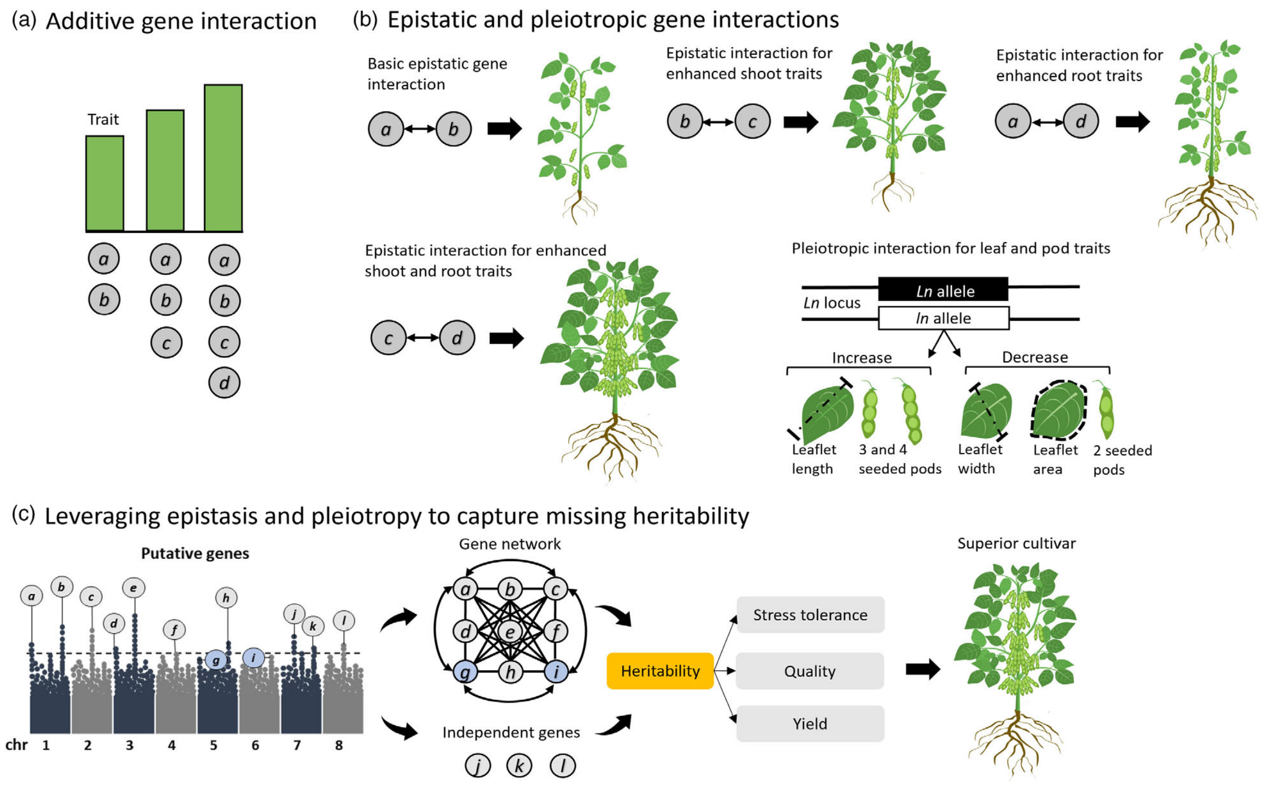

Figure 2 Schematic diagram showing the potential of epigenetic and pleiotropic gene interactions for crop breeding. (a) Basic representation of additive gene interaction, whereby each gene contributes equally to the phenotype. (b) Different scenarios of epistatic interactions, whereby the interaction of genes *b* and *c*, and *a* with *d*, enhances shoot and root traits, respectively, contributing to different yield outcomes that are better than the basic interaction of genes *a* and *b*. The combined effects of the epistatic interaction of genes *c* and *d* result in further enhancement of yield. Furthermore, the *ln* allele affects several traits in soybean in an antagonistic manner, in that the presence of this allele increases leaflet length, the number of three and four-seeded pods, while decreasing leaflet width, leaflet area, and the number of two-seeded pods. (c) Harnessing these interactions has the potential to capture missing heritability, which is not explained by classic additive gene interactions. Genes in blue-shaded circles (*g*, *i*) represent non-significant or low-effect genes but are still important in the gene network and along with independent genes (*j*, *k*, *l*) contribute to capturing the missing heritability of agronomic traits, including tolerance to stresses, superior quality and yield. This provides a promising approach for developing cultivars with superior traits.

tomato. Mimicking these alleles in *Arabidopsis* reveals new inflorescence phenotypes, suggesting conserved pleiotropy in distantly related species (Hendelman *et al*., 2021 and references therein). Browning of freshly cut eggplant (*Solanum melongena*) fruit reduces its commercial value. *Polyphenol oxidase* (*PPO*) contributes to the browning of fruits and vegetables (Hamdan *et al*., 2022), with most of the *PPO* genes clustered on chromosome 8 (Maioli *et al*., 2020). Mutagenesis of the *PPO2* by CRISPR/Cas9 system resulted in a diverse spectrum of phenotypes, ranging from plant dwarfism and architectural differences, early flowering, fruit set and maturation, flower colour, fruits size, shape, colour and weight to reduced and delayed browning of fruit-cuts and fewer seeds, thereby indicating hidden pleiotropic role of *PPO* gene in eggplant (Kodackattumannil *et al*., 2023). Sequencing and cataloguing of large mutant populations, particularly in the protein-coding regions, are powerful tools to uncover hidden variation (or pleiotropy), as evidence in polyploid wheat (Krasileva *et al*., 2017).

Conclusion

Genes associated with crop domestication contribute to phenotypic diversity for multiple traits. Accessing and exploiting epistasis interactions and pleiotropy effects may help breeders predict and control trait expression for the development of an efficient selection strategy (Figure 2). Epistasis bias reduces the speed at which crops can be improved. Advances in generating large-scale genomics and phenomics datasets and factoring all loci including those with small effects in association genetics, with associated computational analysis, may unlock and reduce epistasis bias in trait prediction. New approaches including use of artificial intelligence, use of pangenomes representing diversity in breeding germplasm, the continued improvement and reduction in genotyping cost, and advances in phenomics and data sharing may further facilitate detection of causal variation. Artificial intelligence can be implemented through self-learning algorithms that may be extremely efficient in detecting unanticipated pleiotropic or epistatic interactions and have the potential to make connections between genes, genome, phenotype, and even metabolome from diverse publications. Future developments of digital twins of these complex genetic phenomena may then reveal gene combinations that can be selected in a breeding program giving enhanced characteristics. Pleiotropy and LD cause genetic correlations, and their decoupling may provide new avenues to tailor germplasm modelling pleiotropic QTL, epistasis, and associated SNPs combined with multi-trait GWAS mapping may overcome unfavourable correlations.

The evidence to date suggests significant effects of epistasis and pleiotropy impacting plant phenology, host plant resistance, seed morphology, yield, and quality. A substantial proportion of detectable phenotypic variation is also governed by non-pleiotropic loci. Identifying and exploiting high value pleiotropic genes impacting multiple favourable traits may enhance breeding efficiency. Multi-role pleiotropic genes (*GST* in maize, *BnUDI* in rapeseed, *OsSGL* and *NAL1* (*GPS*) in rice, *PG031*, *POWR1*, and *qSS14* in soybean, and *compactum1*, *Q*, and *TaNPF5.34* in wheat) are known in diverse crops. Pleiotropic effects and tight linkage influence genetic trade-off. Many pleiotropic loci (i.e. *IPAI* in rice, *Ln* in soybean, and *bhtA3*, *bhtA1* in wheat) mediate trade-offs between competing traits. These genes (or their functional orthologs) may be deployed in breeding programs to mediate complex trade-offs.

Epistasis relative to dominance effects contribute more to enhanced performance of hybrid cultivars due to involvement of protein-to-protein interactions. However, epistasis may also cause a negative impact on heterosis provided a higher number of epistatic genes with great effects are involved. Modelling epistasis in genomic selection, which incorporates SNP interactions, provides better genomic prediction accuracy for complex traits. Simultaneous selection for nonallelic interactions and high value pleiotropic loci may model genetic networks affecting multiple traits. Detecting and exploiting higher order epistasis interactions (i.e. protein to protein interaction in F_1) between traits and multiple genes may account partially for missing heritability, though computationally intensive for complex traits. The sequencing and cataloguing of protein-coding region of large mutant populations, however, is a powerful tool to uncover hidden variation (pleiotropy) for use in crop improvement.

Acknowledgements

Appreciation to anonymous reviewers for making useful suggestions to improve the manuscript. Open access publishing facilitated by The University of Western Australia, as part of the Wiley - The University of Western Australia agreement via the Council of Australian University Librarians.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

All authors contributed to the writing and editing of this manuscript, and all approved the final version for submission.

Data availability statement

All data are available in the manuscript or in the cited publications.

References

Abbai, R., Golan, G., Longin, C.F.H. and Schnurbusch, T. (2024) Grain yield trade-offs in spike-branching wheat can be mitigated by elite alleles affecting sink capacity and post-anthesis source activity. *J. Exp. Bot.* **75**, 88–102.

Achary, V.M.M. and Reddy, M.K. (2021) CRISPR-Cas9 mediated mutation in *GRAIN WIDTH* and *WEIGHT2* (*GW2*) locus improves aleurone layer and grain nutritional quality in rice. *Sci. Rep.* **11**, 21941.

Alemu, A., Åstrand, J., Montesinos-López, O.A., Sánchez, J.I., Fernández-González, J., Tadesse, W., Vetukuri, R.R. *et al.* (2024) Genomic selection in plant breeding: key factors shaping two decades of progress. *Mol. Plant*, **17**, 552–578.

Ali, M., Zhang, L., DeLacy, I., Arief, V., Dieters, M., Pfeiffer, W.H., Wang, K. *et al.* (2020) Modeling and simulation of recurrent phenotypic and genomic selections in plant breeding under the presence of epistasis. *Crop J.* **8**, 866–877.

Ang, R.M.L., Chen, A.-A.A., Kern, A.F., Xie, Y. and Fraser, H.B. (2023) Widespread epistasis among beneficial genetic variants revealed by high-throughput genome editing. *Cell Genom.* **3**, 100260.

Anhalt, U.C., Heslop-Harrison, J.S., Piepho, H.P., Byrne, S. and Barth, S. (2009) Quantitative trait loci mapping for biomass yield traits in a *Lolium* inbred line derived F_2 population. *Euphytica*, **170**, 99–107.

Anholt, R.R.A. (2020) Evolution of epistatic networks and the genetic basis of innate behaviors. *Trends Genet.* **36**, 24–29.

Bai, X., Luo, L., Yan, W., Kovi, M.R., Zhan, W. and Xing, Y. (2010) Genetic dissection of rice grain shape using a recombinant inbred line population derived from two contrasting parents and fine mapping a pleiotropic quantitative trait locus *qGL7*. *BMC Genet.* **11**, 16.

Bartual, R., Lacasa, A., Marsal, J.I. and Tello, J.C. (1993) Epistasis in the resistance of pepper to phytophthora stem blight (*Phytophthora capsici* L.) and its significance in the prediction of double cross performances. *Euphytica*, **72**, 149–152.

Bassi, F.M., Sanchez-Garcia, M. and Ortiz, R. (2023) What plant breeding may (and may not) look like in 2050? *Plant Genome*, **17**, e20368.

Bateson, W. (1909) *Mendel's Principles of Heredity*. Cambridge: Cambridge University Press.

Batley, J., Barker, G., O'Sullivan, H., Edwards, K.J. and Edwards, D. (2003) Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Physiol.* **132**, 84–91.

Bayer, P.E., Petereit, J., Danilevicz, M.F., Anderson, R., Batley, J. and Edwards, D. (2021) The application of pangenomics and machine learning in genomic selection in plants. *Plant Genome*, **14**, e20112.

Bayer, P.E., Petereit, J., Durant, É., Monat, C., Rouard, M., Hu, H., Chapman, B. *et al.* (2022) Wheat Panache: A pangenome graph database representing presence–absence variation across sixteen bread wheat genomes. *Plant Genome*, **15**, e20221.

Bhat, J.A., Feng, X., Mir, Z.A., Raina, A. and Siddique, K.H. (2023) Recent advances in artificial intelligence, mechanistic models and speed breeding offer exciting opportunities for precise and accelerated genomics-assisted breeding. *Physiol. Plant.* **175**, e13969.

Boatwright, J.L., Brenton, Z.W., Boyles, R.E., Sapkota, S., Myers, M.T., Jordan, K.E., Dale, S.M. *et al.* (2021) Genetic characterization of a *Sorghum bicolor* multiparent mapping population emphasizing carbon-partitioning dynamics. *G3*, **11**, jkab060.

Boatwright, J.L., Sapkota, S., Myers, M., Kumar, N., Cox, A., Jordan, K.E. and Kresovich, S. (2022) Dissecting the genetic architecture of carbon partitioning in sorghum using multiscale phenotypes. *Front. Plant Sci.* **13**, 790005.

Bocianowski, J. (2014) Estimation of epistasis in doubled haploid barley populations considering interactions between all possible marker pairs. *Euphytica*, **196**, 105–115.

Boehm, F., Chesler, E.J., Yandell, B. and Broman, K.W. (2019a) Testing pleiotropy *vs.* separate QTL in multiparental populations. *G3*, **9**, 2317–2324.

Boehm, F., Yandell, B. and Broman, K.W. (2019b) qtl2pleio: testing pleiotropy vs. separate QTL in multiparental populations. *J. Open Source Softw.* **4**, 1435.

Borel, J.C., Ramalho, M.A. and Abreu, A.F. (2016) Epistasis in intra- and inter-gene pool crosses of the common bean. *Genet. Mol. Res.* **15**, 15017573.

Borrill, P., Harrington, S.A., Simmonds, J. and Uauy, C. (2019) Identification of transcription factors regulating senescence in wheat through gene regulatory network modelling. *Plant Physiol.* **180**, 1740–1755.

Boutet, G., Lavaud, C., Lesné, A., Miteul, H., Pilet-Nayel, M.-L., Andrivon, D., Lejeune-Hénaut, I. *et al.* (2023) Five regions of the pea genome co-control partial resistance to *D. pinodes*, tolerance to frost, and some architectural or phenological traits. *Genes*, **14**, 1399.

Breen, M.S., Kemena, C., Vlasov, P.K., Notredame, C. and Kondrashov, F.A. (2012) Epistasis as the primary factor in molecular evolution. *Nature*, **490**, 535–538.

Carré, C., Carluer, J.B., Chaux, C., Estoup-Streiff, C., Roche, N., Hosy, E., Mas, A. *et al*. (2024) Next-Gen GWAS: full 2D epistatic interaction maps retrieve part of missing heritability and improve phenotypic prediction. *Genome Biol.* **25**, 76.

Chanchu, T., Somta, P., Yimram, T., Laosatit, K., Kaga, A. and Srinives, P. (2023) Antagonistic pleiotropy of *Ln* gene controls trade-off between 2-seeded pods and 4-seeded pods in soybean. *Euphytica*, **219**, 87.

Chang, Y.-C., Wu, J.-T., Hong, M.-Y., Tung, Y.-A., Hsieh, P.-H., Yee, S.W., Giacomini, K.M. *et al*. (2020) GenEpi gene-based epistasis discovery using machine learning. *BMC Bioinform.* **21**, 68.

Chantepie, S. and Chevin, L.-M. (2020) How does the strength of selection influence genetic correlations? *Evol. Lett.* **4**, 468–478.

Chebib, J. and Guillaume, F. (2021) Pleiotropy or linkage? Their relative contributions to the genetic correlation of quantitative traits and detection by multitrait GWA studies. *Genetics*, **219**, iyab159.

Chen, Y. and Lübberstedt, T. (2010) Molecular basis of trait correlations. *Trends Plant Sci.* **15**, 454–461.

Chen, J.-Y., Guo, L., Ma, H., Chen, Y.-Y., Zhang, H.-W., Ying, J.-Z. and Zhuang, J.Y. (2014) Fine mapping of qHd1, a minor heading date QTL with pleiotropism for yield traits in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **127**, 2515–2524.

Chen, Y., Ren, X., Zheng, Y., Zhou, X., Huang, L., Yan, L., Jiao, Y. *et al*. (2017) Genetic mapping of yield traits using RIL population derived from Fuchuan Dahuasheng and ICG6375 of peanut (*Arachis hypogaea* L.). *Mol. Breed.* **37**, 17.

Chen, W., Wu, Y., Fritschi, F.B. and Juenger, T.E. (2021) The genetic basis of the root economics spectrum in a perennial grass. *Proc. Natl. Acad. Sci. USA*, **118**, e2107541118.

Chen, Q., Guo, Z., Shi, X., Wei, M., Fan, Y., Zhu, J., Zheng, T. *et al*. (2022) Increasing the grain yield and grain protein content of common wheat (*Triticum aestivum*) by introducing missense mutations in the *Q* gene. *Int. J. Mol. Sci.* **23**, 10772.

Chen, Y., Wang, W., Yang, Z., Peng, H., Ni, Z., Sun, Q. and Guo, W. (2024) Innovative computational tools provide new insights into the polyploid wheat genome. *aBIOTECH*, **5**, 52–70.

Chung, D., Yang, C., Li, C., Gelernter, J. and Zhao, H. (2014) GPA: A statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation. *PLoS Genet.* **10**, e1004787.

Cooper, M., van Eeuwijk, F.A., Hammer, G.L., Podlich, D.W. and Messina, C. (2009) Modeling QTL for complex traits: detection and context for plant breeding. *Curr. Opin. Plant Biol.* **12**, 231–240.

Costa, W.G., Celeri, M.O., Barbosa, I.P., Silva, G.N., Azevedo, C.F., Borem, A., Nascimento, M. *et al*. (2022) Genomic prediction through machine learning and neural networks for traits with epistasis. *Comput. Struct. Biotechnol. J.* **20**, 5490–5499.

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., Burgueño, J. *et al*. (2017) Genomic selection in plant breeding: Methods, models, and perspectives. *Trends Plant Sci.* **22**, 961–975.

Danilevicz, M.F., Bayer, P.E., Boussaid, F., Bennamoun, M. and Edwards, D. (2021a) Maize yield prediction at an early developmental stage using multispectral images and genotype data for preliminary hybrid selection. *Remote Sens.* **13**, 3976.

Danilevicz, M.F., Bayer, P.E., Nestor, B.J., Bennamoun, M. and Edwards, D. (2021b) Resources for image-based high-throughput phenotyping in crops and data sharing challenges. *Plant Physiol.* **187**, 699–715.

David, I., Elsen, J.-M. and Concordet, D. (2013) CLIP Test: a new fast, simple and powerful method to distinguish between linked or pleiotropic quantitative trait loci in linkage disequilibria analysis. *Hereidty*, **110**, 232–238.

Derbyshire, M.C., Khentry, Y., Severn-Ellis, A., Mwape, V., Saad, N.S.M., Newman, T.E., Taiwo, A. *et al*. (2021) Modeling first order additive × additive epistasis improves accuracy of genomic prediction for sclerotinia stem rot resistance in canola. *Plant Genome*, **14**, e20088.

Desgroux, A., Baudais, V.N., Aubert, V., Roy, G.L., de Larambergue, H., Miteul, H., Aubert, G. *et al*. (2018) Comparative genome-wide-association mapping identifies common loci controlling root system architecture and resistance to *Aphanomyces euteiches* in pea. *Front. Plant Sci.* **8**, 2195.

Desta, Z.A. and Ortiz, R. (2014) Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* **19**, 592–601.

Dong, Z., Wang, Y., Bao, J., Li, Y., Yin, Z., Long, Y. and Wan, X. (2023) The genetic structures and molecular mechanisms underlying ear traits in maize (*Zea mays* L.). *Cells*, **12**, 1900.

Dwivedi, S.L., Reynolds, M.P. and Ortiz, R. (2021) Mitigating tradeoffs in plant breeding. *iScience*, **24**, 102965.

Dwivedi, S.L., Heslop-Harrison, P., Spillane, C., McKeown, P.C., Edwards, D., Goldman, I. and Ortiz, R. (2023) Evolutionary dynamics and adaptive benefits of deleterious mutations in crop gene pools. *Trends Plant Sci.* **28**, 685–697.

Edwards, D. and Batley, J. (2004) Plant bioinformatics: From genome to phenome. *Trends Biotechnol.* **22**, 232–237.

Edwards, D. and Batley, J. (2022) Graph pangenomes find missing heritability. *Nat. Genet.* **54**, 919–920.

Eguchi, Y., Bilolikar, G. and Geiler-Samerotte, K. (2019) Why and how to study genetic changes with context-dependent effects. *Curr. Opin. Genet. Dev.* **58-59**, 95–102.

Fan, X., Cui, F., Ji, J., Zhang, W., Zhao, X., Liu, J., Meng, D. *et al*. (2019) Dissection of pleiotropic QTL regions controlling wheat spike characteristics under different nitrogen treatments using traditional and conditional QTL mapping. *Front. Plant Sci.* **10**, 187.

Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R., Hu, G. *et al*. (2017) Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol.* **18**, 161.

Fei, J., Lu, J., Jiang, Q., Liu, Z., Yao, D., Qu, J., Liu, S. *et al*. (2022) Maize plant architecture trait QTL mapping and candidate gene identification based on multiple environments and double populations. *BMC Plant Biol.* **22**, 110.

Fernandes, S.B. and Lipka, A.E. (2020) simplePHENOTYPES: SIMulation of pleiotropic, linked and epistatic phenotypes. *BMC Bioinform.* **21**, 491.

Fernandes, S.B., Zhang, K.S., Jamann, T.M. and Lipka, A.E. (2021) How well can multivariate and univariate GWAS distinguish between true and spurious pleiotropy? *Front. Plant Sci.* **11**, 602526.

Fernandes, S.B., Casstevens, T.M., Bradbury, P.J. and Lipka, A.E. (2022) A multi-trait multi-locus stepwise approach for conducting GWAS on correlated traits. *Plant Genome*, **15**, e20200.

Flint, J. and Mackay, T.F.C. (2009) Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res.* **19**, 723–733.

Fradgley, N.S., Garner, K., Kerton, M., Swarbręck, S.M. and Bentley, A.R. (2022) Trade-offs in the genetic control of functional and nutritional quality traits in UK winter wheat. *Heredity*, **128**, 420–433.

Fradgley, N., Gardner, K.A., Bentley, A.R., Howell, P., Mackay, I.J., Scott, M.F., Mott, R. *et al*. (2023) Multi-trait ensemble genomic prediction and simulations of recurrent selection highlight importance of complex trait genetic architecture for long-term genetic gains in wheat. *In silico Plants*, **5**, diad002.

Gao, M., Hao, Z., Ning, Y. and He, Z. (2024) Revisiting growth–defence trade-offs and breeding strategies in crops. *Plant Biotechnol. J.* **22**, 1198–1205.

Garland, T., Jr. (2014) Trade-offs. *Curr. Biol.* **24**, R60–R61.

Goettel, W., Zhang, H., Li, Y., Qiao, Z., Jiang, H., Hou, D., Song, Q. *et al*. (2022) *POWR1* is a domestication gene pleiotropically regulating seed quality and yield in soybean. *Nat. Commun.* **13**, 3051.

Gould, F.W. (1975) *The Grasses of Texas*. College Station, TX: Texas A&M University Press.

Guo, Y. and Hong, D. (2010) Novel pleiotropic loci controlling panicle architecture across environments in japonica rice (*Oryza sativa* L.). *J. Genet. Genomics*, **37**, 533–544.

Guo, Y., Zhong, Z., Yang, C., Hu, J., Jiang, Y., Liang, Z., Gao, H. *et al*. (2019) Epi-GTBN: an approach of epistasis mining based on genetic Tabu algorithm and Bayesian network. *BMC Bioinform.* **20**, 444.

Hall, M.C., Baster, C.J. and Willis, J.H. (2006) Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. *Genetics*, **172**, 1829–1844.

Hamdan, N., Lee, C.H., Wong, S.L., Fauzi, C.E.N.C.A., Zamri, N.M.A. and Lee, T.H. (2022) Prevention of enzymatic browning by natural extracts and genome-editing: a review on recent progress. *Molecules*, **27**, 1101.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J. and Goddard, M.E. (2009) Invited review: Genomic selection in dairy cattle: progress and challenges. *J. Dairy Sci.* **92**, 433–443.

He, N., Zhan, G., Huang, F., Abou-Elwafa, S.F. and Yang, D. (2022) Fine mapping and cloning of a major QTL *qph12*, which simultaneously affects the plant height, panicle length, spikelet number and yield in rice (*Oryza sativa* L.). *Front. Plant Sci.* **13**, 878558.

He, T., Angessa, T.T. and Li, C. (2023a) Pleiotropy structures plant height and seed weight scaling in barley despite long history of domestication and breeding selection. *Plant Phenomics*, **5**, 15.

He, X.-G., Li, Z.-R., Guo, S.-C., Zheng, X.-F., Liu, C.-H., Liu, Z.-J., Li, Y.-X. *et al.* (2023b) Epistasis-aware genome-wide association studies provide insights into high-yield and high-quality rice efficient breeding. *J. Integr. Agric.* https://doi.org/10.1016/j.jia.2023.07.021

Hellens, R.P., Moreau, C., Lin-Wang, K., Schwinn, K.E., Thomson, S.J., Fiers, M.W., Frew, T.J. *et al.* (2010) Identification of Mendel's white flower character. *PLoS One*, **5**, e13230.

Hendelman, A., Zebell, S., Rodriguez-Leal, D., Dukler, N., Robitaille, G., Wu, X., Kostyun, X. *et al.* (2021) Conserved pleiotropy of an ancient plant homeobox gene uncovered by *cis*-regulatory dissection. *Cell*, **184**, 1724–1739.

Holland, J.B. (2001) Epistasis and plant breeding. *Plant Breed. Rev.* **1**, 27–92.

Holland, P.W.H., Marlétaz, F., Maeso, I., Dunwell, T.L. and Paps, J. (2017) New genes from old: asymmetric divergence of gene duplicates and the evolution of development. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **372**, 20150480.

Hu, X., Zhao, J., Lin, Z., Wang, Y., Peng, H., Zhao, H., Wan, X. *et al.* (2022) Mendelian randomization for causal inference accounting for pleiotropy and sample structure using genome-wide summary statistics. *Proc. Natl. Acad. Sci. USA*, **119**, e2106858119.

Hughes, R.K., Desforges, N., Selwood, C., Smith, R., Speirs, C.I., Sinnaeve, G., Gorton, P.G. *et al.* (2001) Genes affecting starch biosynthesis exert pleiotropic effects on the protein content and composition of pea seeds. *J. Sci. Food Agric.* **81**, 877–882.

Hui, J., Bai, H., Lyu, X., Ma, S., Chen, X. and Li, S. (2023) A pleiotropic QTL increased economic water use efficiency in bread wheat (*Triticum aestivum* L.). *Front. Plant Sci.* **13**, 1067590.

Ilgokina, A.A., Meshcheryakov, G., Gretsova, M.V., Nuzhdin, S.V. and Samsonova, M.G. (2020) Multi-trait multi-locus SEM model discriminates SNPs of different effects. *BMC Genomics*, **21**(Suppl 8), 490.

Jiang, C. and Zeng, Z.-B. (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics*, **140**, 1111–1127.

Jiang, L., Zhong, H., Jiang, X., Zhang, J., Huang, R., Liao, F., Deng, Y. *et al.* (2022) Identification and pleiotropic effect analysis of *GSE5* on rice chalkiness and grain shape. *Front. Plant Sci.* **12**, 814928.

Kearsey, M.J. and Pooni, H.S. (1996) *The Genetic Analysis of Quantitative Traits*, 1st edn. London: Chapman and Hall. p. 46.

Khaipho-Burch, M., Ferebee, T., Giri, A., Ramstein, G., Monier, B., Yi, E., Romay, M.C. *et al.* (2023) Elucidating the patterns of pleiotropy and its biological relevance in maize. *PLoS Genet.* **19**, e1010664.

Khasanova, A., Lovell, J.T., Bonnette, J., Weng, X., Jenkins, J., Yoshinaga, Y., Schmutz, J. *et al.* (2019) The genetic architecture of shoot and root trait divergence between mesic and xeric ecotypes of a perennial grass. *Front. Plant Sci.* **10**, 366.

Kim, J., Bates, D.G., Postlethwaite, I., Heslop-Harrison, P. and Cho, K.H. (2008) Linear time-varying models can reveal non-linear interactions of biomolecular regulatory networks using multiple time-series data. *Bioinformatics*, **24**, 1286–1292.

Kim, K.H., Kim, J.-Y., Lim, W.-J., Jeong, S., Lee, H.-Y., Cho, Y., Moon, J.K. *et al.* (2020) Genome-wide association and epistatic interactions of flowering time in soybean cultivar. *PLoS One*, **15**, e0228114.

Knight, C.G., Azevedo, R.B.R. and Leroi, A.M. (2001) Testing life-history pleiotropy in Caenorhabditis elegans. *Evolution*, **55**, 1795–1804.

Kodackattumannil, P., Lekshmi, G., Kottackal, M., Sasi, S., Krishnan, S., Senaani, S.A. and Amiri, K.M.A. (2023) Hidden pleiotropy of agronomic traits uncovered by CRISPR-Cas9 mutagenesis of the tyrosinase CuA-binding domain of the polyphenol oxidase 2 of eggplant. *Plant Cell Rep.* **42**, 825–828.

Kou, K., Su, T., Wang, Y., Yang, H., Du, H., He, M., Li, T. *et al.* (2021) Natural variation of the *Dt2* promoter controls plant height and node number in semi-determinant soybean. *Mol. Breed.* **41**, 40.

Krasileva, K.V., Vasquez-Gross, H.A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J. *et al.* (2017) Uncovering hidden variation in polyploid wheat. *Proc. Natl. Acad. Sci. USA*, **114**, E913–E921.

Kuczyńska, A., Mikołajczak, K. and Ćwiek, H. (2014) Pleiotropic effects of the sdw1 locus in barley populations representing different rounds of recombination. *Electron. J. Biotechnol.* **17**, 217–223.

Kumar, S., Hash, C.T., Nepolean, T., Satyavathi, C.T., Singh, G., Mahendrakumar, M.D., Yadav, R.S. *et al.* (2017) Mapping QTLs controlling flowering time and important agronomic traits in pearl millet. *Front. Plant Sci.* **8**, 1731.

Kumar, N., Boatwright, J.L., Boyles, R.E., Brenton, Z.W. and Kresovich, S. (2024) Identification of pleiotropic loci mediating structural and non-structural carbohydrate accumulation within the sorghum bioenergy association panel using high-throughput markers. *Front. Plant Sci.* **15**, 1356619.

Lange, T.M., Heinrich, F., Kopisch-Obuch, F., Keunecke, H., Gültas, M. and Schmitt, A.O. (2023) Improving genomic prediction of rhizomania resistance in sugar beet (*Beta vulgaris* L.) by implementing epistatic effects and feature selection. *F1000Research*, **12**, 280.

LeBlanc, M., Zuber, V., Thompson, W.K., Andreassen, O.A., Schizophrenia and Bipolar Disorder Working Groups of the Psychiatric Genomics Consortium, Frigessi, A. and Andreassen, B.K. (2018) A correction for sample overlap in genome-wide association studies in a polygenic pleiotropy-informed framework. *BMC Genomics*, **19**, 494.

Lee, C.H., Shi, H., Pasaniuc, B., Eskin, E. and Han, B. (2021) PLEIO: a method to map and interpret pleiotropic loci with GWAS summary statistics. *Am. J. Hum. Genet.* **108**, 36–48.

Lee, S.Y., Jeung, J.-U. and Mo, Y. (2024) Allelic combinations of *Hd1*, *Hd16*, and *Ghd7* exhibit pleiotropic effects on agronomic traits in rice. *G3*, **14**, jkad300.

Li, F., Xie, J., Zhu, X., Wang, X., Zhao, Y., Ma, X., Zhang, Z. *et al.* (2018a) Genetic basis underlying correlations among growth duration and yield traits revealed by GWAS in rice (*Oryza sativa* L.). *Front. Plant Sci.* **9**, 650.

Li, X., Tian, R., Kamala, S., Du, H., Li, W., Kong, Y. and Zhang, C. (2018b) Identification and verification of pleiotropic QTL controlling multiple amino acid contents in soybean seed. *Euphytica*, **214**, 93.

Li, H., Jiang, S., Li, C., Liu, L., Lin, Z., He, H., Deng, X.-W. *et al.* (2020) The hybrid protein interactome contributes to rice heterosis as epistatic effects. *Plant J.* **102**, 116–128.

Liller, C.B., Neuhaus, R., von Korff, M., Koornneef, M. and van Esse, W. (2015) Mutations in barley row type genes have pleiotropic effects on shoot branching. *PLoS One*, **10**, e0140246.

Lisker, A., Maurer, A., Schmutzer, T., Kazman, E., Cōster, H., Holzapfel, J., Ebmeyer, E. *et al.* (2022) A haplotype-based GWAS identified trait-improving QTL alleles controlling agronomic traits under contrasting nitrogen fertilization treatments in the MAGIC wheat population WM-800. *Plants*, **11**, 3508.

Liu, R., Jia, H., Cao, X., Huang, J., Li, F., Tao, Y., Qiu, F. *et al.* (2012) Fine mapping and candidate gene prediction of a pleiotropic quantitative trait locus for yield-related trait in *Zea mays*. *PLoS One*, **7**, e49836.

Liu, T., Liu, H., Zhang, H. and Xing, Y. (2013) Validation and characterization of *Ghd7.1*, a major quantitative trait locus with pleiotropic effects on spikelets per panicle, plant height, and heading date in rice (*Oryza sativa* L.). *J. Integr. Plant Biol.* **55**, 917–927.

Liu, D., Zhao, D.-H., Zeng, J.-Q., Shawai, R.H., Tong, J.-Y., Li, M., Li, F. *et al.* (2023) Identification of genetic loci for grain yield-related traits in the wheat population Zhongmai 578/Jimai 22. *J. Integr. Agric.* **22**, 1985–1999.

Lo, S., Muñoz-Amatriaín, M., Hokin, S.A., Cisse, N., Roberts, P.A., Farmer, A.D., Xu, S. *et al.* (2019) A genome-wide association and meta-analysis reveal regions associated with seed size in cowpea [*Vigna unguiculata* (L.) Walp]. *Theor. Appl. Genet.* **132**, 3079–3087.

Luo, X., Ding, Y., Zhang, L., Yue, Y., Snyder, J.H., Ma, C. and Zhu, J. (2017a) Genomic prediction of genotypic effects with epistasis and environment interactions for yield-related traits of rapeseed (*Brassica napus* L.). *Front. Genet.* **8**, 15.

Luo, Z., Wang, M., Long, Y., Huang, Y., Shi, L., Zhang, C., Liu, X. *et al.* (2017b) Incorporating pleiotropic quantitative trait loci in dissection of complex traits: seed yield in rapeseed as an example. *Theor. Appl. Genet.* **130**, 1569–1585.

Lyu, X., Li, Y.-H., Li, Y., Li, D., Han, C., Hong, H., Tian, Y. *et al.* (2023) The domestication associated *L1* gene encodes a eucomic acid synthase pleiotropically modulating pod pigmentation and shattering in soybean. *Mol. Plant*, **16**, 1178–1191.

Maina, F., Harou, A., Hamidou, F. and Morris, G.P. (2022) Genome-wide association studies identify putative pleiotropic locus mediating drought tolerance in sorghum. *Plant Direct*, **6**, e413.

Maioli, A., Gianoglio, S., Moglia, A., Acquardo, A., Valentino, D., Milani, A.M., Prohens, J. *et al*. (2020) Simultaneous CRISPR/Cas-9-editing of three PPO genes reduces fruit flesh browning in *Solanum melongena* L. *Front. Plant Sci.* **11**, 607161.

Maldonado-Taipe, N., Barbier, F., Schmid, K., Jung, C. and Emrani, N. (2022) High-density mapping of quantitative trait loci controlling agronomically important traits in quinoa (*Chenopodium quinoa* Willd.). *Front. Plant Sci.* **13**, 916067.

Malik, P., Kumar, J., Singh, S., Sharma, S., Meher, P.K., Sharma, M.K., Roy, J.K. *et al*. (2021) Single-trait, multi-locus and multi-trait GWAS using four different models for yield traits in bread wheat. *Mol. Breed.* **41**, 46.

Malik, P., Kumar, J., Sharma, S., Meher, P.K., Balyan, H.S., Gupta, P.K. and Sharma, S. (2022) GWAS for main effects and epistatic interactions for grain morphology traits in wheat. *Physiol. Mol. Biol.* **28**, 651–668.

Mather, K. and Jinks, J.L. (1982) *Biometrical Genetics*, 3rd edn. London: Chapman and Hall Ltd.

McGaugh, S.E., Lorenz, A.J. and Flagel, L.E. (2021) The utility of genomic prediction models in evolutionary genetics. *Proc. R. Soc. B*, **288**, 20210693.

de Melo, R.C., Coimbra, J.L.M., Corrêa da Silva, J.G., Cerruti, P.H., Carbonari, L.T.S. and Guidolin, A.F. (2023) Implications of additive × additive epistasis for common bean breeding. *Research*, **220**, 68.

Mendel, G. (1866) *Versuche über Plflanzenhybriden. Verhandlungen des naturforschenden Vereines in Brünn*, Bd. IV für das Jahr 1865, Abhandlungen, 3–47.

Montenegro, J.D., Golicz, A.A., Bayer, P.E., Hurgobin, B., Lee, H., Chan, C.K.K., Visendi, P. *et al*. (2017) The pangenome of hexaploid bread wheat. *Plant J.* **90**, 1007–1013.

Mural, R.V., Grzybowski, M., Miao, C., Damke, A., Sapkota, S., Boyles, R.E., Salas Fernandez, M.G. *et al*. (2021) Meta-analysis identifies pleiotropic loci controlling phenotypic trade-offs in sorghum. *Genetics*, **218**, iyab087.

Ndlovu, N., Kachapur, R.M., Beyene, Y., Das, B., Ogugo, V., Makumbi, D., Spillane, C. *et al*. (2024) Linkage mapping and genomic prediction of grain quality traits in tropical maize (*Zea mays* L.). *Front. Genet.* **15**, 1353289.

Nelson, R.L. (1988) Response to selection for time of flowering in soybean. *Crop Sci.* **28**, 623–626.

Olatoye, M.O., Hu, Z. and Aikpokpodion, P.O. (2019) Epistasis detection and modeling for genomic selection in cowpea (*Vigna unguiculata* L. Walp.). *Front. Genet.* **10**, 677.

Ookawa, T., Hobo, T., Yano, M., Murata, K., Ando, T., Miura, H., Asano, K. *et al*. (2010) New approach for rice improvement using a pleiotropic QTL gene for lodging resistance and yield. *Nat. Commun.* **1**, 132.

Otyama, P.I., Chamberlin, K., Ozias-Akins, P., Graham, M.A., Cannon, E.K.S., Cannon, S.B., MacDonald, G.E. *et al*. (2022) Genome-wide approaches delineate the additive, epistatic, and pleiotropic nature of variants controlling fatty acid composition in peanut (*Arachis hypogaea* L.). *G3*, **12**, jkab382.

Paaby, A.B. and Rockman, M.V. (2013) The many faces of pleiotropy. *Trends Genet.* **29**, 66–73.

Pavlicev, M., Norgard, E.A., Fawcett, G.L. and Cheverud, J.M. (2011) Evolution of pleiotropy: epistatic interaction pattern supports a mechanistic model underlying variation in genotype–phenotype map. *J. Exp. Zool.* **316**, 371–385.

Phillips, P.C. (1998) The language of gene interaction. *Genetics*, **149**, 1167–1171.

Phillips, P.C. (2008) Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867.

Ponte-Fernandez, C., Gonzalez-Dominguez, J., Carvajal-Rodriguez, A. and Martin, M.J. (2022) Evaluation of existing methods for high-order epistasis detection. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **19**, 912–926.

Powell, O.M., Voss-Fels, K.P., Jordan, D.R., Hammer, G. and Cooper, M. (2021) Perspectives on applications of hierarchical gene-to-phenotype (G2P) maps to capture non-stationary effects of alleles in genomic prediction. *Front. Plant Sci.* **12**, 663565.

Raffo, M.A., Sarup, P., Guo, X., Liu, H., Andersen, J.R., Orabi, J., Jahoor, A. *et al*. (2022) Improvement of genomic prediction in advanced wheat breeding lines by including additive-by-additive epistasis. *Theor. Appl. Genet.* **135**, 965–978.

Raherison, E., Majidi, M.M., Goessen, R., Hughes, N., Cuthbert, R., Knox, R. and Lukens, L. (2020) Evidence for the accumulation of nonsynonymous mutations and favorable pleiotropic alleles during wheat breeding. *G3* **10**, 4001–4011.

Rashid, M.A.R., Zhao, Y., Azeem, F., Zhao, Y., Muhu-Din Ahmed, H.G., Atif, R.M., Pan, Y. *et al*. (2022) Unveiling the genetic architecture for lodging resistance in rice (*Oryza sativa*. L) by genome-wide association analyses. *Front. Genet.* **13**, 960007.

Rasmusson, D.C. and Phillips, R.L. (1997) Plant breeding progress and genetic diversity from de novo variation and elevated epistasis. *Crop Sci.* **37**, 303–310.

Rathan, N.D., Krishnappa, G., Singh, A.M. and Govindan, V. (2023) Mapping QTL for phenological and grain-related traits in a mapping population derived from high-zinc-biofortified wheat. *Plan. Theory*, **12**, 220.

Razzaque, S., Heckman, R.W. and Juenger, T.E. (2023) Seed size variation impacts local adaptation and life-history strategies in a perennial grass. *Proc. R. Soc. Biol.* **290**, 20222460.

Rosas, J.E., Escobar, M., Martínez, S., Blanco, P., Pérez, F., Quero, G., Gutiérrez, L. *et al*. (2020) Epistasis and quantitative resistance to *Pyricularia oryzae* revealed by GWAS in advanced rice breeding populations. *Agriculture*, **10**, 622.

Roth, M., Beugnot, A., Mary-Huard, T., Moreau, L., Charcosset, A. and Fiévet, J.B. (2022) Improving genomic predictions with inbreeding and nonadditive effects in two admixed maize hybrid populations in single and multienvironment contexts. *Genetics*, **220**, iyac018.

Saha, S., Perrin, L., Röder, L., Brun, C. and Spinelli, L. (2022) Epi-MEIF: detecting higher order epistatic interactions for complex traits using mixed effect conditional inference forests. *Nucleic Acids Res.* **50**, e114.

Sakai, T., Abe, A., Shimizu, M. and Terauchi, R. (2021) RIL-StEp: epistasis analysis of rice recombinant inbred lines reveals candidate interacting genes that control seed hull color and leaf chlorophyll content. *G3*, **11**, jkab130.

Sang, Z., Wang, H., Yang, Y., Zhang, Z., Liu, X., Li, Z. and Xu, Y. (2022) Epistasis activation contributes substantially to heterosis in temperate by tropical maize hybrids. *Front. Plant Sci.* **13**, 921608.

Sangster, T.A., Lindquist, S. and Queitsch, C. (2004) Under cover: causes, effects and implications of hsp90-mediated genetic capacitance. *Bioessays*, **26**, 348–362.

Sannemann, W., Lisker, A., Maurer, A., León, J., Kazman, E., Cöster, H., Holzapfel, J. *et al*. (2018) Adaptive selection of founder segments and epistatic control of plant height in the MAGIC winter wheat population WM-800. *BMC Genomics*, **19**, 559.

Scarcelli, N., Cheverud, J.M., Schaal, B.A. and Kove, P.X. (2007) Antagonistic pleiotropic effects reduce the potential adaptive value of the FRIGIDA locus. *Proc. Natl. Acad. Sci. USA*, **104**, 16986–16991.

Schaid, D.J., Tong, X., Larrabee, B., Kennedy, R.B., Poland, G.A. and Sinnwell, J.P. (2016) Statistical methods for testing genetic pleiotropy. *Genetics*, **204**, 483–497.

Schaid, D.J., Tong, X., Batzler, A., Sinnwell, J.P., Qing, J. and Biernacka, J.M. (2019) Multivariate generalized linear model for genetic pleiotropy. *Biostatistics*, **20**, 111–128.

Scheben, A., Wolter, F., Batley, J., Puchta, H. and Edwards, D. (2017) Towards CRISPR/CAS crops—bringing together genomics and genome editing. *New Phytol.* **216**, 682–698.

Schulthess, A.W., Reif, J.C., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V. *et al*. (2017) The roles of pleiotropy and close linkage as revealed by association mapping of yield and correlated traits of wheat (*Triticum aestivum* L.). *J. Exp. Bot.* **68**, 4089–4101.

Sedivy, E.J., Akpertey, A., Vela, A., Abadir, A., Khan, A. and Hanzawa, Y. (2020) Identification of non-pleiotropic loci in flowering and maturity control in soybean. *Agronomy*, **10**, 1204.

Shang, Q., Zhang, D., Li, R., Wang, K., Cheng, Z., Zhou, Z., Hao, Z. *et al*. (2020) Mapping quantitative trait loci associated with stem-related traits in maize (*Zea mays* L.). *Plant Mol. Biol.* **104**, 583–595.

Singh, J., Gezan, S.A. and Vallejos, C.E. (2019) Developmental pleiotropy shaped the roots of the domesticated common bean (*Phaseolus vulgaris*). *Plant Physiol.* **180**, 1467–1479.

Slim, L., Chatelain, C., Azencott, C.-A. and Vert, J.P. (2020) Novel methods for epistasis detection in genome-wide association studies. *PLoS One*, **15**, e0242927.

Smith, S.D. (2016) Pleiotropy and the evolution of floral integration. *New Phytol.* **209**, 80–85.

Smith, J.R. and Nelson, R.L. (1986) Selection for seed-filling period in soybean. *Crop Sci.* **26**, 466–4769.

Smith, C.J., Sinnott-Armstrong, N., Cichońska, A., Julkunen, H., Fauman, E.B., Würtz, P. and Pritchard, J.K. (2022) Integrative analysis of metabolite GWAS illuminates the molecular basis of pleiotropy and genetic correlation. *elife*, **11**, e79348.

Sofi, P.A., Rather, A.G. and Warsi, M.Z.K. (2007) Implications of epistasis in maize breeding. *Int. J. Plant Breed. Genet.* **1**, 1–11.

Song, X., Meng, X., Guo, H., Cheng, Q., Jing, Y., Chen, M., Liu, G. *et al.* (2022) Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. *Nat. Biotechnol.* **40**, 1403–1411.

Soyk, S., Lemmon, Z.H., Oved, M., Fisher, J., Liberatore, K.L., Park, S.J., Goren, A. *et al.* (2017) Bypassing negative epistasis on yield in tomato imposed by a domestication gene. *Cell*, **169**, 1142–1155.

Stearns, F.W. (2010) One hundred years of pleiotropy: a retrospective. *Genetics*, **186**, 767–773.

Sun, B., Zhan, X.-D., Lin, Z.-C., Wu, W.-X., Yu, P., Zhang, Y.-X., Sun, L.P. *et al.* (2017) Fine mapping and candidate gene analysis of qHD5, a novel major QTL with pleiotropism for yield-related traits in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **130**, 247–258.

Swinnen, G., Goossens, A. and Pauwels, L. (2016) Lessons from domestication: targeting cis-regulatory elements for crop improvement. *Trends Plant Sci.* **21**, 506–515.

Takai, T., Adachi, S., Taguchi-Shiobara, F., Sanoh-Arai, Y., Iwasawa, N., Yoshinaga, S., Hirose, S. *et al.* (2013) A natural variant of NAL1, selected in high-yield rice breeding programs, pleiotropically increases photosynthesis rate. *Sci. Rep.* **3**, 2149.

Tao, Y., Mace, E.S., Tai, S., Cruickshank, A., Campbell, B.C., Zhao, X., Van Oosterom, E.J. *et al.* (2017) Whole-genome analysis of candidate genes associated with seed size and weight in *Sorghum bicolor* reveals signatures of artificial selection and insights into parallel domestication in cereal crops. *Front. Plant Sci.* **8**, 1237.

Taylor, M.B. and Ehrenreich, I.M. (2015) Higher-order genetic interactions and their contribution to complex traits. *Trends Genet.* **31**, 34–40.

Tian, Z., Wang, X., Dun, X., Zhao, K., Wang, H. and Ren, L. (2024) An integrated QTL mapping and transcriptome sequencing provides further molecular insights and candidate genes for stem strength in rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* **137**, 38.

Tsarev, A., Plugatar, Y., Tsareva, R., Tsarev, V. and Laur, N. (2021) *Promising introduced black cottonwood species for bioenergy and forage production. IOP Conference. Series: Earth and Environmental Science*, 875, 012088.

Tyler, A.L., Asselbergs, F.W., Williams, S.M. and Moore, J.H. (2009) Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. *Bioessays*, **31**, 220–227.

Tyler, A.L., Lu, W., Hendrick, J.J., Philip, V.M. and Carter, G.W. (2013) CAPE: an R package for combined analysis of pleiotropy and epistasis. *PLoS Comput. Biol.* **9**, e1003270.

Tyler, A.L., Emerson, J., Kassaby, E.L., Wells, A.E., Philip, V.M. and Carter, G.W. (2021) The combined analysis of pleiotropy and epistasis (CAPE). *Methods Mol. Biol.* **2212**, 55–67.

Udall, J.A., Quijada, P.A., Lambert, B. and Osborn, T.S. (2006) Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 2. Identification of alleles from unadapted germplasm. *Theor. Appl. Genet.* **113**, 597–609.

Upadhyaya, H.D. and Nigam, S. (1998) Epistasis for vegetative and reproductive traits in peanut. *Crop Sci.* **38**, 44–49.

Upadhyaya, H.D. and Nigam, S.N. (1999) Detection of epistasis for protein and oil contents and oil quality parameters in peanut. *Crop Sci.* **39**, 115–118.

Vaughan, D.A., Balazs, E. and Heslop-Harrison, J.S. (2007) From crop domestication to super-domestication. *Ann. Bot.* **100**, 893–901.

Viana, J.M.S. (2023) The impact of epistasis in the heterosis and combining ability analyses. *Front. Plant Sci.* **14**, 1168419.

Vishnukiran, T., Neeraja, C.N., Jaldhani, V., Vijayalakshmi, P., Rao, P.R., Surahmanyam, D. and Voleti, S.R. (2020) A major pleiotropic QTL identified for yield components and nitrogen content in rice (*Oryza sativa* L.) under differential nitrogen field conditions. *PLoS One*, **15**, e0240854.

de Visser, J.A.G.M., Cooper, T.F. and Elena, S.F. (2011) The causes of epistasis. *Proc. R. Soc. B*, **278**, 3617–3624.

Vivek, M.P., Tyler, A.L. and Carter, G.W. (2014) Dissection of complex gene expression using the combined analysis of pleiotropy and epistasis. *Pac. Symp. Biocomput.* **2014**, 200–211.

Vojgani, E., Pook, T., Martini, J.W.R., Lker, A.C.H., Mayer, M., Schön, C.C. and Simianer, H. (2021a) Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments. *Theor. Appl. Genet.* **134**, 913–2930.

Vojgani, E., Pook, T. and Simianer, H. (2021b) Phenotype prediction under epistasis. *Methods Mol. Biol.* **2212**, 105–120.

Vojgani, E., Hölker, A.C., Mayer, M., Schön, C.C., Simianer, H. and Pook, T. (2023) Genomic prediction using information across years with epistatic models and dimension reduction via haplotype blocks. *PLoS One*, **18**, e0282288.

Von Berg, J., Dam, M.T., van der Laan, S.W. and de Ridder, J. (2022) PolarMorphism enables discovery of shared genetic variants across multiple traits from GWAS summary statistics. *Bioinformatics*, **38**(Suppl 1), i212–i219.

Wang, Z., Liao, B.-Y. and Zhang, J. (2010) Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl. Acad. Sci. USA*, **107**, 18034–18039.

Wang, D., El-Basyoni, S., Baenziger, P.S., Crossa, J., Eskridge, K.M. and Dweikat, I. (2012) Prediction of genetic values of quantitative traits with epistatic effects in plant breeding populations. *Heredity*, **109**, 313–319.

Wang, M., Lu, X., Xu, G., Yin, X., Cui, Y., Huang, L., Rocha, P.S.C.F. *et al.* (2016) *OsSGL*, a novel pleiotropic stress-related gene enhances grain length and yield in rice. *Sci. Rep.* **6**, 38157.

Wang, P., He, S., Sun, G., Pan, Z., Sun, J., Geng, X., Peng, Z. *et al.* (2021) Favorable pleiotropic loci for fiber yield and quality in upland cotton (*Gossypium hirsutum*). *Sci. Rep.* **11**, 15935.

Wang, F., Sun, X., Liu, B., Kong, F., Pan, X. and Zhang, H. (2022a) A polygalacturonase gene *PG031* regulates seed coat permeability with a pleiotropic effect on seed weight in soybean. *Theor. Appl. Genet.* **135**, 1603–1618.

Wang, Z., Rouard, M., Biswas, M.K., Droc, G., Cui, D., Roux, N., Baurens, F.-C. *et al.* (2022b) A chromosome-level reference genome of *Ensete glaucum* gives insight into diversity and chromosomal and repetitive sequence evolution in the Musaceae. *Gigascience*, **11**, giac027.

Watanabe, S., Harada, K. and Abe, J. (2012) Genetic and molecular bases of photoperiod responses of flowering in soybean. *Breed. Sci.* **61**, 531–543.

Watanabe, K., Stringer, S., Frei, O., Umićević, M.M., de Leeuw, C., Polderman, T.J.C., van der Sluis, S. *et al.* (2019) A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348.

Weeks, N.T., Luecke, G.R., Groth, B.M., Kraeva, M., Ma, L., Kramer, L., Koltes, J.E. *et al.* (2018) High-performance epistasis detection in quantitative trait GWAS. *Intl. J. High Perform. Comput. Appl.* **32**, 321–336.

Wei, K., Wang, J., Sang, M., Zhang, S., Zhou, H., Jiang, L., Clavijo Michelangeli, J.A. *et al.* (2018) An ecophysiologically based mapping model identifies a major pleiotropic QTL for leaf growth trajectories of *Phaseolus vulgaris*. *Plant J.* **95**, 775–784.

Wei, X., Qiu, J., Yong, K., Fan, J., Zhang, Q., Hua, H., Liu, J. *et al.* (2021) A quantitative genomics map of rice provides genetic insights and guides breeding. *Nat. Genet.* **53**, 243–253.

Weighill, D., Jones, P., Shah, M., Ranjan, P., Muchero, W., Schmutz, J., Sreedasyam, A. *et al.* (2018) Pleiotropic and epistatic network-based discovery: integrated networks for target gene discovery. *Front. Energy Res.* **6**, 30.

Weighill, D., Jones, P., Bleker, C., Ranjan, P., Shah, M., Zhao, N., Martin, M. *et al.* (2019) Multi-phenotype association decomposition: unraveling complex gene-phenotype relationships. *Front. Genet.* **10**, 417.

Wen, M., Su, J., Jiao, C., Zhang, X., Xu, T., Wang, T., Liu, X. *et al.* (2022) Pleiotropic effect of the *compactum* gene and its combined effects with other loci for spike and grain-related traits in wheat. *Plants*, **11**, 1837.

Williams, O.R., Vander Schoor, J.K., Butler, J.B., Hecht, V.F.G. and Weller, J.L. (2024) Physical seed dormancy in pea is genetically separable from seed coat thickness and roughness. *Front. Plant Sci.* **15**, 1359226.

Wisser, R.J., Kolkman, J.M., Patzoldt, M.E. and Balint-Kurti, P.J. (2011) Multivariate analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a *GST* gene. *Proc. Natl. Acad. Sci. USA*, **108**, 7339–7344.

Xiang, Y., Helal, M.M.U., Liang, L., Zhang, M., Han, H., Dai, X., Zhao, J. *et al.* (2023) Genome-wide association study exposed the pleiotropic genes for yield-related and oil quality traits in *Brassica napus* L. *Oil Crop Sci.* **8**, 156–164.

Xie, Q., Li, N., Yang, Y., Lv, L., Yao, H., Wei, R., Sparkes, D.L. *et al.* (2018) Pleiotropic effects of the wheat domestication gene Q on yield and grain morphology. *Planta*, **247**, 1089–1098.

Xie, L.-H., Zhu, Y.-J., Tang, S.-Q., Wei, X.-J., Sheng, Z.-H., Jiao, G.-A., Hu, P.S. *et al*. (2020) Pleiotropic effects of rice florigen gene *RFT1* on the amino acid content of unmilled rice. *Front. Genet.* **11**, 13.

Xu, F., Zhou, R., Dossou, S.S.K., Song, S. and Wang, L. (2021) Fine mapping of a major pleiotropic QTL associated with sesamin and sesamolin variation in sesame (*Sesamum indicum* L.). *Plants*, **10**, 1343.

Xu, Y., Zhang, X., Li, H., Zheng, H., Zhang, J., Olsen, M.S., Varshney, R.K. *et al*. (2022) Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. *Mol. Plant*, **15**, 1664–1695.

Xu, D., Bian, Y., Luo, X., Jia, C., Hao, Q., Tian, X., Cao, Q. *et al*. (2023) Dissecting pleiotropic functions of the wheat Green Revolution gene Rht-B1b in plant morphogenesis and yield formation. *Development*, **150**, dev201601.

Yan, W.-H., Wang, P., Chen, H.-X., Zhou, H.-J., Li, Q.-P., Wang, C.-R., Ding, Z.H. *et al*. (2011) A major QTL, Ghd8, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol. Plant*, **4**, 319–330.

Yang, M., Chen, J., Chang, Y., Wan, S., Zhao, Z., Ni, F. and Guan, R. (2016) Fine mapping of a pleiotropic locus (*BnUD1*) responsible for the up-curling leaves and downward-pointing siliques in *Brassica napus*. *Int. J. Mol. Sci.* **24**, 3069.

Yang, X., Lai, Y., Wang, L., Zhao, M., Wang, J., Li, M., Chi, L. *et al*. (2023a) Isolation of a novel QTL, *qSCM4*, associated with strong culm affects lodging resistance and panicle branch number in rice. *Int. J. Mol. Sci.* **24**, 812.

Yang, Y., Zhang, Y., Li, J., Xu, P., Wu, Z., Deng, X., Pu, Q. *et al*. (2023b) Three QTL from *Oryza meridionalis* could improve panicle architecture in Asian cultivated rice. *Rice*, **16**, 22.

Yu, S.B., Li, J.X., Xu, C.G., Tan, Y.F., Gao, Y.J., Li, X.H., Zhang, Q. *et al*. (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. USA*, **94**, 9226–9231.

Yu, K., Wang, J., Sun, C., Liu, X., Xu, H., Yang, Y., Dong, L. *et al*. (2020) High-density QTL mapping of leaf-related traits and chlorophyll content in three soybean RIL populations. *BMC Plant Biol.* **20**, 470.

Yuan, B., Qi, G., Yuan, C., Wang, Y., Zhao, H., Li, Y., Wang, Y. *et al*. (2023) Major genetic locus with pleiotropism determined seed-related traits in cultivated and wild soybeans. *Theor. Appl. Genet.* **136**, 125.

Zanini, S.F., Bayer, P.E., Wells, R., Snowdon, R.J., Batley, J., Varshney, R.K., Nguyen, H.T. *et al*. (2021) Pangenomics in crop improvement—from coding structural variations to finding regulatory variants with pangenome graphs. *Plant Genome*, **15**, e20177.

Zhan, X., Sun, B., Lin, Z., Gao, Z., Yu, P., Liu, Q., Shen, X. *et al*. (2015) Genetic mapping of a QTL controlling source–sink size and heading date in rice. *Gene*, **571**, 263–270.

Zhang, J. (2023) Patterns and evolutionary consequences of pleiotropy. *Annu. Rev. Ecol. Evol. Syst.* **54**, 1–19.

Zhang, Y.-S., Wu, W. and Xu, C.-G. (2008) Two pleiotropic intervals of rice were assessed using NILs constructed by two methods. *Yi Chuan*, **30**, 781–787.

Zhang, J., Singh, A., Mueller, D.S. and Singh, A.K. (2015) Genome-wide association and epistasis studies unravel the genetic architecture of sudden death syndrome resistance in soybean. *Plant J.* **84**, 1124–1136.

Zhang, X., Ren, Z., Luo, B., Zhong, H., Ma, P., Zhang, H., Hu, H. *et al*. (2022) Genetic architecture of maize yield traits dissected by QTL mapping and GWAS in maize. *Crop J.* **10**, 436–446.

Zhou, Z.H., Liu, G.X., Su, L.T., Han, L. and Yan, L. (2014) Detecting epistasis by LASSO-penalized-model search algorithm in human genome-wide association studies. *Adv. Mater. Res.* **989–994**, 2426–2430.

Zhou, Y., Zhang, Z., Bao, Z., Li, H., Lyu, Y., Zan, Y., Wu, Y. *et al*. (2022) Graph pangenome captures missing heritability and empowers tomato breeding. *Nature*, **606**, 527–534.

Zhou, J., Wan, Z., Gali, K.K., Jha, A.B., Nickerson, M.T., House, J.D., Tarán, B. *et al*. (2023) Quantitative trait loci associated with amino acid concentration and *in vitro* protein digestibility in pea (*Pisum sativum* L.). *Front. Plant Sci.* **14**, 1083086.