Harnessing Agronomics through Genomics and Phenomics in Plant Breeding: A Review

Chunpeng James Chen[1], Jessica Rutkoski[2], James C. Schnable[3], Seth C. Murray[4], Lizhi Wang[5], Xiuliang Jin[6], Benjamin Stich[7], Jose Crossa[8], Ben J. Hayes[9], and Zhiwu Zhang[1*]

[1]Department of Crop and Soil Sciences, Washington State University, Pullman, WA, USA;
[2]Department of Crop Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA;
[3]Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, USA;
[4]Department of Soil and Crop Sciences, Texas A&M University, College Station, TX, USA;
[5]Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, IA, USA;
[6]Institute of Crop Sciences, Chinese Academy of Agricultural Sciences/Key Laboratory of Crop Physiology and Ecology, Ministry of Agriculture, Beijing 100081, China;
[7]Institute for Quantitative Genetics and Genomics of Plants;
Heinrich Heine University, Düsseldorf, Germany;
[8]International Maize and Wheat Improvement Center (CIMMYT), Km 45 Carretera México-Veracruz, Texcoco, Edo. de México, México;
[9]Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Australia.
[*]Correspondence should be addressed to ZZ (email: Zhiwu.Zhang@WSU.EDU).

Table of Contents

Highlights

- Phenomics allows genetic mapping of traits throughout plant growth and a better understanding of both pleiotropy and endophenotypes.
- Many GWAS identified loci are not validated and remain unused.
- Genomics provides a path to validate phenomics traits, especially the ones without direct links to agronomics traits.
- Multiple trait GBLUP remains the most sophisticated method to integrate agronomics and phenomics traits but does not demonstrate consistent advantage across traits and environments.
- Bayesian methods have been extended to incorporate multiple traits, multiple environments, and prior biological knowledge.
- Prediction using explainable machine learning incorporates domain knowledge with complex modules allowing to learn from data but guided by confirmed domain knowledge rather than less reliable human expert experiences.

Abstract

Plant breeding primarily focuses on improving conventional agronomic traits, e.g. yield, quality, and resistance to biotic and abiotic stress; however, genetic improvement methods are being rapidly enhanced through genomics and phenomics. In the Genomics-Phenomics-Agronomics (GPA) paradigm, diverse research approaches have been conducted to bridge any two of these elements, and recently, all of them together. This review first highlights the progress to link i) genomics to agronomics; ii) genomics to phenomics; and iii) phenomics to agronomics. Secondly, the GPA domain is dissected into different layers, each addressing the three elements simultaneously. These dissected layers include genetic dissection through gene mapping using genome-wide association studies and genomic selection using Best Linear Unbiased Prediction, Bayesian approaches, and machine learning. The objective of the review is to help readers to grasp the core developments among the exponentially growing literature in each of these fields. Through this review, the connections among the three elements of the GPA paradigm are coherently integrated toward the prospect of sustainable development of agronomic traits through both genomics and phenomics.

Introduction

A key goal of agronomy is to understand and improve phenotypic characteristics such as yield, end-use quality, nutrition response, and resilience to abiotic and abiotic stress. Phenotypic variation is a product of the interactions between genotype and environment. Such interactions are inaccessible without detailed phenotypic data, even with genotypes measured accurately at both individual and genomic levels[1]. The need for accurate and efficient assessment of agronomic traits led to the creation of phenomics approaches, in which phenotypic data is collected in ways that go beyond conventional measurements such as weights and counts. From this aspect, phenomics can be classified into forward phenomics and reverse phenomics [2]. Forward phenomics screens traditionally valuable traits in either a high-throughput and low-resolution manner, or a higher-resolution and lower-throughput manner. Reverse phenomics captures traits to explore mechanisms of physiological, biochemical, or biophysical processes or collect data useful for phenomic selection [3,4]. The forward phenomic traits include previously intractable traits, such as yield evaluated at a large scale using satellites [5–7]. Reverse phenomics traits are much newer, with evolving methods described throughout this review. Some reverse phenomic traits even blur the boundary with genomics. For instance, microbial phenomics measures the expression of specific genomes through sequencing [6,8].

Merging genomics and phenomics with agronomy provides a new path toward mapping from genes to phenotypic variation. As agronomic traits of interest remain mostly the same for plant breeding [9], the technologies in genomics and phenomics continue to advance at an accelerating pace. In the Genomics-Phenomics-Agronomy (GPA) paradigm, diverse research approaches have been conducted to bridge any two of these elements, and recently, all of them together. The latter case not only brings a new path of enhancement, but also unprecedented complexity [10]. With correlation to agronomic traits, phenomic traits can be useful for the prediction of agronomic traits [11]. The impact can be further enhanced in conjunction with genomics. Latent factors can be derived to reduce the dimensions of raw phenomic traits. The prediction of agronomic traits can be improved by incorporating the genetic loci identified for the latent traits (**Figure 1a**). This has been done for lipid content in oats[12]. One hundred latent factors were derived from the 1668 metabolites. After dividing the whole genome into the regions associated with the latent factors and the rest to derive kinship separately, the genomic BLUP (Best Linear Unbiased Prediction) with the two separated kinship outperformed the genomic BLUP with the regular genomic kinship to predict lipid content.

An exponentially growing body of literature has focused on the benefits and the challenges resulting from GPA integration. In this review, we first recap the current state of understanding the pairwise relationships among genomics, phenomics, and agronomics (sections I, II, and III). Then, we provide an overview of how the three components as a whole are dissected through genetic mapping using genome-wide association studies (GWAS), and genomic prediction using best linear unbiased prediction, Bayesian methods, and Machine Learning (ML) (section IV to

VII) (**Figure 1b and 1c**). Finally, we give concluding remarks on the challenges and potential solutions. Although we discuss the strong impact of GWAS, we do not seek to specifically review the methods which have been well described in other reviews[13–17].
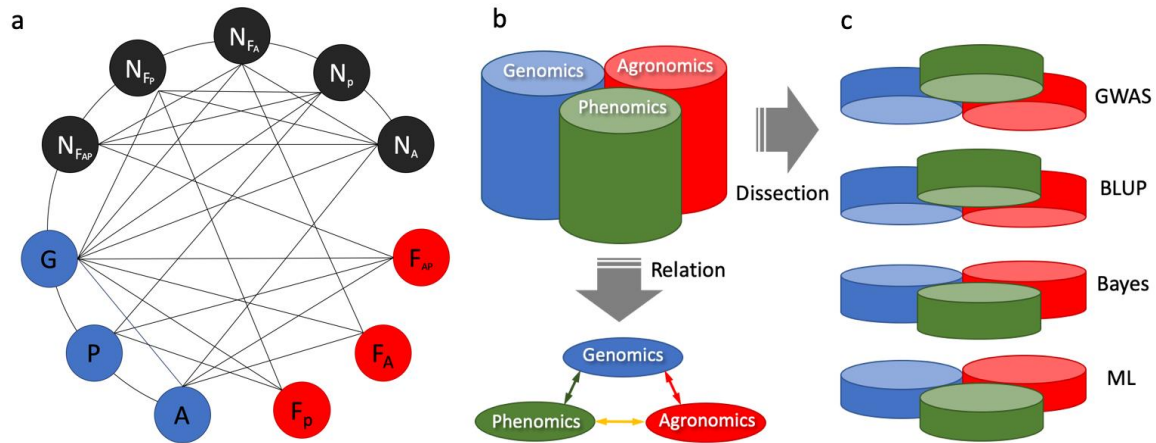

Figure 1. Relationship and dissection among genomics, phenomics, and agronomics.
Genomics (G) and Phenomics (P) are two major factors related to and interacting with Agronomics (A). Factor analyses can be conducted on A and P separately (F_A and F_P) or jointly (F_{AP}). A, P, and their factors (F_A, F_P, and F_{AP}) can be applied to G for genomic selection or through genome-wide association studies (GWAS) to map causal nucleotide mutations (N). The linkage and interaction among the causal nucleotide mutations are beneficial to leverage in both GWAS discovery and genomic selection on A (a). Among the possible relations among GPA, three of them are critical, from G to A, from P to A, and between P and A (b). There are at least four ways to dissect these relations (c), including genetic mapping using GWAS, and prediction of agronomic traits using Best Linear Unbiased Prediction (BLUP), Bayes approaches, and Machine Learning (ML).

I. Agronomics and Genomics

Upon the development of new molecular biological technologies to assess allelic diversity in the 1980s, Marker Assisted Selection (MAS) was introduced. MAS is particularly powerful for traits with simple genetic architecture. MAS was quickly adapted in plant breeding, such as for SUB1 in rice [18], but largely unpublished in the private sector (such as disease resistance). However, many economically important traits have complex genetic architecture with the majority of genetic variance resulting from large numbers of genes with individually small effects. This has limited the number of breeding applications for which MAS is suitable.

In many plant species, genetically identical individuals can be produced in large numbers. This allows for experiments to be repeated across multiple locations and over multiple years and results in a high statistical power. In contrast, phenotypic data for most animal breeding programs come from large numbers of genetically unique individuals raised and phenotyped on production farms[19]. Therefore, the pressure to develop methods and tools to analyze large and highly unbalanced data sets with many nuisance factors was much higher in the animal breeding field[20]. This led to the development of new and more advanced methods such as Best Linear Unbiased Prediction (BLUP) by Henderson in the framework of the Mixed Linear Model (MLM) [21–23]. Multiple software packages have been developed to implement the BLUP method in practical breeding. Some of them are free to the public, including DMU [24] and MTDFREML[25].

To accommodate genes with a small effect, Henderson's BLUP was extended by Fernando and Grossman to MAS BLUP (MBLUP) with two random genetic effects, one for a major quantitative trait locus (QTL) and the other for the total genetic effects of the remaining QTLs [26]. The covariance among individuals was defined by both pedigree and the linked marker for the QTL effect, and by pedigree only for the remaining QTLs. To deal with traits completely lacking major genes, Bernardo reduced the two random genetic effects in MAS BLUP back to one term only on the genes with small effects with covariance proportional to kinship defined by all available genetic markers across the genome [27]. This earliest genomic BLUP used RFLP markers to predict the performance of combinations of inbred maize lines. Before genomic BLUP gained its current popular abbreviation, GBLUP, the same method was independently reinvented at least twice, once in 1997 [28] and once in 2007 [29]. The 1997 reinvention proposed to replace the pedigree relationship with the marker-based relationship and demonstrated that marker-based relationship outperformed the pedigree relationship using simulation with 5, 30, and 100 QTLs. The 2007 reinvention added a new function in the MTDFREML package to conduct BLUP with kinship calculated from all available genomic markers. In 2008, the GBLUP method was implemented by VanRaden in the national dairy genetic evaluation [30], marking a milestone for practical applications of genomic selection (GS).

The success of GBLUP in dairy cattle breeding first demonstrated the advantage of GS to select individuals with intractable traits, especially for early selection. Breeding dairy cattle primarily requires selection of bulls. Biological realities mean there is no way to directly evaluate the potential of a bull for milk production. Traditional dairy breeding therefore used data from the daughters of a bull to evaluate the milk production potential of that bull's genetics. However, bulls were up to seven years old before they had sufficient daughters with milking records to derive accurate enough estimated breeding values for selection [31]. In contrast, in a breeding program with genomic selection, bulls can be accurately selected to provide semen to the industry for artificial insemination at the age of only 12 months. This reduction of the breeding cycle has doubled the rate of genetic gain, compared with the gain of selection under the progeny testing system [31]. The same advantage of genomic selection observed in dairy was also demonstrated in other species, especially for those with long generation times such as trees [32], or low multiplication rates such as clonally propagated species [33], but also those species with shorter generation cycles such as maize [34–36].

The application of GS from dairy cattle to other species including plants emanated not only from the GBLUP invention and reinventions, but also from the adaptation of the GS concept by Meuwissen, Hayes, and Goddard in 2001, forming what we currently known as GS [37]. Using a totally different approach, the authors expanded MAS to incorporate all genetic markers, rather than only the significant ones. The authors simultaneously fit all markers as random effects to avoid overparameterization. The marker effects were assumed to have a normal distribution with a zero mean and a uniform variance. Furthermore, several prior distributions were assumed for the variances of the marker effect distribution. One prior distribution is a constant, which is equivalent to Ridge Regression. The authors named this method BLUP, which is currently known as Ridge Regression BLUP (rrBLUP). Another prior distribution is inverse Chi-square. Bayesian approaches were used to solve this type of model with the distribution of distribution parameters, including Gibbs sampling. As some genes have larger effects than others, the authors restricted some of the markers to have an effect of zero and named the Bayesian methods as BayesA and BayesB for without and with the application of the restriction, respectively. The 2001 study demonstrated that rrBLUP was much better than MAS. For some traits, BayesB outperformed BayesA, and the order reversed for other traits. Since then, the Bayes alphabet has gotten longer due to selecting prior distributions and restrictions, including BayesR [38] and BayesCpi [39]. It has become a widely accepted concept to use genomic selection with genetic markers covering the whole genome to select individuals even when their phenotypes are absent.

There was a long delay to apply GS in any species since the introduction of GBLUP in 1994 [27] and then Bayesian methods in 2001[27]. Two factors mainly contributed to the delay. First, GS required genotyping large numbers of markers across all individuals being evaluated and/or used to construct prediction models. Existing technologies were sufficient to genotype small population numbers to evaluate GS, but at the time genotyping remained comparatively

expensive at the scale needed for widespread use of GS. Second, while Bayesian methods increased public awareness of GS by demonstrating that GS results in more accurate predictions than MAS, the computing time required to produce predictions was substantially higher than existing methods. The existing computational software used conventional BLUP, which involves only a few random effects such as the additive genetic effect. The Bayesian method is the opposite, containing tens of thousands of random marker effects. The existing genetic evaluation systems implemented the BLUP models which contain only several random effects. There is no straight transition without explicitly replacing the existing software.

Table 1. BLUP methods to predict breeding values of individuals by incorporating their kinship

Method*	Year	Kinship to define the variance structure of random genetic effect
BLUP	1975	Pedigree-based kinship with direct inverse matrix [23,40,41]
MBLUP	1989	Two kinships are involved; one is based on a pedigree maker linked to a QTL to capture the random QTL effect, and the other is based on pedigree to capture the remaining QTLs [26]
GBLUP	1994	Genomic markers [27]
TBLUP	2010	Genomic markers weighted by their strength in GWAS [42]
SBLUP	2016	Associated markers identified by SUPER GWAS method [43]
CBLUP	2016	Individuals are compressed into groups and predicted accordingly [43]

*BLUP – Best Linear Unbiased Prediction, SUPER – Settlement of MLM Under Progressively Exclusive Relationship, QTL – Quantitative Trait Loci, MBLUP – MAS BLUP, GBLUP – Genomic BLUP, TBLUP – Trait specific BLUP, SBLUP – SUPER (Settlement of MLM Under Progressively Exclusive Relationship) BLUP, CBLUP – Compression BLUP.

VanRaden demonstrated that rrBLUP had identical results with GBLUP when the kinship was derived from his efficient algorithm[30]. GBLUP has the advantage, only requiring minor modification to the existing software that uses pedigrees to derive kinship. The only addition is to import genetic markers and derive kinship for substitution[29]. Since GBLUP has higher computational efficiency than Bayesian methods, more methods have been developed to extend

the BLUP alphabet, including the transformation of the maximum likelihood methods (MLM) used for GWAS with improved statistical power over the regular MLM. In 2018, SUPER (Settlement of MLM Under Progressively Exclusive Relationship) and Compressed MLM GWAS methods were transformed for genomic selection. The corresponding GS methods were named SUPER BLUP (SBLUP) and Compressed BLUP (CBLUP). The introduction of these new methods enriched the BLUP method series, forming a BLUP alphabet, including GBLUP, SBLUP, CBLUP, and single step BLUP (SSBLUP) to integrate relative individuals without marker information (**Table 1**).

Many studies have been conducted to compare prediction accuracy between the Bayesian methods and GBLUP, and these revealed that superiority between the two methods depends on the genetic architecture of the specific traits evaluated (e.g., [44]). In general, Bayes B has the accuracy advantage over GBLUP for traits controlled by smaller numbers of genes; otherwise, GBLUP has the advantage [45]. Both real and simulated data demonstrated that SBLUP is superior for traits controlled by a smaller number of genes, such as Mendelian traits. Compressed BLUP is superior for traits with low heritability [43]. Recently, a method based on cloud computing was developed for GS (MMAP) that automatically selects the best prediction method for a given particular dataset. The system also avoids the difficulties users often experience when installing programs and finding computing resources. Users can easily upload the data and download the results [46].

II. Genomics and Phenomics

A major limitation in efforts to link genotype to phenotype in the past has been the time and cost required to score traits across sufficiently large populations for later quantitative genetic analyses. In recent years, major investments in plant phenotyping technologies have been made[47–50] However, while a great deal of motivation for developing new phenomic technologies has focused on reducing the cost of data per plot or increasing the accuracy of each data point [51], many of the advances in phenomics technology have not significantly reduced the effort or cost of collecting the first data point per plot. Instead, they have dramatically reduced the marginal cost and effort required to collect additional measurements per plot beyond the first. This can include both lowering the cost or effort to score the same trait multiple times (temporally) in the growing season and lowering costs and effort to score multiple properties of a single plot at the same time. Plant phenotyping approaches in controlled environments [52–54], though less directly relevant to plant breeding and agronomic questions than field phenotyping, can often illustrate the challenges and opportunities also present in field phenotyping [51].

Fixed field platforms provide the opportunity to collect sensor data at different time points, including in intemperate weather, and to employ sensors either too heavy or requiring too much power to be carried on flying platforms [55–57]. However, a key challenge with these systems is to collect data from a sufficiently large number of plots to conduct quantitative genetic analyses or train genomic prediction models. Gantry-based systems in Rothamsten (UK) [55] and Arizona

(USA) [56] are able to collect data from 0.12 and 0.61 hectares of land, respectively, enough land to grow either 130 or 650 standardized maize yield plots [55,56]. Cable-based systems, while being of lower cost to construct, face similar limitations in the number of plots from a quantitative genetics and breeding perspective; one cable-based system deployed in Nebraska (USA) covers a 0.4-hectare field – estimated 430 plots – [57] and another at ETH-Zurich covering an area of 1 hectare – estimated 1070 plots [58]. The reuse of the same field for small plot phenotyping trials from one year to the next can also introduce additional spatial variation from the small plot work in previous years. Fixed field platforms are incredibly powerful as test beds, but scaling to large-scale quantitative genetic models requires data collected from more and larger fields across more environments. Efforts in the field phenotyping for quantitative genetic traits outside of field platforms fall into three large categories from most to least labor intensive.

The most labor intensive of these is likely to use technology to measure new traits in the field, which still requires individual human intervention on a per plant/per plot basis. For example, manual excavation and measurement of roots can identify QTL controlling root phenotypes under field conditions [59]. More recent efforts have automated the quantification of measurements for manually extracted roots [60] and reduced the amount of time required per root dug, making it logistically practical to collect sufficient measurements to conduct GWAS for maize and sorghum roots under field conditions [61]. Manually collected hyperspectral reflectance data similarly requires a human to walk through the field with a backpack spectrometer collecting measurements from each leaf. However, recent advances demonstrate that one sensor data type can be used to predict a range of plant nutrient, plant morphological, and photosynthesis-related parameters in maize, wheat, and tobacco [62–67].

Intermediate in required labor intensity are approaches to phenotyping that use mobile robotic platforms. These approaches require only a moderate amount of human intervention per plant/plot measured but presently still require a human supervisor whenever the robotic platform is in action. Ground-based robots can take two strategies, either larger platforms which drive above the plant canopy [68,69], or smaller rovers that fit between individual rows [70,71]. Both approaches can measure a range of heritable (genetically controlled) traits with sufficient accuracy to identify QTLs controlling variation in natural or structured populations. The Phenobot 1.0 system was able to quantify variation in leaf angle independently at the top, middle, and bottom of sorghum plants from an association panel grown at field density, conduct GWAS, and identify separate loci controlling variation in leaf angle in different parts of the plant which were validated by manual measurements [72]. TerraSense below-canopy rovers were able to decompose the 3D structure of the canopies of 890 maize hybrids into a series of latent space phenotypes [73] with a heritability of up to 0.44 [74]. However, scaling robotic platforms for large-scale breeding and quantitative genetics research will require solving problems of navigation, particularly under the leaf canopy which can interfere with high-precision Global Positioning System (GPS) signals. Automated navigation and data processing for flying robotic

platforms are, in some ways, more advanced than ground-based ones [75,76]. In estimates from flying robotic platforms, the proportion of variance in plant height explained by genetic factors was equal to or greater than ground truth measurements collected in sorghum or maize [77]. Within the United States, flying robot platforms require a dedicated handler during data collection, like ground-based systems. However, capability for autonomous robotic data collection from flying platforms does exist [78], and autonomous phenotyping from flying platforms is already being employed elsewhere in the world where regulations permit [79]. Once a mobile robotic phenotyping platform is constructed or purchased and an operator has been trained, the barriers to frequent data collection are much lower, enabling the collection of many more data points throughout a growing season than would be practical with manual measurements [80,81].

The third class of phenotyping approaches have the theoretical potential to provide a level of phenotyping and plot evaluation with essentially no manual labor required per datapoint collected. A representative example of this group of phenotyping technologies is the estimation of plant traits from satellite images. Satellite data has been previously employed to estimate plant genetics and plant breeding relevant traits including plant height, nutrient status [82], and leaf area index [83], but with a spatial resolution of multiple meters, making it an impractical tool for plant breeding. For reference, a standard corn yield plot might measure on the order 1.5 meters x 5 meters. Commercially available satellite imagery with spatial resolutions of 0.3-0.5 meters per pixel is now available from multiple sources (As reviewed in [84]) with revisit times (how frequently new images can be collected from the same location) as short as every one to five days [85], though cloud cover can remain an issue. To date, only a modest number of studies have evaluated the potential of this technology in a quantitative genetics or breeding context. Phenotyping of large-scale yield plots of spring wheat (2.5 meters x 8.5 meters) with satellites achieved a correlation of 0.53 between satellite data and yield and a correlation of 0.58 between satellite data and biomass [86]. The seed yield and biomass of a bean diversity panel grown in 3 x 2 meter plots could be predicted with an accuracy of 0.52 and 0.55, respectively, using satellite data with a spatial resolution of 0.5 meters [87], while the grain yield of 1.5 x 5 meter plots maize hybrids could be predicted with an accuracy of 0.34 [88]. When overall correlations between predicted and ground-truth measurements are low, it is particularly important to quantify the proportion of variance explained by genetic factors both for predicted measurements and for the residual between ground truth and predicted measurements. For estimates of maize biomass from proximal image data, more than half the variance in the direction and magnitude of measurement error in estimated biomass is explained by genotype-to-genotype variation [89]. Replicated data collected from diversity panels such as described by [87] provides an opportunity to evaluate the genetic control of error in satellite data, an important question to be addressed before satellite-based estimates can be incorporated into breeding and genetics research.

Many of the technologies described are on the cusp of being adopted for genomics and breeding, but current publications focus primarily on proof of concepts and evaluations of feasibility. As these technologies progress from demonstration to application, new questions will have to be addressed. For many of the technologies described above, the marginal cost of collecting data from additional time points is much lower than the cost of deploying the phenomics infrastructure (hardware, software, and human) necessary to collect data from the first time point. Time series data can provide greater power to map causal variants, and more insight into how individual genetic variants influence phenotypes over time [80,81,90,91]. But it will take experimentation and evaluation to develop best practices for how many time points can be collected before the marginal value of new data exceeds even the low marginal cost of having a drone drive through or fly above a field in different types of experiments seeking to address different questions.

The split between sensor data collection and the extraction of numerical traits, which are unified in manual phenotyping but divided in high-throughput phenotyping, also creates unique challenges and opportunities. If sensor data are collected and stored in standardized formats, as new algorithms are developed or trained to estimate new traits, it will be possible to conduct QTL and GWAS mapping or train new genomic selection models entirely in silico using published genetic marker and sensor data. A recent example of the potential for in silico quantitative genetic research is a QTL mapping study conducted in a Setaria RIL population using published image and marker data [73]. This creates opportunities to address questions of both pleiotropy and genotype-by-environment (G×E) interactions. Pleiotropy will be detected where diverse sensor data collected from genetic mapping and GWAS populations, results in larger numbers of distinct traits being scored from a common population than would ever be feasible with manual phenotyping. GxE interactions will be dissected when development and validation of new prediction algorithms in several environments are used to measure the same trait in populations grown across other environments using previously collected and published sensor data.

III. Phenomics and Agronomics

Phenomics seeks to measure all physical properties of a plant, whether or not those properties are of direct agronomic relevance. Advances in high-throughput phenotyping technologies have made it possible to quantify those properties under a wide range of environmental conditions. However, high-throughput phenotyping generally collects raw sensor data, which must first be processed to extract meaningful features. These features may be agronomic traits themselves, indirect traits to plant phenotypes, or novel digital phenotypes with unknown, if any, relationship to agronomically relevant traits.

Different spectral sensors have been used to collect images and proximal measurements for extracting plant phenotyping traits. Different kinds of agronomic phenotyping traits have been obtained using image characteristics processing algorithms and ML methods [92,93]. The data of photosynthesis sensors without images are generally used to evaluate the crop photosynthesis status under different stress conditions by calculating the difference between carbon dioxide and water [94]. Compared with photosynthesis sensor data, images from fluorescence sensors have the advantage to explore the spatial distribution of leaf photosynthetic status and early monitoring of plant stress based on calculated maximum fluorescence, initial fluorescence, photosynthetic quantum yield, and photochemical-quenching coefficient [95–97].

Stereo cameras can be used to get three-dimensional images of plants [98]. Three-dimensional segmentation algorithms and ML methods can be applied to these images for obtaining structural phenotyping traits of plant growth and development [99]. Similarly, point clouds of Light Detection and Ranging (LIDAR) can be processed to obtain more accurate three-dimensional structural phenotyping traits based on classification and segmentation methods [100–103]. The images of RGB cameras are analyzed based on binary image processing and ML methods to obtain plant phenotyping traits, including plant density [104], canopy coverage [105], height [106], and leaf rolling [107]. Images from the multispectral and hyperspectral cameras with the high-spectral resolution are explored to estimate green area index [108], nitrogen content [109], and chlorophyll content [110] using optical physical models, data classification methods, and ML methods. Images from thermal cameras are used to estimate canopy temperature [111] and lodging [112] using a texture extracting algorithm, image segmentation algorithms, and ML methods.

In general, these sensors can be loaded by near-ground phenotyping platforms. The sensor images from the unoccupied aerial vehicles (UAVs) are first processed to export the orthomosaic and the digital elevation models (DEMs) using commercial software (such as Pix4D; Progeny (now Phenix); Hyphen; and OpenDroneMap Agisoft Photoscan Professional (Version 1.2.2, Agisoft, Saint Petersburg, Russia). The orthomosaic and DEMs data can be used to analyze dynamic traits at the plot scale throughout the entire growing season. The sensor images from the stationary platforms [55] and phenomobiles [113] are first calibrated to further estimate plant phenotyping traits of interest. Plant phenotyping traits of the images from UAVs, stationary platforms, and phenomobiles are extracted using a multisource combination method including the image extracting algorithms[114–116], ML methods[117], optical physical model[118], and data dimensionality reduction method[92]. However, the processing protocols of the stationary platforms and phenomobiles are largely different due to different image processing strategies.

Many sensors are effective to measure traits that are observable and not observable by human eyes. The observable traits include color index [108], plant density [104], ear density [119], leaf rolling [107], lodging [112,120], and other traits. The RGB, multispectral, hyperspectral,

thermal, and stereo cameras and LIDAR can be used to quickly and accurately get these plant phenotyping traits in high-throughput fashion. The images from these sensors have been combined to improve the estimation accuracy [105,112]. The non-observable traits include chlorophyll content [108], nitrogen content [96], and canopy temperature [112]. RGB, multispectral, hyperspectral, and thermal cameras; photosynthesis sensors; and fluorescence sensors could be used to quickly and efficiently monitor these traits indirectly.

In addition to direct interpretable traits, phenomics allows to create novel digital traits built by computer vision. For example, the convex hull and plant aspect ratio [121–123], convex hull mesh [124], side projected area [125], bi-angular convex-hull area ratio [126], eccentricity and compactness [127,128], as well as circularity and sphericity [129] have not yet demonstrated an actual value for plant breeders; yet phenomic selection approaches suggest value may yet be found [3]. Genomics has demonstrated its role to interpret these novel digital traits and apply them in plant breeding. For 1668 metabolites, GWAS identified genomic loci associated with the latent factors of the metabolites. The incorporation of the associated loci in GBLUP improved the prediction accuracy for lipid content, one of the breeding goals in oat [130].

The fast development of micro- and nanotechnology and advanced flexible electronic materials has enabled the design of implantable sensors for real-time, continuous, in vivo monitoring of plant molecular indicators and vital signs; optimized agrochemical and water allocation; and finished high-throughput plant physiological state measurements during all growth stages [131–133]. Implantable sensors can be used to translate plant physiological and biochemical signals into digital images with high spatiotemporal resolution. Advances in flexible electronic devices, microfluidics, and nanofabrication technologies have combined with the emergence of the internet of things paradigm to achieve continuous, minimally invasive or noninvasive, long-term and real-time high-throughput plant physiological and chemical phenotypic trait detection [134,135]. The implantable nano-sensors have been successfully applied to measure plant reactive oxygen species [136], wound-induced H_2O_2 waves [137], glucose [138,139], sucrose [140], Ca^{2+} [141,142], nitric oxide [143], ethylene [144], jasmonic acid [145], methyl salicylate [146], abscisic acid [147,148], and pH [149]. Sensors that make it possible to precisely and accurately monitor the onset of stress in individual plants will promote enhancement of plant phenotyping approaches for plant breeding programs, utilizing high-throughput measurements of physiological and biochemical traits of stress-tolerant plant varieties in the future. Besides, implantable sensors will be used to monitor plant health status to improve resource use efficiency.

IV. Merge GPA through GWAS

Integrating genetic mapping with phenomics allows a better understanding of biological function across time, different phenotypes, and connecting genetic loci to agronomics. Across time, low

12

marginal costs in high-throughput phenomics enable four-dimensional (4D) data collection (three spatial dimensions plus time), taking repeated measurements throughout growth. Among the first 4D genetic mapping reported was [150] using a ground vehicle to measure a cotton linkage population for multiple traits throughout growth. Clear temporal patterns were identified, showing the most significant loci at early growth stages may not be detected by the end of the season [150]. Since this time, numerous other authors have had similar findings through temporal mapping in sorghum [91,151] and maize [81]. A wheat study showed a more complex pattern with two of seven primary detected QTL observed at multiple time points, despite the majority of loci only being significant at specific times [152]. The temporal importance of loci suggests mapping restricted to seedlings or juvenile stages alone is unlikely to provide insight on agronomic end traits. Such insight may require temporal phenotyping and modeling that links across development stages and does not treat each temporal measurement as a discrete phenotype. For example, when using UAVs to temporally measure maize plant height, parameters of a logistic growth model fit across growth stages resulted in more robust heritability and yield predictions than from individual temporal measurements [153]; such models also permit comparisons between disparate studies.

Phenomics can also connect genetic mapping to agronomics through the measurement of additional traits. Endophenotypes are phenotypes more closely aligned to underlying biological function, are controlled by fewer loci, and are more heritable than agronomic phenotypes of interest. As a simple example, yield components (e.g., kernel weight or number), are endophenotypes for grain yield. Furthermore, kernel weight can be dissected into endophenotypes of kernel shape and density. When incorporating digital image analysis to phenotype grain in a GWAS panel of synthetic wheat, loci for kernel length, width, thickness, and other shape traits were identified to dissect the trait of interest, grain weight [154]; In estimating the value of these endophenotypes, path analysis was used to understand the network of trait interactions. Integrating measurements into networks is important to identify the biological basis of traits [155].

Genetic mapping studies have often focused on few traits due to the effort and expense of phenotyping and can therefore miss valuable endophenotype loci; phenomics approaches are eliminating such barriers. Combining both known (e.g., plant height) and novel (e.g., plant perimeter side view) phenomic measurements in a maize linkage population, 988 QTLs were mapped for 106 different traits over 16 time points using an automated platform in a greenhouse [36]. Among the 988 QTL, several hotspots were found for multiple QTL, suggesting underlying pleiotropy to exploit or to avoid in agronomic improvement when antagonistic effects are present. Yet many such relationships may be due to major phenology loci. Frequent co-localization of biomass yield and composition are found by measuring 58 known phenotypes relating to aboveground biomass and composition in a linkage mapping population of field-grown sorghum [156].

The two largest such effects were due to known large effect flowering time and plant height phenology loci. A similar observation was for rice [52]. However, numerous other co-localizations demonstrated challenges arising from attribution of significant loci to a single measured phenotype, when other (unmeasured) pleiotropic phenotypes may also be affected. Pleiotropic endophenotype discovery can further be extended to various forms of stress resistance. Various abiotic and biotic stresses were jointly leveraged in mapping to identify loci with major pleiotropic effects in Arabidopsis by a novel multi-trait GWAS approach [157].

Genetic mapping with phenomics timepoints and traits collected in the same environments to get at function is straightforward, but it is more difficult to connect measurements across environments; crop modeling is again helpful to overcome such challenges. As a simplistic example, flowering time, measured in days, is not directly comparable between locations or even the same location over years. A rudimentary model of flowering date informed by weather information converts to growing degree days, comparable across environments and biologically more meaningful. An ecophysiological modeling approach can bridge more complicated environmental interactions with crop growth. An early such study [158] collected time-series measurement of leaf elongation ratios, directly tying these to temperature, evaporative demand, and water pressure. These parameters were used in genetic mapping, improving detection. Endophenotypes like flowering or leaf elongation turn out to be surprisingly complicated, interacting with the environment at the gene, cell, and organism level. Recently, how temperature and photoperiod time affect individual sorghum flowering genes, how these genes interact with the environment, and how the genes interact epistatically with each other were demonstrated through genetic mapping [159].

Because individual genes in a network each respond to different signals (e.g., light, temperature, or time), but can have pleiotropic effects, understanding the interplay of agronomic traits will be improved from complex simulation modeling of their endophenotypes [160,161]. Most complex crop models to date are being applied within the context of GS but would also be effective in genetic mapping by using BLUP output from such models for known or new traits [162,163].

Without *a priori* knowledge, such as required for crop modeling, latent space phenotyping is an approach to generate new heritable phenotypes for mapping agronomic value and distill them down. Here, abstract phenotypes are extracted from images or point clouds using ML based on consistent differences observed between plots [73,74]. Latent space approaches were found repeatable across maize plant structure [74], strawberry fruit shape [164], and several other plant image and synthetic datasets [73]. To date, only one study have used latent space phenotyping to successfully map traits, using Setaria linkage mapping and synthetic Arabidopsis plants using GWAS [73].

Genetic mapping of additional phenomic traits is not without challenges. Computational demands are linear with the number of traits and become restrictive when marker numbers reach hundreds of thousands or when significance thresholds are simulated, as in permutation tests [165]. Data visualization and presentations of more than a few traits and loci quickly become difficult to interpret and use. Integrating related traits can also potentially lead to issues with multicollinearity, which need to be considered and dealt with appropriately [166]. There are also multiple testing issues created in large p, small n studies. Statistical tests run on mapping each trait are typically corrected for marker number using one of an increasing number of false discovery rate approaches [167]. Approaches informing and reducing the number of markers selected based on gene expression or other criteria suggesting function will also reduce multiple testing issues [168,169]. All approaches are geared towards controlling positives from GWAS marker predictors, focusing on one or a few traits; however, control for many phenotype-dependent trait testing is also now needed to prevent false positives as phenomics is merged[170].

V. Merge GPA through BLUP

When extending the GBLUP model to include multiple traits as response variables, the model becomes multi-trait GBLUP (MT-GBLUP) [171]. MT-GBLUP is expected to improve breeding value prediction accuracy relative to single-trait GBLUP by enabling information to be borrowed among correlated traits. The effectiveness of this approach is demonstrated in a genomic recurrent selection program for multiple quality traits in winter squash (*Cucurbita moschata*)[172]. Another advantage of MT-GBLUP is that the breeding value estimates obtained from the model can be used in conjunction with economic or preference weights to estimate an optimal selection index [173].

Traits included in MT-GBLUP can be a combination of primary traits of interest for breeding and correlated 'secondary traits' which are included solely to increase prediction accuracy. Simulation studies have confirmed that MT-GBLUP can improve breeding value prediction accuracy on unobserved individuals, especially for low-heritability traits that have not been recorded on the entire training population [171,174]. However, empirical results have indicated that in many cases, MT-GBLUP may only improve accuracy for a given trait when one or more correlated traits are observed on both the training and validation populations [163,175–180].

Given that breeding programs already phenotype multiple agronomic and quality traits, MT-GBLUP can be readily applied to improve prediction accuracies. It's found that for agronomic, disease resistance, and quality traits in Cassava, MT-GBLUP improved GS accuracy for unobserved individuals, especially for low-heritability traits with moderate to high genetic correlations with high-heritability traits [181]. A study that evaluated MT-GBLUP for grain yield and percent protein in the rye (*Secale cereale* L.) found that for the prediction of untested individuals, MT-GBLUP was more accurate than single-trait GBLUP [182], and the

improvement in accuracy was greatest for the low-heritability trait, percent protein, when that trait was scarcely recorded on the training set.

Similarly, MT-GBLUP was found to be more accurate than single-trait GBLUP for scarcely recorded end-use quality traits in wheat (*Triticum aestivum* L.). For each quality trait, the corresponding NIR- or NMR-based prediction are used as secondary traits [183].

Other studies using MT-GBLUP to model agronomic and quality traits have only observed improvements in accuracy when predicting breeding values of individuals with some phenotypic data [177–180]. This implies that MT-GBLUP will usually not be useful for accelerating breeding cycles, because individuals would still need to be phenotyped to some degree before selection. However, secondary trait data was demonstrated that they could be taken on single plants in a greenhouse and then used in MT-GBLUP to improve accuracy while maintaining rapid breeding cycles [184].

With high-throughput phenotyping (HTP) becoming mainstream, there are thousands of new secondary traits that could potentially be used in MT-GBLUP to improve prediction accuracy for traits of economic importance. Vegetation indices (VIs) and canopy temperature (CT) evaluated using an aerial HTP platform is demonstrated to improve MT-GBLUP accuracy for grain yield in wheat when measured on both the training and validation populations [176]. The authors used a simple repeatability model to combine data from time points within a growth stage, and trait-growth stage combinations were considered as separate traits in MT-GBLUP. This simple repeatability approach to pre-processing the HTP data is not ideal, because data from different time points are incorrectly assumed independent. However, a study that evaluated the use of simple repeatability, random regression, and multivariate models to combine data from different time points found that all methods performed equally well [163]. This outcome may have been due to the relatively low number of data collection time points in the dataset. A study evaluating GS models incorporating HTP data from 20 time points found that random regression outperformed MT-GBLUP [185].

Although MT-GBLUP has not yet been evaluated in an applied breeding scenario where predictions will be made across breeding cycles and at the early stages of testing, validation studies have indicated that MT-GBLUP will be effective in this context. A study that evaluated MT-GBLUP for grain yield in wheat with VIs and CT as secondary traits found that MT-GBLUP outperformed single-trait GBLUP, especially when predicting across breeding cohorts [186]. In addition to across-cohort prediction, breeding programs aspire to predict selection candidates at seed-limited stages before yield testing. Early results evaluating this approach have shown that MT-GBLUP including VIs can improve GS accuracy even when VIs are evaluated on small plots [47].

Another option for modeling large amounts of HTP in a BLUP framework is to treat the HTP variables as predictors rather than as responses; the advantage is that a large number of variables can be more easily included in the model. In 2018, VIs and CT are evaluated to aid GS for grain yield using each trait-time point combination as a fixed effect predictor variable in a GS model and using each trait across time points as a response variable in MT-GBLUP. The authors found that on average, MT-GBLUP performed better than methods that treated secondary traits as predictor variables [187]. Other studies have found that HTP data can be used in the same way as marker data in GS models to improve accuracy [11,188,189]. With this approach, a relationship matrix is estimated using trait values, which may be reflectance values from many spectral bands. The phenotypic relationship matrix is then included as an additional kernel in GBLUP. Just like with markers, another kernel can be added for the interaction of traits-by-environments [11], and the model can be extended to include multiple responses [189]. Most recently, it has been suggested that low-cost HTP data can entirely replace marker data in BLUP to save cost and potentially even improve accuracy [3]. Alternatively, HTP data could be used to impute genotypic data on large numbers of individuals that have not been genotyped [190]. In this way, breeding programs could replace some of their genotyping with low-cost phenotyping to improve efficiency.

VI. Merge GPA through Bayesian methods

Bayesian methods for genomic prediction were developed to incorporate the "prior knowledge" that, for some traits, there were mutations with moderate to large effects, as well as many mutations with small effects [37]. This is in contrast to rrBLUP or GBLUP, where the effect of all mutations on the trait is assumed to be very small and to be derived from a normal distribution. In rrBLUP, this prior assumption is very "strong" – if the single nucleotide polymorphism (SNP) effects predicted from the rrBLUP model are investigated, they are always very small. Note that this prior assumption is shared with GBLUP, which is a mathematical equivalent of rrBLUP [30]. Bayesian methods have been given different names according to the distribution of mutation effects they assume (**Table 2**). In practice, the effect of SNPs in linkage disequilibrium with the actual mutations causing the effects is estimated, rather than the mutation effects directly. The exception to this may occur if whole-genome sequence information is used rather than SNP array or genotyping-by-sequencing (GBS) data.

Many studies have been conducted to compare BLUP and Bayesian methods and the algorithms within each category with inconsistent conclusions or claiming the similarity among the methods. In addition to the differences in statistical power and traits investigated, systematic biases of evaluation criteria also led to the mixed results[191,192] Pearson correlation coefficient is typically used to compare observations observed and predicted using cross-validation. There are two ways to calculate the correlation coefficients. One way is to calculate the correlation instantly soon after the predictions are available for a fold of individuals. The final prediction accuracy is calculated as the averaged correlation across folds, named instant accuracy. The

other way is holding the calculation of correlation until completion of iteration so that only one correlation coefficient is calculated on all individuals, named hold accuracy. The two types of accuracy (instant vs. hold) are not equal and both have biases. The magnitude of hold accuracy is proportional to the number of the fold and reaches the maximum bias with leaving one out (Jack Knife) validation. The instant accuracy has bias only in the case that the number of individuals in each fold is small [192]. As the bias due to small samples can be corrected using the Olkin & Pratt method[193], the instant accuracy is recommended for assessment of prediction accuracy[192].

Table 2. Bayesian methods for genomic prediction

Method*	Year	Assumption of SNP effect distribution
rrBLUP	2001	All SNPs have an effect and effects follow a normal distribution (leading to very small effects for all SNPs) [37]
BayesA	2001	All SNP effects follow a Student's t-distribution (allowing a small proportion of moderate to large effects) [37]
BayesB	2001	Many SNPs with zero effect, the remainder with t-distribution of effects [37]
BayesCπ	2011	A proportion (π) of SNPs with zero effect, the remainder with a normal distribution of effects [194]
Bayesian LASSO	2008	An exponential distribution of SNP effects [195]
BayesR	2012	The distribution of SNP effects is assumed to be from multiple normal distributions, the first with zero variance (giving zero effects), the second with very small variance (very small effects), the third with small variance (small effects), and the fourth with moderate variance (small effects) [196]
BMTME	2016	Normal distribution of effects assumed [197]

*rrBLUP-Ridge Regression Best Linear Unbiased Prediction, BMTME-Bayesian multi-trait multi-environment.

The accuracy of GBLUP and BayesR for 14 traits in maize, cattle, and pigs are compared in a comprehensive study [198]. They found largely similar results for the accuracy of genomic predictions from GBLUP and BayesR. However, the accuracy of prediction from BayesR was higher when traits had a moderate to high heritability, and/or there were mutations of moderate to large effect segregating. Biomass was examined on the accuracy of genomic predictions from GBLUP, BayesA, BayesB, and BayesCpi using 200 sorghum lines as a reference [199]. The study found very similar accuracies of prediction from all methods. In contrast, BayesB was found to give the highest accuracy of genomic prediction for drought tolerance in maize when GBLUP, BayesA, LASSO, and a variety of other methods were compared [200]. The accuracy of genomic prediction for GBLUP and BayesR for leaf rust, stem rust, and stripe rust in a diverse set of wheat landraces are evaluated [201]. For leaf rust and stripe rust, the accuracy of genomic prediction was similar for GBLUP and BayesR; however, for stem rust, the accuracy of genomic prediction for BayesR was considerably higher, likely because a mutation with a large effect was segregating for this trait.

The Bayesian multi-trait and multi-environment (BMTME) model in the original R package was improved for analyzing breeding data [202]. The authors introduced Bayesian multi-output regressor stacking (BMORS) functions that are efficient in terms of computational resources. This improved version of BMTME allows general covariance matrices by using the matrix normal distribution that facilitates easy derivation of all full conditional distributions and also permits a more efficient model in terms of computing time for implementation [203,204].

To summarize results across these and other studies, Bayesian methods for genomic prediction that use a prior distribution of mutation effects with a small but non-zero probability of moderate to large effects (e.g., BayesA, BayesB, BayesCpi, BayesR) appear to have an advantage, in terms of increased accuracy of prediction, when there are mutations of moderate to large effect segregating for that trait. Results from livestock suggest that the Bayesian methods have an advantage (in terms of accuracy of prediction) as the data set size increases to tens of thousands of individuals [205]. If whole-genome sequence data is used rather than random genome-wide SNP markers, Bayesian methods also appear to have an advantage, which is assumed to be a result of the possibility of setting the effects of some mutations to affecting equal to zero (e.g., BayesB, BayesCpi, BayesR) [206,207].

In practice, GBLUP is used much more widely for routine genomic evaluations than the Bayesian methods. This is because the Bayesian methods are typically implemented with Gibbs sampling or Metropolis-Hastings Markov Chain Monte Carlo (MCMC) approaches, which require long compute times[208]. Recently, several strategies have been reported which greatly reduce compute times required for the Bayesian methods. These include
- a BayesR approach which uses GWAS summary statistics (effect estimate and standard error for each SNP) as input [209]

- a hybrid expectation-maximization/Gibbs sampling algorithm implementation of BayesR [210]
- variational Bayes methods (VBM) (e.g., [211]). The VMB methods accelerated approximate inference for the parameters in the Bayesian models, representing a deterministic alternative to Monte Carlo methods [212].
- integrating ML with other approaches to rapidly infer parameters in the Bayesian models [213].

All of these approaches are leading to much wider application of the Bayesian methods. In several cases, the Bayesian methods have been applied to data sets with tens of thousands, or even one hundred thousand, individuals with (imputed) whole-genome sequence data [209,210].

The Bayesian methods have also been extended to incorporate biological knowledge in genomic predictions. Biological knowledge could include annotation of functional elements (e.g., gene space, regulatory elements), knowledge of which genes are differentially expressed in particular treatments (e.g., high and low disease resistance), and information from genome-wide association studies (which genes or genome regions harbor mutations associated with trait variants). The BayesR method was extended to incorporate such information, by allowing each class of variant to have a different distribution of effects (e.g., SNPs in and around genes that are differentially expressed between high and low yielding varieties might have more moderate to large effects than SNPs that are not in such genes) [206]. This new method was termed BayesRC. The BayesRC approach was used to incorporate prior information on mutations (from GWAS) affecting blackleg resistance in Canola [214]. They reported significantly higher accuracies of genomic prediction for blackleg resistance using BayesRC than BayesR. As more biological information becomes available for crop species (e.g., [215]), the accuracy of genomic predictions incorporating such information should improve (e.g., [216]).

There is increasing interest in multi-trait methods to incorporate both novel phenotypes (e.g., hyper-spectral drone images and NIR for grain quality) and G×E interaction in genomic prediction. The Bayesian methods have been extended to multi-trait implementations. For example, a Bayesian multi-trait-multi-environment (BMTME) [197] method was used to predict wheat quality [217]. Genomic predictions from this model had higher accuracy than genomic predictions from a multi-trait ridge regression model (a model very similar to multi-trait BLUP) for many quality traits. Extensions of BayesR for multi-trait applications have also been described [218]. As the cost of acquiring both novel phenotypes and whole-genome sequence data continues to fall, a future can be envisaged where these types of data are available on hundreds of thousands of individual varieties/lines/clones. Computationally efficient multi-trait Bayesian methods will enable accurate genomic predictions from such data.

Both visible and invisible spectral bands are modeled simultaneously by Bayesian functional regression analyses, which gave better prediction accuracy than vegetation indices constructed for predicting grain yield and biomass [219]. In maize trials, using all bands can increases prediction accuracy over vegetation indices [220]. Genomic and phenomic information were linked to model the effect of genomic, G×E, band, and band × environment (B×E) interactions [221]. The authors proposed Bayesian functional regression models with B-spline and Fourier as basis functions that take into account all available bands, genomic or/and pedigree information, the main effects of lines and environments, as well as G×E and B×E interaction effects. The authors observed that the models with B×E were the most accurate models at different time points and environments. The functional regression models are more parsimonious and computationally more efficient than the conventional regression models.

VII. Merge GPA through ML

Both BLUP and Bayesian models for analyzing the relationship between genomics, phenomics, and agronomics are top-down models, which are based on domain knowledge and scientific hypotheses using pre-programmed formulations readily solvable by specific statistical methods. The ability to integrate human knowledge from the inception of these models has resulted in parametric methods with satisfactory performance, but it also introduces explicit or implicit assumptions that may turn out to be either false or overly restrictive. On the other hand, ML models are built from the bottom up using non-parametric methods with different and complementary strategies. These methods have the flexibility to model multiple types of genetic effects, including not only additive and dominant effects, but also epistasis effects. Examples include reproducing kernel Hilbert space (RKHS) regression [222], Support Vector Regression with non-linear kernels [223], Random Forest (RF) [224], and, recently, Deep Learning (DL) [203,204,225–229].

DL is one type of ML using neural networks using a large number of simple computational units growing into more complex modeling systems. As a type of non-parametric ML approach, DL can be considered a subfield of ML, because it provides flexibility to adapt to complicated associations between data and output. An important strength of DL is the ability to adapt to patterns of unknown structure. DL models are inspired by the functioning of the brain, as they try to mimic how the brain works when performing a complex task and are based on multilayer ("deep") artificial neural networks, in which different "neurons" receive input from the layer of the lower hierarchical level being activated according to set activation rules [230,231]. The activation defines the output sent to the next layer, which receives the information as input. Because these models incorporate fewer assumptions, it is often the case that they only become competitive with expert models as the total size of available datasets increases to train the models.

Without pre-programmed human knowledge or biases, an ML algorithm attempts to learn its lesson from massive datasets, which have successfully outperformed traditional models in many studies that are too complicated to model with traditional methods, but struggled in others where straight relationships can be identified. Prediction accuracy of ten different statistical and ML genomic selection models were compared for 18 traits from 8 datasets [232]. Although their study was unable to identify an all-purpose model or combination of models, the authors recommended Bayesian Lasso, weighted Bayesian shrinkage regression, and random forest models. Two ML methods, multi-layer perceptron and support vector machine, are compared against the Bayesian and GBLUP models for predicting ordinal traits in plant breeding based on seven datasets. It's found the ML methods to be not as competitive as the Bayesian model and discussed the disadvantages of these methods that need to be overcome in future research [227].

Applications of DL for GP for single and multiple traits (continuous, binary, and ordinal traits under a univariate and multivariate framework) have shown that DL models provide predictions that are competitive with conventional statistical ML models, especially in cases where larger data sets are available. Nine multi-environment maize (309 lines across three environments) and wheat data sets (250-2,304 lines in different environments) from field breeding trials were analyzed using GBLUP and DL with and without G×E interaction [204]. The DL method was superior when the G×E interaction term was not included in the method under the grid of parameters implemented. This can be attributed to the fact that DL methods are capable of capturing complex relationships hidden in the data without requiring strong assumptions about the underlying mechanisms, which are frequently unknown or insufficiently defined. DL is a general-purpose approach for learning functional relationships from data that do not require any prior information, as do the GBLUP and other genomic Bayesian methods. Furthermore, continuous multi-trait deep learning (MTDL) models are compared with the BMTME model proposed in 2016 [197] with and without GxE [203]. Among models without GxE, the MTDL model was the best, while among models with GxE, the BMTME model was superior. General results indicated that the MTDL models are very competitive for performing predictions in the context of GS, with the important practical advantage that they require fewer computational resources than the BMTME model. However, it's pointed out two main disadvantages of DL: (a) it is hard to train a DL method, because it requires testing different combinations of hyperparameters corresponding to the number of layers, the number of units, the number of epochs, the type of regularization (and the dropout percentage in the context of dropout regularization), and the type of activation function in each layer; and (b) the computational time required to implement a DL method increases as the number of layers and units increases [204].

Despite their promising performances, ML models have also shown their limitations as a bottom-up approach. For example, ML requires large training datasets, and the prediction results are often inexplicable due to the black-box nature of bottom-up models. A more recent trend in the field of ML is explainable algorithms. The challenges faced by black box models were discussed

by including the numerous application domains where transparency is absolutely critical [233]. Moreover, some ML models are not only hard to explain but also easy to fool by noise or deliberately generated data. It's argued that even if it may not be feasible to convert ML models into fully transparent systems, partial insights into the interpretability and explainability should be developed [234]. Several recent studies have made noticeable progress in such a direction. Statistical approaches [235,236] have been proposed to interpret prediction results with respect to feature contributions. Emerging approaches and challenges for genomics and phenomics are presented using multi-omics big data integration [237]. They suggest that farmers, plant breeders, and ML researchers work together to build the next generation of ML platforms for plant breeding. Explainable ML models can be considered as a hybrid of the top-down and bottom-up approaches. They incorporate domain knowledge in the design of more complex modules from the bottom-up, still allowing the algorithm to learn from data, but guided by confirmed domain knowledge rather than less reliable human expert experiences. As such, results from these hybrid ML models are much more explainable, yet they could also become computationally more challenging due to complex modeling structures. Consequently, operations research and optimization algorithms have been combined with ML models to address such challenges in explainable ML.

One of the success stories of hybrid ML models is the Syngenta Crop Challenge in Analytics, which was inaugurated in 2016 and has become a unique competition that provides plant scientists and data scientists the opportunity to help solve some of the world's biggest agricultural challenges with access to real-world crop data. The majority of winning solutions have used hybrid models that integrate agronomic and biological knowledge with cutting-edge ML techniques. Their major properties are highlighted for the recent four years (2017-2020).

2017: A portfolio optimization model was designed for soybean seed selection, which was built on an ensemble of machine learning models for predicting the performance of seeds under diverse weather scenarios [238]. They conducted global sensitivity and uncertainty analyses and used the results to help select the optimal seed portfolio to balance expected return and risk.

2018: One of the first deep neural network models was proposed for predicting crop yield using genotype and environment data [239]. Computational results suggested that this model significantly outperformed other popular methods such as Lasso, shallow neural networks, and regression trees. As a follow-up, a convolutional neural network - recurrent neural network (CNN-RNN) model was designed for crop yield prediction [240]. The model was designed to capture the time dependencies of environmental factors and the genetic improvement of seeds over time without having their genotype information. It demonstrated the capability to generalize the yield prediction to untested environments without a significant drop in the prediction accuracy. Coupled with the backpropagation method, the model could reveal the extent to which

weather conditions, the accuracy of weather predictions, soil conditions, and management practices were able to explain the variation in the crop yields.

2019: A random forest model was used to predict the presence of drought and heat stress in corn hybrids [241]. To feed the random forest model with the most meaningful features, they calculated meteorological features for four maize growth stages. Such biological insights helped the model produce a more accurate prediction of yield performances across a range of environmental scenarios.

2020: A combinatorial optimization model was combined with a random forest for predicting the yield performance of crossing testers and inbreds in a plant breeding process [242]. The combinatorial optimization model was designed to detect genotype-by-environment interactions, and the random forest model was able to capture other types of linear and nonlinear effects.

Concluding Remarks

Genomics and phenomics have enhanced plant breeding for agronomic traits both directly and indirectly. On the genomics side, genetic mapping using GWAS directly aims to identify genetic loci linked to the causal mutations consequently used as criteria for selection. However, many of these genetic loci are not validated and remain unused, see Outstanding Questions. In contrast, genomic selection is widely used in breeding for agronomic traits. Genomic selection increases the frequencies of favorite alleles, even those specific favorable alleles and causal genes which are unknown, as is frequently the case for genes with small effects. Nevertheless, there are substantially more peer-reviewed publications using GWAS than GS, partially due to the discovery nature of GWAS versus the applied results of GS. Importantly, GWAS results can enhance GS, especially once a particular locus has been validated across multiple studies. Among the three major categories of GS methods (BLUP, Bayes, and ML), the best method varies across traits and datasets depending on genetic architecture, such as heritability, number of genes, and gene effect distribution. Within each of these categories, multiple methods have been developed for adaptation to the genetic architecture. On the phenomics side, many of the advances in phenomics technology have not significantly reduced the effort or cost of collecting the first data point, but instead dramatically reduced the marginal cost and effort required to collect additional measurements beyond the first. Multiple traits and multi-time point models can improve prediction accuracy through the integration of relationships among traits, including both traditional agronomic traits and high-throughput phenomic traits.

Outstanding Questions Box

- GBLUP is used much more widely for routine genomic evaluations than the Bayesian methods regardless of the advantage of the Bayesian methods on traits with specific genetic architecture and interaction with growing conditions such as locations and periods. Adjustments of Gibbs sampling or Metropolis-Hastings Markov Chain Monte Carlo approaches remain the most challenge to dramatically reduce computing times required for the Bayesian methods.
- It is an important and open question for a biological understanding of crops if segregating loci uniquely identified as contributing to different temporal growth stages are real, repeatable, and useful in understanding the biological basis of these traits, let alone in the context of genetic improvement.
- It is a statistical challenge to correct for multiple testing of hundreds to thousands of phenotypes to prevent false discoveries in the way that multiple testing of genetic markers has become routine in large p small n problems.
- It is an important open question how to validate and compare unique high-throughput phenotyping across environments. While crop modeling approaches may play a role, new approaches are needed to parameterize crop models using high throughput traits that cannot be manually validated.
- It is a practical question as to what phenomic selection approaches are capturing; is it plant/grain composition (e.g. protein vs. yield has a known tradeoff) which would allow such models to be more portable to other untested environments and populations, or is it genetic relationships, like genomic selection, which would result in the same limitations in predicting untested genotypes in untested environments that genomic selection has had.
- Predictions from bottom-up and top-down hybrid machine learning methods are much more explainable, yet they could also become computationally more challenging due to complex modeling structures.

Conflict of interest declared: None

References

1   Wallace, J.G. *et al.* (2018) On the Road to Breeding 4.0: Unraveling the Good, the Bad, and the Boring of Crop Quantitative Genomics. *Annu. Rev. Genet.* 52, 421–444

2   Furbank, R.T. and Tester, M. (2011) Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644

3   Rincent, R. *et al.* (2018) Phenomic Selection Is a Low-Cost and High-Throughput Method Based on Indirect Predictions: Proof of Concept on Wheat and Poplar. *G3 Genes, Genomes, Genet.* 8, 3961–3972

4   Lane, H.M. *et al.* (2020) Phenomic selection and prediction of maize grain yield from near-infrared reflectance spectroscopy of kernels. *Plant Phenome J.* 3, e20002

5   Lobell, D.B. *et al.* (2015) A scalable satellite-based crop yield mapper. *Remote Sens. Environ.* 164, 324–333

6   Acin-Albiac, M. *et al.* (2020) Microbial high throughput phenomics: The potential of an irreplaceable omics. *Comput. Struct. Biotechnol. J.* 18, 2290–2299

7   Panek, E. and Gozdowski, D. (2020) Analysis of relationship between cereal yield and NDVI for selected regions of Central Europe based on MODIS satellite data. *Remote Sens. Appl. Soc. Environ.* 17, 100286

8   Hong, J.-K. *et al.* (2021) Microbial phenomics linking the phenotype to fonction: The potential of Raman spectroscopy. *J. Microbiol.* 59, 249–258

9   Garg, M. *et al.* (2018) Biofortified Crops Generated by Breeding, Agronomy, and Transgenic Approaches Are Improving Lives of Millions of People around the World. *Front. Nutr.* 5, 12

10   Wallace, J.G. *et al.* (2018) On the road to breeding 4.0: Unraveling the good, the bad, and the boring of crop quantitative genomics. *Annu. Rev. Genet.* 52, 421–444

11   Krause, M.R. *et al.* (2019) Hyperspectral Reflectance-Derived Relationship Matrices for Genomic Prediction of Grain Yield in Wheat. *G3 Genes, Genomes, Genet.* 9, 1231–1247

12   Campbell, M.T. *et al.* (2020) Translating insights from the seed metabolome into improved prediction for healthful compounds in oat (Avena sativa L.). *bioRxiv* DOI: 10.1101/2020.07.06.190512

13   Tam, V. *et al.* (2019) Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 20, 467–484

14   Cantor, R.M. *et al.* (2010) Prioritizing GWAS Results: A Review of Statistical Methods and Recommendations for Their Application. *Am. J. Hum. Genet.* 86, 6–22

15   Sun, N. and Zhao, H. (2020) Statistical Methods in Genome-Wide Association Studies. *Annu. Rev. Biomed. Data Sci.* 3, 265–288

16   Tibbs Cortes, L. *et al.* (2021) Status and prospects of genome-wide association studies in plants. *Plant Genome* n/a, e20077

17   Teo, Y.Y. (2008) Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. *Curr Opin Lipidol* 19, 133–143

18   Neeraja, C.N. *et al.* (2007) A marker-assisted backcross approach for developing submergence-tolerant rice cultivars. *Theor. Appl. Genet.* 115, 767–776

19   Xu, Y. *et al.* (2019) Enhancing Genetic Gain through Genomic Selection: From Livestock to Plants. *Plant Commun.* 1, 100005

20   Weigel, K.A. *et al.* (2017) A 100-Year Review: Methods and impact of genetic selection in dairy cattle—From daughter–dam comparisons to deep learning algorithms. *J. Dairy*

Sci. 100, 10234–10250

21   Henderson, C.R. (1963) Selection index and expected genetic advance. *Stat. Genet. plant breeding. Hanson, W.D., Robinson, H.F. NAS-NRC, Washington, DC. Pub. No 982*

22   Henderson, C.R. (1973) SIRE EVALUATION AND GENETIC TRENDS. *J. Anim. Sci.* 1973, 10–41

23   Henderson, C.R. (1975) Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics* 31, 423–447

24   Madsen, P. *et al.* (2006), DMU - A Package for Analyzing Multivariate Mixed Models in Quantitative Genetics and Genomics.

25   Boldman, K.G. *et al.* (1993) A manual for use of MTDFREML. *USDA-ARS, Clay Center, Nebraska*

26   Fernando, R.L. and Grossman, M. (1989) Marker assisted selection using best linear unbiased prediction. *Genet. Sel. Evol* 21, 467–477

27   Bernardo, R. (1994) Prediction of Maize Single-Cross Performance Using RFLPs and Information from Related Hybrids. *Crop Sci.* 34, 20–25

28   Nejati-Javaremi, A. *et al.* (1997) Effect of total allelic relationship on accuracy of evaluation and response to selection. *J Anim Sci* 75, 1738–1745

29   Zhang, Z. *et al.* (2007) Technical note: Use of marker-based relationships with multiple-trait derivative-free restricted maximal likelihood. *J. Anim. Sci.* 85, 881–885

30   VanRaden, P.M. (2008) Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423

31   García-Ruiz, A. *et al.* (2016) Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci.* 113, E3995–E4004

32   Grattapaglia, D. *et al.* (2018) Quantitative genetics and genomics converge to accelerate forest tree breeding. *Front. Plant Sci.* 871, 1–10

33   Stich, B. and Inghelandt, D. Van (2018) Prospects and potential uses of genomic prediction of key performance traits in tetraploid potato. *Front. Plant Sci.*

34   Massman, J.M. *et al.* (2013) Genomewide Selection versus Marker-assisted Recurrent Selection to Improve Grain Yield and Stover-quality Traits for Cellulosic Ethanol in Maize. *Crop Sci.* 53, 58–66

35   Beyene, Y. *et al.* (2015) Genetic Gains in Grain Yield Through Genomic Selection in Eight Bi-parental Maize Populations under Drought Stress. *Crop Sci.* 55, 154–163

36   Zhang, X. *et al.* (2017) High-throughput phenotyping and QTL mapping reveals the genetic architecture of maize plant growth. *Plant Physiol.* 173, 1554–1564

37   Meuwissen, T.H. *et al.* (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829

38   Wang, T. *et al.* (2015) A computationally efficient algorithm for genomic prediction using a Bayesian model. *Genet. Sel. Evol.* 47, 34

39   Gianola, D. *et al.* (2009) Additive genetic variability and the Bayesian alphabet. *Genetics* 183, 347–363

40   Henderson, C.R. (1976) Simple Method for Computing Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. *Biometrics* 32, 69–83

41   Henderson, C.R. (1975) Rapid Method for Computing the Inverse of a Relationship Matrix. *J. Dairy Sci.* 58, 1727–1730

42   Zhang, Z. *et al.* (2010) Mixed linear model approach adapted for genome-wide association

studies. *Nat. Genet.* 42, 355–60

43   Wang, X. *et al.* (2018) ScienceDirect Genomic selection methods for crop improvement : Current status and prospects. *Crop J.* 6, 330–340

44   Verbyla, K.L. *et al.* (2009) Accuracy of genomic selection using stochastic search variable selection in Australian holstein friesian dairy cattle. *Genet. Res. (Camb).* 91, 307–311

45   Daetwyler, H.D. *et al.* (2010) The impact of genetic architecture on genome-wide evaluation methods. *Genetics* 185, 1021–1031

46   Wei Huang *et al.* (2020) MMAP : a cloud computing platform for mining the maximum accuracy of predicting phenotypes from genotypes. *BIOINFORMATICS* btaa824, 11–14

47   Krause, M.R. *et al.* (2020) Aerial high-throughput phenotyping enables indirect selection for grain yield at the early generation, seed-limited stages in breeding programs. *Crop Sci.* 60, 3096–3114

48   Sagan, V. *et al.* (2019) UAV-Based High Resolution Thermal Imaging for Vegetation Monitoring, and Plant Phenotyping Using ICI 8640 P, FLIR Vue Pro R 640, and thermoMap Cameras. *Remote Sens.* 11,

49   Li, M. *et al.* (2020) Comprehensive 3D phenotyping reveals continuous morphological variation across genetically diverse sorghum inflorescences. *New Phytol.* 226, 1873–1885

50   Wang, X. *et al.* (2019) High-throughput phenotyping with deep learning gives insight into the genetic architecture of flowering time in wheat. *Gigascience* 8,

51   Araus, J.L. *et al.* (2018) Translating High-Throughput Phenotyping into Genetic Gain. *Trends Plant Sci.* 23, 451–466

52   Yang, W. *et al.* (2014) Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* 5,

53   Junker, A. *et al.* (2015) Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. *Front. Plant Sci.* 5, 1–21

54   Ge, Y. *et al.* (2016) Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Comput. Electron. Agric.* 127, 625–632

55   Virlet, N. *et al.* (2017) Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. *Funct. Plant Biol.* 44, 143

56   Burnette, M. *et al.* (2018), TERRA-REF Data Processing Infrastructure., pp. 1–7

57   Bai, G. *et al.* (2019) NU-Spidercam: A large-scale, cable-driven, integrated sensing and robotic system for advanced phenotyping, remote sensing, and agronomic research. *Comput. Electron. Agric.* 160, 71–81

58   Kirchgessner, N. *et al.* (2017) The ETH field phenotyping platform FIP: A cable-suspended multi-sensor system. *Funct. Plant Biol.* 44, 154–168

59   Cai, H. *et al.* (2012) Mapping QTLs for root system architecture of maize (Zea mays L.) in the field at different developmental stages. *Theor. Appl. Genet.* 125, 1313–1324

60   Das, A. *et al.* (2015) Digital imaging of root traits (DIRT): A high-throughput computing and collaboration platform for field-based root phenomics. *Plant Methods* 11,

61   Zheng, Z. *et al.* (2020) Shared Genetic Control of Root System Architecture between Zea mays and Sorghum bicolor. *Plant Physiol.* 182, 977–991

62   Yendrek, C.R. *et al.* (2017) High-Throughput Phenotyping of Maize Leaf Physiological and Biochemical Traits Using Hyperspectral Reflectance. *Plant Physiol.* 173, 614–626

63   Silva-Perez, V. *et al.* (2018) Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *J. Exp. Bot.* 69, 483–496

64   Ge, Y. *et al.* (2019) High-throughput analysis of leaf physiological and chemical traits with {VIS}–{NIR}–{SWIR} spectroscopy: a case study with a maize diversity panel. *Plant Methods* 15, 66

65   Vergara-Diaz, O. *et al.* (2020) Assessing durum wheat ear and leaf metabolomes in the field through hyperspectral data. *Plant J.* 102, 615–630

66   Wang, S. *et al.* (2021) Unique contributions of chlorophyll and nitrogen to predict crop photosynthetic capacity from leaf spectroscopy. *J. Exp. Bot.* 72, 341–354

67   Cotrozzi, L. *et al.* (2020) Spectral Phenotyping of Physiological and Anatomical Leaf Traits Related with Maize Water Status. *Plant Physiol.* 184, 1363–1377

68   Bao, Y. *et al.* (2018) Field-based Robot Phenotyping of Sorghum Plant Architecture using Stereo Vision. *J. F. Robot.* DOI: 10.1002/rob.21830

69   Sun, S. *et al.* (2018) In-field High Throughput Phenotyping and Cotton Plant Growth Analysis Using LiDAR. *Front. Plant Sci.* 9,

70   Higuti, V.A.H. *et al.* (2019) Under canopy light detection and ranging-based autonomous navigation. *J. F. Robot.* 36, 547–567

71   Stager, A. *et al.* (2019) Design and Construction of Unmanned Ground Vehicles for Sub-Canopy Plant Phenotyping. *arXiv1903.10608 [cs]*

72   Mantilla-Perez, M.B. *et al.* (2020) Towards "smart canopy" sorghum: discovery of the genetic control of leaf angle across layers. *Plant Physiol.* DOI: 10.1104/pp.20.00632

73   Ubbens, J. *et al.* (2020) Latent Space Phenotyping: Automatic Image-Based Phenotyping for Treatment Studies. *Plant Phenomics* 2020, 1–13

74   Gage, J.L. *et al.* (2019) In-Field Whole-Plant Maize Architecture Characterized by Subcanopy Rovers and Latent Space Phenotyping. *Plant Phenome J.* 2, 190011

75   Sankaran, S. *et al.* (2015) Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: {A} review.

76   Shi, Y. *et al.* (2016) Unmanned Aerial Vehicles for High-Throughput Phenotyping and Agronomic Research. *PLoS One* 11, e0159781

77   Pugh, N.A. *et al.* (2018) Temporal Estimates of Crop Growth in Sorghum and Maize Breeding Enabled by Unmanned Aerial Systems. *Plant Phenome J.* 1, 170006

78   Han, X. *et al.* (2018), Calibrated plant height estimates with structure from motion from fixed-wing {UAV} images., in *Autonomous {Air} and {Ground} {Sensing} {Systems} for {Agricultural} {Optimization} and {Phenotyping} {III}*, 10664, pp. 106640D

79   Kawamura, K. *et al.* (2020) Field phenotyping of plant height in an upland rice field in Laos using low-cost small unmanned aerial vehicles (UAVs). *Plant Prod. Sci.* 23, 452–465

80   Wang, X. *et al.* (2019) Dynamic plant height QTL revealed in maize through remote sensing phenotyping using a high-throughput unmanned aerial vehicle (UAV). *Sci. Rep.* 9, 3458

81   Anderson, S.L. *et al.* (2020) Unoccupied aerial system enabled functional modeling of maize height reveals dynamic expression of loci. *Plant Direct* 4,

82   Jin, Z. *et al.* (2017) Crop model- and satellite imagery-based recommendation tool for variable rate N fertilizer application for the US Corn system. *Precis. Agric.* 18, 779–800

83   Price, J.C. (1993) Estimating leaf area index from satellite data. *IEEE Trans. Geosci. Remote Sens.* 31, 727–734

84   Zhang, C. *et al.* (2020) High-resolution satellite imagery applications in crop phenotyping : An overview. *Comput. Electron. Agric.* 175, 105584

85   Yang, C. (2018) High resolution satellite imaging sensors for precision agriculture. *Front. Agric. Sci. Eng.* 0, 0

86   Tattaris, M. *et al.* (2016) A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding. *Front. Plant Sci.* 7,

87   Sankaran, S. *et al.* (2019) Unmanned aerial system and satellite-based high resolution imagery for high-throughput phenotyping in dry bean. *Comput. Electron. Agric.* 165, 104965

88   Sankaran, S. *et al.* (2021) Can High-Resolution Satellite Multispectral Imagery Be Used to Phenotype Canopy Traits and Yield Potential in Field Conditions? *Trans. ASABE* 0, 0

89   Liang, Z. *et al.* (2018) Conventional and hyperspectral time-series imaging of maize lines widely used in field trials. *Gigascience* 7, 1–11

90   Moore, C.R. *et al.* (2013) High-Throughput Computer Vision Introduces the Time Axis to a Quantitative Trait Map of a Plant Growth Response. *Genetics* 195, 1077–1086

91   Miao, C. *et al.* (2020) Increased Power and Accuracy of Causal Locus Identification in Time Series Genome-wide Association in Sorghum. *Plant Physiol.* 183, 1898–1909

92   Jin, X. *et al.* (2020) High-throughput estimation of crop traits: A review of ground and aerial phenotyping platforms. *IEEE Geosci. Remote Sens. Mag.* DOI: 10.1109/MGRS.2020.2998816

93   Jiang, Y. and Li, C. (2020) Convolutional Neural Networks for Image-Based High-Throughput Plant Phenotyping: A Review. *Plant Phenomics* 2020, 1–22

94   Carmo-Silva, E. *et al.* (2017) Phenotyping of field-grown wheat in the UK highlights contribution of light response of photosynthesis and flag leaf longevity to grain yield. *J. Exp. Bot.* 68, 3473–3486

95   Jansen, M. *et al.* (2009) Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREENFLUORO allows detection of stress tolerance in. *Funct. PLANT Biol.* 36, 902–914

96   Thoren, D. and Schmidhalter, U. (2009) Nitrogen status and biomass determination of oilseed rape by laser-induced chlorophyll fluorescence. *Eur. J. Agron.* 30, 238–242

97   Sharma, D. *et al.* (2012) Phenotyping of wheat cultivars for heat tolerance using chlorophyll a fluorescence. *Funct. Plant Biol.* 39, 936–947

98   Li, W. *et al.* (2016) Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecol. Indic.* 67, 637–648

99   Biskup, B. *et al.* (2007) A stereo imaging system for measuring structural parameters of plant canopies. *Plant. Cell Environ.* 30, 1299–1308

100   Hosoi, F. and Omasa, K. (2009) Estimating vertical plant area density profile and growth parameters of a wheat canopy at different growth stages using three-dimensional portable lidar imaging. *ISPRS J. Photogramm. Remote Sens.* 64, 151–158

101   Hui, F. *et al.* (2018) Image-based dynamic quantification and high-accuracy 3D evaluation of canopy structure of plant populations. *Ann. Bot.* 121, 1079–1088

102   Jimenez-Berni, J.A. *et al.* (2018) High Throughput Determination of Plant Height, Ground Cover, and Above-Ground Biomass in Wheat with LiDAR. *Front. Plant Sci.* 9,

103   Thapa, S. *et al.* (2018) A Novel LiDAR-Based Instrument for High-Throughput, 3D Measurement of Morphological Traits in Maize and Sorghum. *Sensors* 18, 1187

104   Jin, X. *et al.* (2017) Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sens. Environ.* 198, 105–114

105   Makanza, R. *et al.* (2018) High-Throughput Phenotyping of Canopy Cover and Senescence in Maize Field Trials Using Aerial Digital Canopy Imaging. *Remote Sens.* 10, 330

106   Bendig, J. *et al.* (2015) Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinf.* 39, 79–87

107   Baret, F. *et al.* (2018) Leaf-rolling in maize crops: from leaf scoring to canopy-level measurements for phenotyping. *J. Exp. Bot.* 69, 2705–2716

108   Jay, S. *et al.* (2017) Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. *F. Crop. Res.* 210, 33–46

109   Baresel, J.P. *et al.* (2017) Use of a digital camera as alternative method for non-destructive detection of the leaf chlorophyll content and the nitrogen nutrition status in wheat. *Comput. Electron. Agric.* 140, 25–33

110   Roosjen, P.P.J. *et al.* (2018) Improved estimation of leaf area index and leaf chlorophyll content of a potato crop using multi-angle spectral data – potential of unmanned aerial vehicle imagery. *Int. J. Appl. Earth Obs. Geoinf.* 66, 14–26

111   Amani, I. *et al.* (1996) Evaluation of canopy temperature as a screening tool for heat tolerance in spring wheat. *J. Agron. Crop Sci.* 176, 119–129

112   Liu, T. *et al.* (2018) Estimates of rice lodging using indices derived from UAV visible and thermal infrared images. *Agric. For. Meteorol.* 252, 144–154

113   Comar, A. *et al.* (2012) A semi-automatic system for high throughput phenotyping wheat cultivars in-field conditions: description and first results. *Funct. Plant Biol.* 39, 914–924

114   Chen, C.J. and Zhang, Z. (2020) GRID: A Python Package for Field Plot Phenotyping Using Aerial Images. *Remote Sens.* 12,

115   Matias, F.I. *et al.* (2020) FIELDimageR: An R package to analyze orthomosaic images from agricultural field trials. *Plant Phenome J.* 3, e20005

116   Fernandez-Gallego, J.A. *et al.* (2018) Wheat ear counting in-field conditions: high throughput and low-cost approach using RGB images. *Plant Methods* 14, 22

117   Schwalbert, R.A. *et al.* (2020) Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. *Agric. For. Meteorol.* 284, 107886

118   Jin, X. *et al.* (2014) Newly Combined Spectral Indices to Improve Estimation of Total Leaf Chlorophyll Content in Cotton. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 4589–4600

119   Madec, S. *et al.* (2019) Ear density estimation from high resolution RGB imagery using deep learning technique. *Agric. For. Meteorol.*

120   Chu, T. *et al.* (2017) Assessing Lodging Severity over an Experimental Maize (Zea mays L.) Field Using UAS Images. *Remote Sens.* 9, 923

121   Rose, J.C. *et al.* (2015) Accuracy Analysis of a Multi-View Stereo Approach for Phenotyping of Tomato Plants at the Organ Level. *Sensors* 15, 9651–9665

122   Chang, S. *et al.* (2020) High-Throughput Phenotyping (HTP) Data Reveal Dosage Effect at Growth Stages in Arabidopsis thaliana Irradiated by Gamma Rays. *Plants* 9,

123   Mazis, A. *et al.* (2020) Application of high-throughput plant phenotyping for assessing biophysical traits and drought response in two oak species under controlled environment. *For. Ecol. Manage.* 465, 118101

124　Patrick, A. and Li, C. (2017) High Throughput Phenotyping of Blueberry Bush Morphological Traits Using Unmanned Aerial Systems. *Remote Sens.* 9, 1250

125　Tolley, S. *et al.* (2020) High-throughput phenotyping identifies plant growth differences under well-watered and drought treatments. *J. Integr. Agric.* 19, 2429–2438

126　Das Choudhury, S. *et al.* (2016), Automated Vegetative Stage Phenotyping Analysis of Maize Plants using Visible Light Images DS-FEW '.

127　Anandan, A. *et al.* (2020) Non-destructive phenotyping for early seedling vigor in direct-seeded rice. *Plant Methods* 16, 127

128　Kim, S.L. *et al.* (2020) High-throughput phenotyping platform for analyzing drought tolerance in rice. *Planta* 252, 38

129　Xiong, X. *et al.* (2017) A high-throughput stereo-imaging system for quantifying rape leaf traits during the seedling stage. *Plant Methods* 13, 7

130　Campbell, M.T. *et al.* (2021) Translating insights from the seed metabolome into improved prediction for lipid-composition traits in oat (Avena sativa L.). *Genetics* DOI: 10.1093/genetics/iyaa043

131　Ling, W. *et al.* (2018) Materials and Techniques for Implantable Nutrient Sensing Using Flexible Sensors Integrated with Metal–Organic Frameworks. *Adv. Mater.* 30, 1800917

132　Giraldo, J.P. *et al.* (2019) Nanobiotechnology approaches for engineering smart plant sensors. *Nat. Nanotechnol.* 14, 541–553

133　Oren, S. *et al.* (2017) High-Resolution Patterning and Transferring of Graphene-Based Nanomaterials onto Tape toward Roll-to-Roll Production of Tape-Based Wearable Sensors. *Adv. Mater. Technol.* 2, 1700223

134　Han, S.-T. *et al.* (2017) An Overview of the Development of Flexible Sensors. *Adv. Mater.* 29, 1700375

135　Byun, S.-H. *et al.* (2019) Mechanically transformative electronics, sensors, and implantable devices. *Sci. Adv.* 5, eaay0418

136　Fichman, Y. *et al.* (2019) Whole-Plant Live Imaging of Reactive Oxygen Species. *Mol. Plant* 12, 1203–1210

137　Lew, T.T.S. *et al.* (2020) Real-time detection of wound-induced H 2 O 2 signalling waves in plants with optical nanosensors. *Nat. Plants* 6, 404–415

138　Deuschle, K. *et al.* (2006) Rapid metabolism of glucose detected with FRET glucose nanosensors in epidermal cells and intact roots of Arabidopsis RNA-silencing mutants. *Plant Cell* 18, 2314–2325

139　Zhu, Q. *et al.* (2017) FRET-based glucose imaging identifies glucose signalling in response to biotic and abiotic stresses in rice roots. *J. Plant Physiol.* 215, 65–72

140　Chaudhuri, B. *et al.* (2008) Protonophore- and pH-insensitive glucose and sucrose accumulation detected by FRET nanosensors in Arabidopsis root tips. *Plant J. Cell Mol. Biol.* 56, 948–962

141　Krebs, M. *et al.* (2012) FRET-based genetically encoded sensors allow high-resolution live cell imaging of Ca2+ dynamics. *Plant J.* 69, 181–192

142　Matsuda, T. *et al.* (2013) Highlighted Ca 2+ imaging with a genetically encoded 'caged' indicator. *Sci. Rep.* 3, 1398

143　Giraldo, J.P. *et al.* (2015) A Ratiometric Sensor Using Single Chirality Near-Infrared Fluorescent Carbon Nanotubes: Application to In Vivo Monitoring. *Small* 11, 3973–3984

144　Esser, B. *et al.* (2012) Selective Detection of Ethylene Gas Using Carbon Nanotube-based Devices: Utility in Determination of Fruit Ripeness. *Angew. Chemie Int. Ed.* 51, 5752–

5756

145   Larrieu, A. *et al.* (2015) A fluorescent hormone biosensor reveals the dynamics of jasmonate signalling in plants. *Nat. Commun.* 6, 6043

146   Li, Y. *et al.* (2018) A reagent-assisted method in SERS detection of methyl salicylate. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* 195, 172–175

147   Jones, A.M. *et al.* (2014) Abscisic acid dynamics in roots detected with genetically encoded FRET sensors. *Elife* 3, e01741

148   Waadt, R. *et al.* (2014) FRET-based reporters for the direct visualization of abscisic acid concentration changes and distribution in Arabidopsis. *Elife* 3, e01739

149   Fasano, J.M. *et al.* (2001) Changes in Root Cap pH Are Required for the Gravity Response of the Arabidopsis Root. *Plant Cell* 13, 907–921

150   Pauli, D. *et al.* (2016) Field-based high-throughput plant phenotyping reveals the temporal patterns of quantitative trait loci associated with stress-responsive traits in cotton. *G3 Genes, Genomes, Genet.* 6, 865–879

151   Shakoor, N. *et al.* (2019) Genomewide association study reveals transient loci underlying the genetic architecture of biomass accumulation under cold stress in Sorghum. *BioRxiv*

152   Lyra, D.H. *et al.* (2020) Functional QTL mapping and genomic prediction of canopy height in wheat measured using a robotic field phenotyping platform. *J. Exp. Bot.* 71, 1885–1898

153   Anderson, S.L. *et al.* (2019) Prediction of maize grain yield before maturity using improved temporal height estimates of unmanned aerial systems. *Plant Phenome J.* 2, 1–15

154   Rasheed, A. *et al.* (2014) Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. *BMC Plant Biol.* 14, 128

155   York, L.M. (2019) Functional phenomics: an emerging field integrating high-throughput phenotyping, physiology, and bioinformatics. *J. Exp. Bot.* 70, 379–386

156   Murray, S.C. *et al.* (2008) Genetic improvement of sorghum as a biofuel feedstock: II. QTL for stem and leaf structural carbohydrates. *Crop Sci.* 48, 2180–2193

157   Thoen, M.P.M. *et al.* (2017) Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping. *New Phytol.* 213, 1346–1362

158   Reymond, M. *et al.* (2003) Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiol.* 131, 664–675

159   Li, X. *et al.* (2018) Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proc. Natl. Acad. Sci. U. S. A.* 115, 6679–6684

160   Hammer, G. *et al.* (2006) Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci.* 11, 587–593

161   Muller, B. and Martre, P. (2019) Plant and crop simulation models: powerful tools to link physiology, genetics, and phenomics. *J. Exp. Bot.* 70, 2339–2344

162   Messina, C.D. *et al.* (2018) Leveraging biological insight and environmental variation to improve phenotypic prediction: Integrating crop growth models (CGM) with whole genome prediction (WGP). *Eur. J. Agron.* 100, 151–162

163   Sun, J. *et al.* (2017) Multitrait, Random Regression, or Simple Repeatability Model in High-Throughput Phenotyping Data Improve Genomic Prediction for Wheat Grain Yield. *Plant Genome* 10,

164   Feldmann, M.J. *et al.* (2020) Multi-dimensional machine learning approaches for fruit

shape phenotyping in strawberry. *Gigascience* 9, giaa030

165   Churchill, G.A. and Doerge, R.W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138, 963–971

166   Dormann, C.F. *et al.* (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography (Cop.).* 36, 27–46

167   Brzyski, D. *et al.* (2017) Controlling the Rate of GWAS False Discoveries. *Genetics* 205, 61–75

168   Ramstein, G.P. *et al.* (2019) Breaking the curse of dimensionality to identify causal variants in Breeding 4. *Theor. Appl. Genet.* 132, 559–567

169   Yurko, R. *et al.* (2020) A selective inference approach for false discovery rate control using multiomics covariates yields insights into disease risk. *Proc. Natl. Acad. Sci.*

170   Murray, S.C. and Wisser, R.J. (2014) Genetic Inference on Quantitative Traits Through Linkage and Association Studies. In *Genetics, Genomics and Breeding of Maize* pp. 91–115, CRC Press

171   Calus, M.P.L. and Veerkamp, R.F. (2011) Accuracy of multi-trait genomic selection using different methods. *Genet. Sel. Evol.* 43, 26

172   Hernandez, C.O. *et al.* (2020) Genomic Prediction and Selection for Fruit Traits in Winter Squash. *G3 Genes, Genomes, Genet.* 10, 3601–3610

173   Mrode, R.A. (2014) *Linear Models for the Prediction of Animal Breeding Values: 3rd Edition*, CABI.

174   Guo, G. *et al.* (2014) Comparison of single-trait and multiple-trait genomic prediction models. *BMC Genet.* 15, 1–7

175   Pszczola, M. *et al.* (2013) Effect of predictor traits on accuracy of genomic breeding values for feed intake based on a limited cow reference population. *Anim. An Int. J. Anim. Biosci.* 7, 1759–1768

176   Rutkoski, J. *et al.* (2016) Canopy Temperature and Vegetation Indices from High-Throughput Phenotyping Improve Accuracy of Pedigree and Genomic Selection for Grain Yield in Wheat. *G3 Genes|Genomes|Genetics* 6, 2799–2808

177   Lado, B. *et al.* (2018) Resource allocation optimization with multi-trait genomic prediction for bread wheat (Triticum aestivum L.) baking quality. *Theor. Appl. Genet.* 131, 2719–2731

178   Fernandes, S.B. *et al.* (2018) Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. *Theor. Appl. Genet.* 131, 747–755

179   Ward, B.P. *et al.* (2019) Multienvironment and Multitrait Genomic Selection Models in Unbalanced Early-Generation Wheat Yield Trials. *Crop Sci.* 59, 491–507

180   Bhatta, M. *et al.* (2020) Multi-trait Genomic Prediction Model Increased the Predictive Ability for Agronomic and Malting Quality Traits in Barley (Hordeum vulgare L.). *G3 (Bethesda).* 10, 1113–1124

181   Okeke, U.G. *et al.* (2017) Accuracies of univariate and multivariate genomic prediction models in African cassava. *Genet. Sel. Evol.* 49,

182   Schulthess, A.W. *et al.* (2016) Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. *Theor. Appl. Genet.* 129, 273–287

183   Hayes, B.J. *et al.* (2017) Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. *Theor. Appl. Genet.* 130, 2505–2519

184   Watson, A. *et al.* (2019) Multivariate Genomic Selection and Potential of Rapid Indirect Selection with Speed Breeding in Spring Wheat. *Crop Sci.* 59, 1945–1959

185   Momen, M. *et al.* (2019) Predicting Longitudinal Traits Derived from High-Throughput Phenomics in Contrasting Environments Using Genomic Legendre Polynomials and B-Splines. *G3 Genes, Genomes, Genet.* 9, 3369–3380

186   Sun, J. *et al.* (2019) High-throughput phenotyping platforms enhance genomic selection for wheat grain yield across populations and cycles in early stage. *Theor. Appl. Genet.* 132, 1705–1720

187   Crain, J. *et al.* (2018) Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding. *Plant Genome* 11, 170043

188   Wang, X. *et al.* (2017) Predicting rice hybrid performance using univariate and multivariate GBLUP models based on North Carolina mating design II. *Hered.* 118, 302–310

189   Galán, R.J. *et al.* (2020) Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye. *Theor. Appl. Genet.* DOI: 10.1007/s00122-020-03651-8

190   Whalen, A. *et al.* (2020) Using high-throughput phenotypes to enable genomic selection by inferring genotypes. *bioRxiv* DOI: 10.1101/2020.02.28.969600

191   Runcie, D. and Cheng, H. (2019) Pitfalls and Remedies for Cross Validation with Multi-trait Genomic Prediction Methods. *G3 Genes|Genomes|Genetics* 9, 3727–3741

192   Zhou, Y. *et al.* (2016) Systematic bias of correlation coefficient may explain negative accuracy of genomic prediction. *Briefings Bioinforma.* at <http://bib.oxfordjournals.org/content/early/2016/07/18/bib.bbw064.abstract>

193   Olkin, I. and Pratt, J.W. (1958) Unbiased Estimation of Certain Correlation Coefficients. *Ann. Math. Stat.* 29, 201–211

194   Habier, D. *et al.* (2011) Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics* 12, 186

195   Park, T. and Casella, G. (2008) The Bayesian Lasso. *J. Am. Stat. Assoc.* 103, 681–686

196   Erbe, M. *et al.* (2012) Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *J. Dairy Sci.* 95, 4114–4129

197   Montesinos-López, O.A. *et al.* (2016) A Genomic Bayesian Multi-trait and Multi-environment Model. *G3 Genes|Genomes|Genetics* 6, 2725 LP – 2744

198   Zhang, H. *et al.* (2019) Factors Affecting the Accuracy of Genomic Selection for Agricultural Economic Traits in Maize, Cattle, and Pig Populations. *Front. Genet.* 10, 189

199   de Oliveira, A.A. *et al.* (2018) Genomic prediction applied to high-biomass sorghum for bioenergy production. *Mol. Breed.* 38,

200   Shikha, M. *et al.* (2017) Genomic Selection for Drought Tolerance Using Genome-Wide SNPs in Maize. *Front. Plant Sci.* 8,

201   Daetwyler, H.D. *et al.* (2014) Genomic prediction for rust resistance in diverse wheat landraces. *Theor. Appl. Genet.* 127, 1795–1803

202   Montesinos-López, O.A. *et al.* (2019) An R Package for Bayesian Analysis of Multi-environment and Multi-trait Multi-environment Data for Genome-Based Prediction. *G3 Genes, Genomes, Genet.* 9, 1355–1369

203   Montesinos-López, O.A. *et al.* (2018) Multi-trait, Multi-environment Deep Learning

Modeling for Genomic-Enabled Prediction of Plant Traits. *G3 Genes|Genomes|Genetics* 8, 3829 LP – 3840

204 Montesinos-López, A. *et al.* (2018) Multi-environment Genomic Prediction of Plant Traits Using Deep Learners With Dense Architecture. *G3 Genes|Genomes|Genetics* 8, 3813 LP – 3828

205 Kemper, K.E. *et al.* (2015) Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. *Genet. Sel. Evol.* 47, 29

206 MacLeod, I.M. *et al.* (2016) Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* 17, 144

207 Warburton, C.L. *et al.* (2020) Use of whole-genome sequence data and novel genomic selection strategies to improve selection for age at puberty in tropically-adapted beef heifers. *Genet. Sel. Evol.* 52, 28

208 Pérez, P. and De Los Campos, G. (2008), BGLR: A Statistical Package for Whole Genome Regression and Prediction., *https://cran.r-project.org/web/packages/BGLR/vignettes/BGLR-extdoc.pdf*.

209 Lloyd-Jones, L.R. *et al.* (2019) Improved polygenic prediction by Bayesian multiple regression on summary statistics. *Nat. Commun.* 10, 5086

210 Wang, T. *et al.* (2017) Application of a Bayesian non-linear model hybrid scheme to sequence data for genomic prediction and QTL mapping. *BMC Genomics* 18, 618

211 Montesinos-López, O.A. *et al.* (2017) A Variational Bayes Genomic-Enabled Prediction Model with Genotype × Environment Interaction. *G3 (Bethesda).* 7, 1833–1853

212 Ormerod, J.T. and Wand, M.P. (2012) Gaussian Variational Approximate Inference for Generalized Linear Mixed Models. *J. Comput. Graph. Stat.* 21, 2–17

213 Yin, L. *et al.* (2020) KAML: improving genomic prediction accuracy of complex traits using machine learning determined parameters. *Genome Biol.* 21, 146

214 Fikere, M. *et al.* (2018) Genomic Prediction Using Prior Quantitative Trait Loci Information Reveals a Large Reservoir of Underutilised Blackleg Resistance in Diverse Canola (Brassica napus L.) Lines. *Plant Genome* 11, 170100

215 Tu, X. *et al.* (2020) Reconstructing the maize leaf regulatory network using ChIP-seq data of 104 transcription factors. *Nat. Commun.* 11, 5089

216 Ramstein, G.P. *et al.* (2020) Dominance Effects and Functional Enrichments Improve Prediction of Agronomic Traits in Hybrid Maize. *Genetics* 215, 215–230

217 Ibba, M.I. *et al.* (2020) Genome-based prediction of multiple wheat quality traits in multiple years. *Plant Genome* 13, e20034

218 Kemper, K.E. *et al.* (2018) A multi-trait Bayesian method for mapping QTL and genomic prediction. *Genet. Sel. Evol.* 50, 10

219 Montesinos-López, O.A. *et al.* (2017) Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods* 13, 4

220 Aguate, F.M. *et al.* (2017) Use of Hyperspectral Image Data Outperforms Vegetation Indices in Prediction of Maize Yield. *Crop Sci.* 57, 2517–2524

221 Montesinos-López, A. *et al.* (2017) Genomic Bayesian functional regression models with interactions for predicting wheat grain yield using hyper-spectral image data. *Plant Methods* 13, 62

222 Gianola, D. *et al.* (2006) Genomic-assisted prediction of genetic value with semiparametric procedures. *Genetics* 173, 1761–1776

223 Long, N. *et al.* (2011) Application of support vector regression to genome-assisted prediction of quantitative. *Theor. Appl. Genet.* 123, 1065–1074

224 González-Recio, O. and Forni, S. (2011) Genome-wide prediction of discrete traits using bayesian regressions and machine learning. DOI: 10.1186/1297-9686-43-7

225 Ehret, A. *et al.* (2015) Application of neural networks with back-propagation to genome-enabled prediction of complex traits in Holstein-Friesian and German Fleckvieh cattle. *Genet. Sel. Evol.* 47, 22

226 Bellot, P. *et al.* (2018) Can deep learning improve genomic prediction of complex human traits? *Genetics* 210, 809–819

227 Montesinos-López, O.A. *et al.* (2019) A Benchmarking Between Deep Learning, Support Vector Machine and Bayesian Threshold Best Linear Unbiased Prediction for Predicting Ordinal Traits in Plant Breeding. *G3 Genes|Genomes|Genetics* 9, 601 LP – 618

228 Montesinos-López, O.A. *et al.* (2019) New Deep Learning Genomic-Based Prediction Model for Multiple Traits with Binary, Ordinal, and Continuous Phenotypes. *G3 Genes|Genomes|Genetics* 9, 1545 LP – 1556

229 Sandhu, K. *et al.* (2021) Deep Learning for Predicting Complex Traits in Spring Wheat Breeding Program. *Front. Plant Sci.* 11, 613325

230 Kohonen, T. (2001) *Self-Organizing Maps*, (3rd edn) Springer-Verlag.

231 Lewis, N.D. (2016) *Deep Learning Made Easy with R: A Gentle Introduction For Data Science*, CreateSpace Independent Publishing Platform.

232 Heslot, N. *et al.* (2012) Genomic Selection in Plant Breeding: A Comparison of Models. *Crop Sci.* 52, 146

233 Castelvecchi, D. (2016) Can we open the black box of {AI}? *Nat. News* 538, 20

234 Emmert-Streib, F. *et al.* (2020) Artificial Intelligence: A Clarification of Misconceptions, Myths and Desired Status. *Front. Artif. Intell.* 3,

235 Lundberg, S. and Lee, S.-I. (2017) A Unified Approach to Interpreting Model Predictions. *arXiv1705.07874 [cs, stat]*

236 Štrumbelj, E. and Kononenko, I. (2014) Explaining prediction models and individual predictions with feature contributions. *Knowl. Inf. Syst.* 41, 647–665

237 Harfouche, A.L. *et al.* (2019) Accelerating Climate Resilient Plant Breeding by Applying Next-Generation Artificial Intelligence. *Trends Biotechnol.* 37, 1217–1235

238 Marko, O. *et al.* (2017) Portfolio optimization for seed selection in diverse weather scenarios. *PLoS One* 12,

239 Khaki, S. and Wang, L. (2019) Crop Yield Prediction Using Deep Neural Networks. 10, 1–10

240 Khaki, S. *et al.* (2020) A CNN-RNN Framework for Crop Yield Prediction. *Front. Plant Sci.* 10, 1–14

241 Mimić, G. *et al.* (2020) Engineering Meteorological Features to Select Stress Tolerant Hybrids in Maize. *Sci. Rep.* 10, 1–10

242 Ansarifar, J. *et al.* (2020) Performance prediction of crosses in plant breeding through genotype by environment interactions. *Sci. Rep.* 10, 11533