Journal of Advanced Research 35 (2022) 215–230

Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Advanced high-throughput plant phenotyping techniques for genome-wide association studies: A review

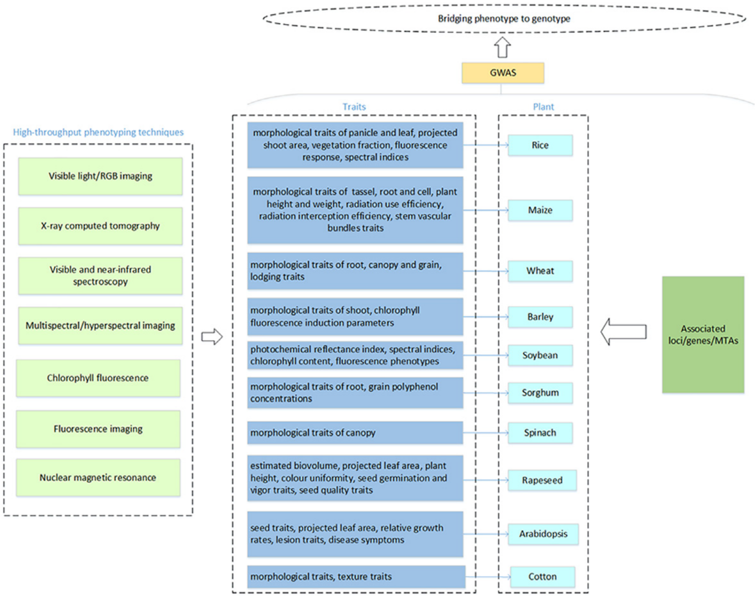
Qinlin Xiao [a,b], Xiulin Bai [a,b], Chu Zhang [c,*], Yong He [a,b,*]

[a] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China
[b] Key Laboratory of Spectroscopy Sensing, Ministry of Agriculture and Rural Affairs, Hangzhou 310058, China
[c] School of Information Engineering, Huzhou University, Huzhou 313000, China

HIGHLIGHTS

- Traits obtained by high-throughput phenotyping perform similarly or even better in GWAS than those obtained by traditional, manual methods.
- Dynamic phenotyping contributing a lot for GWAS to identify time-specific loci.
- High-throughput phenotyping, allowing noncontact and dynamic measurement, possesses great potential to provide high quality trait data for GWAS.
- Future research should focus on the development of low-cost high-throughput phenotyping techniques and efficiency data/image analysis algorithm.
- Research of using various high-throughput phenotyping techniques and GWAS on more diverse plant species and traits is urgently needed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 18 August 2020
Revised 5 May 2021
Accepted 9 May 2021
Available online 12 May 2021

Keywords:
Traits
Gene

ABSTRACT

Background: Linking phenotypes and genotypes to identify genetic architectures that regulate important traits is crucial for plant breeding and the development of plant genomics. In recent years, genome-wide association studies (GWASs) have been applied extensively to interpret relationships between genes and traits. Successful GWAS application requires comprehensive genomic and phenotypic data from large populations. Although multiple high-throughput DNA sequencing approaches are available for the generation of genomics data, the capacity to generate high-quality phenotypic data is lagging far behind. Traditional methods for plant phenotyping mostly rely on manual measurements, which are laborious, inaccurate, and time-consuming, greatly impairing the acquisition of phenotypic data from large populations. In

---

Peer review under responsibility of Cairo University.

* Corresponding authors at: School of Information Engineering, Huzhou University, Huzhou 313000, China (C. Zhang); College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China (Y. He).

E-mail addresses: chuzh@zjhu.edu.cn (C. Zhang), yhe@zju.edu.cn (Y. He).

https://doi.org/10.1016/j.jare.2021.05.002

Phenotype
Imaging
Spectroscopy

contrast, high-throughput phenotyping has unique advantages, facilitating rapid, non-destructive, and high-throughput detection, and, in turn, addressing the shortcomings of traditional methods.

Aim of Review: This review summarizes the current status with regard to the integration of high-throughput phenotyping and GWAS in plants, in addition to discussing the inherent challenges and future prospects.

Key Scientific Concepts of Review: High-throughput phenotyping, which facilitates non-contact and dynamic measurements, has the potential to offer high-quality trait data for GWAS and, in turn, to enhance the unraveling of genetic structures of complex plant traits. In conclusion, high-throughput phenotyping integration with GWAS could facilitate the revealing of coding information in plant genomes.

Introduction

The plant phenotype refers to all morphological, physiological, and biochemical characteristics reflecting the structure, composition, and growth of a plant [1,2]. It encompasses not only agronomic traits such as structure, size, and color, but also the physiological status in the course of development. Conversely, genes are nucleotide sequences that encode polypeptide chains or functional RNA, and therefore, are the basic genetic units regulating trait expressions. Alleles are variant forms of a gene at the same position on a pair of homologous chromosomes, which control different forms of the same trait. Alleles are classified as either dominant or recessive, and they generate functional RNA or proteins that determine whether traits become dominant and recessive. The genotype, which is the sum of all genes obtained from both parents, represents the genetic makeup of a plant. The plant phenotype is influenced by the genotype as well as the environment [3]. According to Mendel's genetic theory, a recessive allele will not be expressed when a dominant allele present. In addition, as the expression of alleles is under the influence of environmental factors (e.g., temperature, light, and soil. [4,5]), dominant traits may emerge only under certain environmental conditions. Therefore, plant phenotypes are the aggregate of three-dimensional (3D) spatiotemporal expression information derived from interactions among genotypes and environmental factors. Diverse phenotypes are formed due to the selective expression of plant genetic information under various environmental conditions.

Over the past few years, tremendous progress has been made in plant genome sequencing [6–8], which has facilitated research on the integration of genotyping and phenotyping for crop improvement. However, traditional phenotyping methods largely rely on manual measurements, which are laborious and time-consuming, and hinder the acquisition of comprehensive phenotypic data from individuals in large populations. Moreover, manual measurements are subjective and error-prone; therefore, data accuracy and reliability cannot be guaranteed [9]. Besides, owing to the workforce, cost, and other contextual limitations, manual measurements can only be exploited for limited features during key stages of plant growth. In addition, phenotypic changes cannot be fully tracked throughout the plant life cycle. High-throughput phenotyping overcomes the above shortcomings of traditional methods, and therefore, has emerged as a powerful tool for evaluating plant phenotypes. High-throughput phenotyping techniques, such as visible light imaging, hyperspectral imaging, and fluorescence imaging, have been successfully applied in evaluating plant growth, biomass, and nutritional status [10–14]. All the techniques above have unique advantages of allowing rapid, non-destructive, and high-throughput detection. Detailed information on the phenotyping techniques can be found in [15].

The robust development of high-throughput phenotyping techniques and gene sequencing technologies has promoted the study of the genetic structures of plant traits. Methods linking phenotypes and genotypes to identify genetic architectures that regulate important traits, such as quantitative trait locus (QTL) mapping, candidate-gene association studies, and genome-wide association studies (GWASs), have been used to study various aspects of plant architecture, development, and responses to environmental factors [16]. GWAS, for example, provides high-resolution genetic data and has a high capacity to link small-effect genes/QTLs on a genome-wide scale. The integration of GWAS and high-throughput phenotyping has enhanced our understanding of plant growth and facilitated crop breeding. Although high-throughput phenotyping has been used to plant phenotyping for a long time, few studies have been conducted on their GWAS applications. Reyazul et al. [16] and Moreira et al. [17] reviewed the progress made in linking significant genes and crop traits by high-throughput phenotyping techniques and various genetic analysis approaches. However, GWAS was rarely mentioned in their reviews. Some reviews have focused on GWAS for specific traits of specific plants [18–21]. Few review papers focusing on progress in high-throughput phenotyping integration with GWAS in plant genetic architecture exploration have been published recently. Therefore, intending to keep abreast of current developments, this review summarizes the status concerning integrating high-throughput phenotyping and GWAS in plants and expounds the loci found through the multi-scale plant traits obtained by various high-throughput phenotyping techniques in GWAS studies. The inherent challenges and future prospects are further discussed to enhance our understanding of the approach of integrating high-throughput phenotyping and GWAS, as well as to provide guidance for future research.

High-throughput phenotyping

High-throughput phenotyping platforms integrate data acquisition equipment, a control terminal, and a data analysis platform. They mainly collect phenotypic data through non-invasive imaging and spectroscopy techniques and adopt high-performance computing equipment to rapidly analyze plant growth activities and physiological status. Compared with traditional phenotyping methods, high-throughput phenotyping facilitates simultaneous data acquisition for multiple traits in large populations and dynamic observation of plants at different growth stages. Second, compared to traditional methods such as visual scoring, which are prone to subjective interpretation, trait characterization based on spectra or images is more objective. Third, it allows non-destructive estimation of biochemical parameters based on modeling, thus avoiding laborious operations.

Over the past years, great efforts have been made to develop high-throughput phenotyping techniques for different targets, such as cells, seeds, shoots, leaves, roots, individual plants, and canopy [22]. In terms of cells and tissues, high-resolution imaging techniques, such as micro-computed tomography (micro-CT) imaging and microscopic imaging, could be used to determine the number of cells [23], changes in cell structure [24], growth rate of cells [25], and tissue morphology [26]. As for seed phenotyping, visible light imaging has been widely utilized for the characterization of morphological traits, such as color [27], the length of coleoptile [28], and germination rate [29]. A smartphone-based portable instrument for seed morphological parameter phenotyping has been developed [30]. Besides, X-ray imaging has been used to evaluate seed morphometric features and tissue integrity [31]. For the assessment of the biochemical components, near-infrared spectroscopy and time-domain pulsed nuclear magnetic resonance (NMR) demonstrated advantages in determining the content of protein, oil content, and fatty acids in seeds [32–34]. As for organs, individual plant, and canopy, the characteristics and potential of various high-throughput phenotyping techniques for obtaining phenotypes at these scales have been documented. Studies of the applications of image-based high-throughput phenotyping techniques related to the acquisition of plant morphological, physiological, and pathological traits in recent years have been summarised by Zhang Y and Zhang N [35]; Yang et al. [36] outlined the commonly used high-throughput phenotyping techniques for plant and its application, especially in the assessment of abiotic stress, pest stress and yield quality in rice and other crops. Shakoor et al. [37] provided a review of high-throughput phenotyping platforms and sensors. The potential applications of high-throughput phenotyping techniques in disease assessment were also detailed. Besides, Liu et al. [38] gave a thorough review of hyperspectral imaging and three-dimensional sensing for plant phenotyping. Jang et al. [39] and Yang et al. [40] focused on reviewing the applications on unmanned aerial vehicles in field crop phenotyping. They summarized the deployed sensors that can be mounted on unmanned aerial vehicles and their characteristics in detail.

In addition to the aforementioned high-throughput phenotyping techniques, some methods have been used for plant phenotyping, although few examples are found. Positron-emission tomography has made a breakthrough in manganese uptake and transport of maize seedlings phenotyping [41] and crop tolerance to western corn rootworm [42]. Besides, microwave resonator is another novel technique that can assess water status [43] and water distribution [44]. High-throughput phenotyping techniques can be used in combination, aiming to provide more comprehensive and accurate results. RGB cameras and multispectral sensors have been integrated for yield estimation [45]. The combination of thermal sensors, RGB cameras and multi-spectral sensors has been applied to yield assessment [46], biochemical parameters determination, and biophysical parameters quantitation [47]. Furthermore, ultrasonic distance sensors with spectral, thermal, and other sensors have been conducted for crop canopy phenotyping [48] and yield prediction [49].

High-throughput phenotyping provides a solid foundation for unlocking the genetic characteristics underlying plant phenotypes [16,22]. In particular, high-throughput phenotyping techniques can be used to achieve short-interval continuous imaging of plants and monitor dynamic changes, which facilitate efficient analysis of all activities in plant growth [50–53]. High-throughput phenotyping tools are generally used to obtain high-resolution images of samples, from which features are extracted by data/image processing algorithms. Sometimes, derived parameters calculated from the acquired values are used, such as the height/width ratio. Therefore, robust data/image processing algorithms are crucial for accurate and efficient phenotypes of plants. Mochida et al. [54] fully discussed machine learning algorithms in the preprocessing, segmentation, feature extraction, and classification of plant images. With advanced computer technology development, deep learning shows outstanding advantages in big data processing [55]. Some studies using deep learning algorithms have gradually emerged in plant phenotyping research. Detailed information on applying deep learning in plant stress phenotyping can be found in the literature [56]. Besides, deep learning has performed well on a board range of plant phenotyping tasks: prediction of relative moisture content [57], flowering time and the rate of flowering detection [58], diagnosis of cold damage [59], panicle and spikes recognization and counting [60–64], seedling development detection [65], leaf to panicle ratio calculation [66], yield estimation and prediction [46,67], estimation of vegetation indices [68], fruit number quantification [69], hypocotyl or coleoptile length determination [70], disease detection and quantitation [71–73], leaf counting [74], temporal phenotype/genotype classification [75], ear density estimation [76], and the segmentation and grading of plant products [77]. It is foreseeable that applying deep learning in plant phenotyping is an inevitable trend, and there will be more researches in this field in the future.

Genome-wide association studies

GWAS uses statistical methods to search for associations between genomic polymorphisms (e.g., SNPs, insertions and deletions, and structural variations) and phenotypic variation. GWAS highly facilitates the analyzing the genetic architectures associated with complex traits [3] and delves into the genetic basis for plant phenotypic diversity [78]. Compared with traditional genetic mapping methods that require substantial effort to construct mapping groups, the following are the key advantages of GWAS that have been highlighted. First, GWAS uses natural populations as research materials, which greatly reduces the research time. Second, GWAS has a high efficiency and high resolution. Third, GWAS is conducted on a genome-wide scale and allows simultaneous detection of multiple alleles at the same locus. Finally, natural populations have high genetic diversity; the germplasm materials and genetic information from a single natural population can be used for the genetic analysis of multiple traits, avoiding the need for repeated population establishment, and greatly reducing the costs of gene sequencing. Important agronomic traits, such as yield, crop quality, and disease resistance, are generally controlled by multiple genes, which are prone to influence by continuously changing environments. Compared with traits controlled by a single gene, the genetic architecture of a complex trait is more intricate. Therefore, the application of GWAS in the exploration of the genetic basis of plant traits has attracted considerable attention. The specific analytical process for high-throughput phenotyping combined with GWAS is illustrated in Fig. 1.

Application of GWAS integrated with high-throughput phenotyping

Rice

Rice (*Oryza sativa* L.) is one of the most important food crops globally. Rice production is closely linked to national food security in some countries. Precise phenotyping of rice traits and the exploration of the underlying genetic structures are essential for improving crop yield and quality. A schematic overview of the application of high throughput phenotyping in GWAS studies for rice is presented in Fig. 2.

Yang et al. [79] established an automated rice phenotyping facility for greenhouse use by integrating X-ray computed tomography (CT) and visible light imaging. Fifteen agronomic traits, including plant height, tiller number, and shoot fresh weight, were measured. Among the 141 loci identified, 25 loci were located close to known genes, such as *SD1*, *Hd1*, and *OsGH3-2*. In addition, Crowell et al. [80] introduced an image skeletonization phenotyping platform into inflorescence phenotyping for field-grown material.

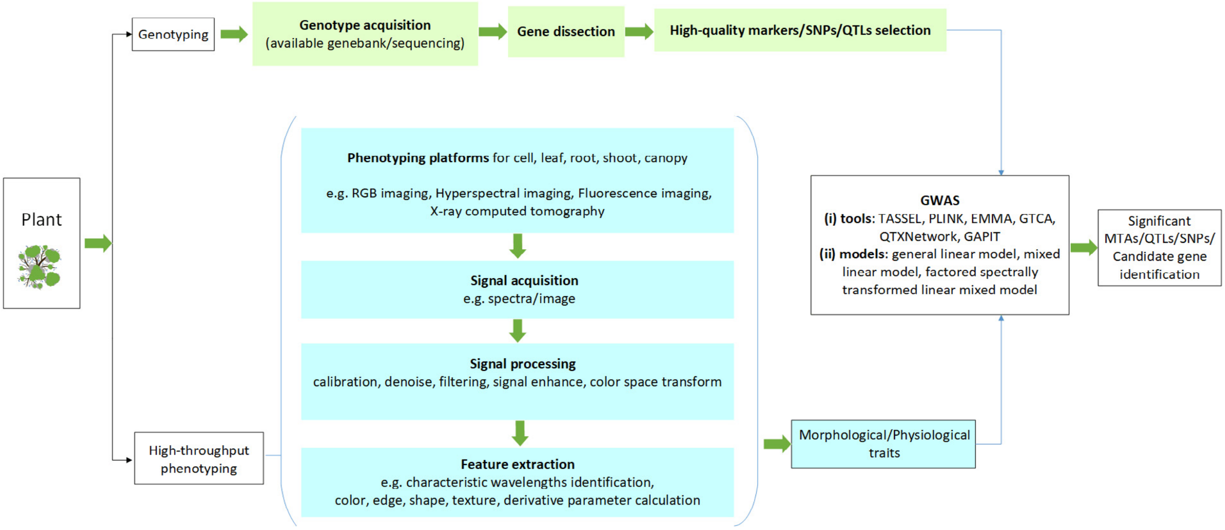


Fig. 1. General analysis process of GWAS equipped with high-throughput phenotyping.

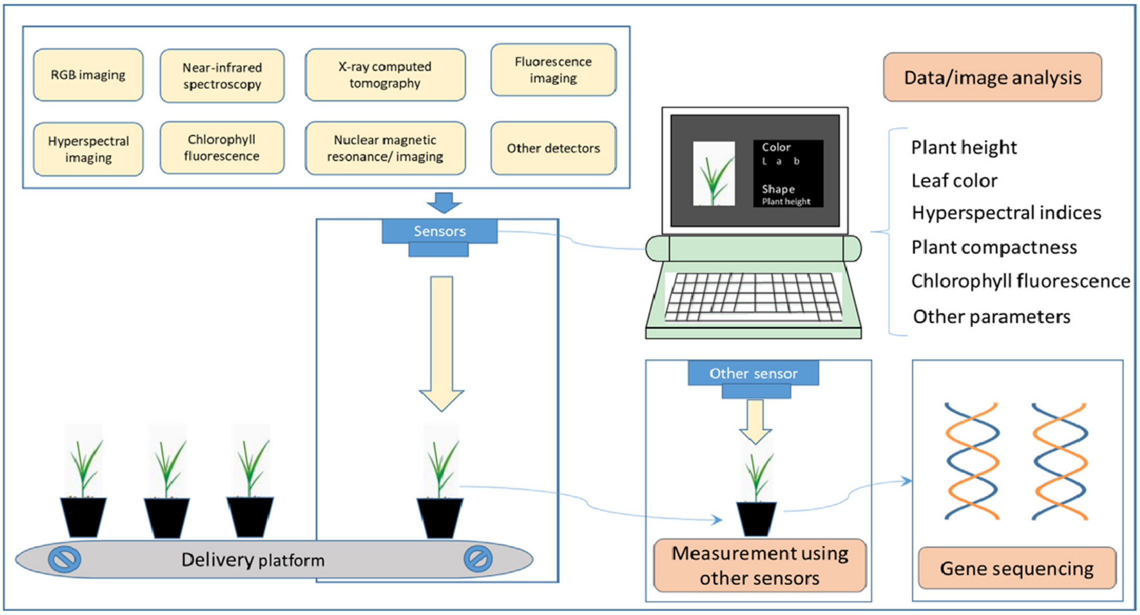


Fig. 2. Schematic overview of high throughput phenotyping used in GWAS studies for rice.

Forty-nine panicle traits were captured. Through GWAS based on 242 accessions, ten candidate genes were identified near significant GWAS peaks. Likewise, Rebolledo et al. [3] conducted comprehensive RGB imaging with GWAS for panicle traits of harvested samples, revealing subtle associations between panicle traits and numerous markers, which had also been reported by Crowell et al [80].

In comparison with cross-sectional phenotyping, dynamic phenotyping greatly improves temporal resolution. The combination of dynamic RGB imaging and GWAS has facilitated the discovery of loci associated with leaf-related traits [81] and shoot growth [51,52,82]. Such approach provides a feasible solution for the detection of QTLs that persist throughout multiple growth stages, as well as time-specific and transient QTLs [52]. By analyzing 739 tiller traits at nine time points using micro-CT-RGB imaging system, GWAS facilitated the identification of 402 loci associated with tiller growth and a gene controlling tiller angle (*TAC1*) [83]. The data revealed that there are three *TAC1* haplotypes and that the tiller angles of rice accessions containing haplotype H3 are substantially smaller. Notably, a dynamic relationship between loci and tiller traits in different developmental stages was revealed [83].

Visible and near-infrared spectroscopy is a powerful tool for obtaining effective spectral parameters. Based on the normalized difference spectral index calculated from spectral data, SNPs similar to those identified on the basis of seed protein content determined by traditional destructive methods as well as new SNPs associated with rice protein content that were not identified through biochemical analysis of protein content were found [84]. In contrast to visible light imaging and spectroscopy, hyperspectral imaging allows trait characterization at the 2D spatial scale as well as the 1D spectral scale. Feng et al. [13] developed a hyperspectral imaging system that enables the acquisition of thousands of hyperspectral indices during multiple growth stages. Based on indices acquired in three growth stages, nine hundred and eighty-nine rice loci associated with growth regulation, many of which could not be identified by traditional agronomic trait analysis, were identified. More importantly, the study revealed that the red edge (680–

760 nm) is valuable for generating phenotypic and genetic information in rice research. In addition, hyperspectral imaging integrated with GWAS also shows the potential in identifying genetic architectures associated with grain quality traits [85]. In this study, the results of GWAS revealed that grain chalk and hyperspectral variants share genomic regions, which contain several plausible candidate genes for grain chalkiness.

Rice plants experience various stresses in the course of growth such as salt and drought stress. Hence, there is a vital requirement to study phenotypes under stress conditions and identify QTLs facilitating stress tolerance. RGB imaging [86,87] and fluorescence imaging [87] have been used in GWAS to explore genetic architecture under salt stress responses. Morphological features, fluorescence responses, and dynamic growth indicators such as relative growth rate, transpiration rate, and transpiration use efficiency have been derived from RGB and fluorescence images [86,87]. The studies revealed that loci on chromosomes 3, 1 [87] and 11 [86] are involved in early responses to salt stress. Dynamic RGB imaging has also been used for phenotyping of individual rice plants under drought stress [88,89]. A study using 507 rice accessions and image-based traits revealed that 443 loci, most of which co-localizing with previously known genes, were identified to be associated with drought stress regulation [88]. Besides, candidate genes of *OSM34*, *COLD1*, *NAL1*, *GF14c*, and *OsMPS* were identified to be associated with shoot growth trajectories under water deficit [89]. Detailed information about studies on high-throughput phenotyping and GWAS in rice is presented in Table 1.

Nearly all studies discussed in this section demonstrated that high-throughput phenotyping integrated with GWAS could accelerate genetic architecture analysis and the dissection of complex traits in rice for the following reasons: (1) Complex traits appear to be regulated by multiple small-effect loci. Dynamic phenotyping provides insights into the genetic architecture that is missed under cross-sectional phenotyping and that is likely to contribute to the detection of small-effect loci. Efficient and non-destructive high-throughput phenotyping meets the requirements for convenient and rapid dynamic phenotyping. (2) Some small-effect loci have a weak correlation with traits and are difficult to detect using traditional methods. By adding novel parameters that encompass different perspectives through spectroscopy, fluorescence imaging, and X-ray CT imaging, high-throughput phenotyping facilitates the localization of novel heritable loci. For example, numerous hyperspectral and fluorescence signals could represent a specific biochemical condition and could reveal phenotypic and genetic relationships with traditional agronomic traits. The acquisition of such features could aid the identification of more small-effect loci that warrant further dissection and analysis at the gene expression level.

Maize

In maize (*Zea mays* L.), to find selection features underlying male inflorescence transformation, Gage et al. [90] conducted a GWAS of 942 inbred accessions and three 200-line biparental populations. Fifteen morphological traits, including branch number and tassel length, were monitored and 242 SNPs were observed to be associated with the traits. In addition, the heritability of the traits measured using manual and image-based methods was investigated. In a follow-up study, the performance of the manual and image-based methods in GWAS were compared based on receiver-operating characteristic (ROC) curves; no significant difference was observed [91]. Similarly, Pace et al. [92] compared the performance of GWAS using a novel root phenotyping platform that they developed, automatic root image analysis (ARIA), with that of GWAS using established software (WinRHIZO). Both software identified significant SNPs associated with total root length in similar genomic regions. A tool for semi-automatic root excavation and cleaning in the field and extracting features based on images (Core Root Excavation using Compressed-air, CREAMD) has been used in GWAS of maize and sorghum [93]. One hundred and thirty-nine root system architecture-associated genes, including *Bige1*, which encodes a MATE transporter, were detected in maize, and seven pairs of syntenic genes were identified in maize and sorghum (*Sorghum bicolor* (L.) Moench) [93].

A study by Wang et al. [94] revealed that plant height-related QTLs vary at different growth stages. In the study, the plant growth rate was recorded by aerial imaging and used in GWAS. Multiple candidate genes involved in plant height regulation, including *SAUR61*, which encodes an auxin response protein, were identified. Similarly, the growth rate was measured in a GWAS of biomass, which accumulates gradually during plant growth [10]. The identification of genetic loci regulating biomass through specific-stage phenotyping is substantially less comprehensive than dynamic phenotyping, and non-destructive imaging makes spatio-temporal dynamic monitoring feasible. The most significant marker-trait associations explained more than 12% of the natural phenotypic variation in biomass accumulation. The study also showed that plant biomass is controlled by numerous small-effect loci, some of which act at specific growth stages. Chen et al. [11] analyzed the genetic and environmental components affecting biomass accumulation in the canopy. Radiation interception efficiency, radiation use efficiency, and biomass were evaluated, and the Monteith equation was used to identify environmental influences on interception efficiency and radiation use efficiency. Simulated multi-genotype canopies were used to determine the extent to which canopy heterogeneity influences biomass accumulation of individual genotypes [11]. Microscope-RGB imaging-assisted GWAS has also been used to identify candidate genes associated with bulliform cell number and width [95]. In the study, bulliform cell column number and width were quantified from tens of thousands of leaf epidermal glue-impression images using convolutional neural networks. Likewise, micro-CT imaging was also used to explore the genetic architecture associated with maize stem vascular bundles [96]. Detailed information on the above studies is presented in Table 1.

In maize, high-throughput phenotyping integrated with GWAS has been successfully applied to investigate cell and root traits, as well as to identify selection signatures for male inflorescence transformation, which opens avenues for future research on changes in genetic information in the course of evolution [90]. In addition, studies have shown that for complex traits, some genes act in certain developmental phases, whereas others play a more general role and function in multiple stages [10,94]. Dynamic phenotyping is especially advantageous for traits that develop gradually, such as biomass accumulation and plant height.

Wheat

In a GWAS of 231 synthetic hexaploid wheat (*Triticum aestivum* L.) accessions to identify QTLs controlling grain morphology, visible light/RGB imaging was used to measure 29 traits related to grain morphology [97]. Candidate genes, including *TaCwi-2A*, *TaSus-6B*, *TaCKX-6D*, and *TaGW2-2B*, which are known to regulate grain size and weight, were identified. Furthermore, relationships between some favorable alleles and grain phenotypes were revealed [97]. A semiautomatic system including spectrometers was used to capture the canopy reflectance of wheat under normal nitrogen supply and nitrogen deficiency [12]. Three vegetation indices derived from canopy reflectance were used in a GWAS. The study identified a set of loci associated with canopy traits as well as *PPD-D1*, which is involved in photoperiod response regulation. Another study using unmanned aerial systems-based imaging

Table 1
Applications of high-throughput phenotyping integrated with GWAS in plant.

Crop	Population size	Technique	Environment	Traits	Number of associated loci/SNP/QTL/candidate genes	Reference
rice	533 accessions	x-ray CT, color-imaging	Greenhouse	15 morphological traits (e.g. plant height, tiller number, green leaf area)	141 loci	[79]
rice	242 accessions	visible light imaging	Field	49 panicle traits (e.g. panicle length, rachis length, primary branch number)	10 candidate genes	[80]
rice	225 accessions	RGB imaging	Laboratory	number of spikelets per panicle (NSP), the number of primary (PBN) and secondary branches (SBN), the length of primary (PBL) and secondary (SBL) branches	17 GWAS sites for NSP, 10 for PBN, 11 for SBN, 7 for PBL, 11 for SBL	[3]
rice	533 accessions	RGB imaging	Laboratory	29 leaf traits (6 size–related traits, 7 color–related traits, 16 shape–related traits)	73 loci for size-related traits, 123 for color-related traits, 177 for shape related traits	[81]
rice	360 accessions	visible light/RGB imaging	Greenhouse	projected shoot area	7 QTLs	[51]
rice	357 accessions	RGB imaging	Greenhouse	projected shoot area	442 SNPs	[52]
rice	165 lines	unmanned aerial vehicle, RGB imaging	Field	vegetation fraction	4 QTLs	[82]
rice	234 accessions	micro-CT, RGB imaging	Laboratory	739 traits during tillering process (e.g. tiller number, convex hull area, total tiller area, height/width ratio)	402 loci	[83]
rice	80 accessions	visible/near-infrared spectroscopy	Laboratory	5 hyperspectral traits (reflectance at wavelength 1177 nm and 1227 nm, normalized difference spectral index (NDSI), differential spectral index, simple ratio index)	NDSI: 65 genes	[84]
rice	529 accessions	hyperspectral imaging	Laboratory	1540 hyperspectral indices	989 loci	[13]
rice	221 accessions	visible and near-infrared hyperspectral imaging	Laboratory	the first principal component for spectral values in the range 702–922 nm	44 chromosomal candidate regions	[85]
rice	553 genotypes	visible light/RGB imaging	Greenhouse	relative growth rate, transpiration rate, transpiration use efficiency (TUE)	previously undetected loci affecting TUE on chromosome 11	[86]
rice	378 genotypes	visible light/RGB imaging, fluorescence imaging	Greenhouse	97 digital traits (7 morphology and growth traits, 90 fluorescence responses traits)	visible light/RGB imaging: a genomic region on chromosome 3; fluorescence imaging: 4 genomic regions	[87]
rice	507 accessions	RGB imaging	Greenhouse	51 image-based traits (e.g. total projected area, plant compactness, height/width ratio)	443 loci	[88]
rice	378 accessions	visible light/RGB imaging,	Greenhouse	projected shoot area	\	[89]
maize	942 inbred accessions, three 200 biparental lines	visible light /RGB imaging	Field	15 tassel morphological traits (e.g. branch number, tassel length, tortuosity)	242 SNPs	[90]
maize	942 inbred lines	visible light/RGB imaging	Field	tassel weight, tassel length, spike length, branch number	\	[91]
maize	384 inbred lines	visible light /RGB imaging	Laboratory	total root length, total surface area	\	[92]
maize, sorghum	maize: 369 inbred lines; sorghum: 294 accessions	visible light/RGB imaging	Field	root area, convex hull area, median width, maximum width, width-profile angle, adjusted depth	maize: 139 genes, sorghum: 115 SNPs	[93]
maize	252 inbred lines	aerial visible light/RGB imaging	Field	plant height at 4 different growth stages, the growth rate of plant height	68 QTLs	[94]
maize	252 inbred lines	near-infrared, visible light/ RGB and fluorescence imaging	Laboratory	plant fresh weight, plant dry weight, biovolume estimation at 11 different developmental time points	12 MTAs, 6 pairs of epistatic interactions	[10]
maize	255 hybrids lines	RGB imaging	Laboratory	biomass, radiation interception efficiency, radiation use efficiency	\	[11]
maize	468 inbred lines	Microscope, RGB imaging	Laboratory	bulliform cell column number and width	5 candidate genes	[95]
maize	480 inbred lines	micro-CT imaging	Laboratory	48 stem vascular bundles traits	1562 SNPs	[96]
bread wheat	231 synthetic hexaploids	visible light /RGB imaging	Field	grain size and shape traits (e.g. length, width, volume of seed)	197 loci using general linear model, 79 loci using mixed linear model	[97]
bread wheat	211 bread wheat cultivars	visible/near-infrared reflectance spectroscopy	Field	the modified canopy adjusted ratio index 2 (MCARI2), the MERIS terrestrial chlorophyll index (MTCI), the normalized difference vegetation index (NDVI)	105 QTLs for MTCI, 97 QTLs for NDVI, 159 QTLs for MCARI2	[12]

Table 1 (*continued*)

Crop	Population size	Technique	Environment	Traits	Number of associated loci/SNP/QTL/candidate genes	Reference
spring wheat	1185 lines	unmanned aerial system integrate RGB and RedEdge multispectral imaging	Field	lodging traits: 3 visual scores of lodging namely intensity, severity and lodging index per plot and additional supporting agronomic measurements per plot, 2 digital lodging scores obtained by taking overall summary mean per plot or combined lodging index of normal mixture parameters	a key genomic region on chromosome 2A	[98]
Durum wheat	248 accessions	unmanned aerial vehicles integrated with multi-spectral imaging, tractor-based system integrated with GreenSeeker spectral sensors	Field	normalized difference vegetation index (NDVI)	46 QTLs	[99]
winter wheat	335–352 genotypes	light detection and ranging (LIDAR)	Field	canopy height, average daily stem elongation rates	10 MTAs for final height, 3 MTAs for temperature response, 4 MTAs for vigour	[100]
winter wheat	335 cultivars	RGB imaging	Field	septoria tritici blotch infected traits (e.g. the percentage of leaf area covered by lesions, average pycnidia density within lesions, pycnidia size, pycnidia gray value)	26 chromosome segments	[101]
winter wheat	215 lines	visible light/RGB scan	Laboratory	total seminal root (TSR) length, root diameter (RD), the length of seminal axis roots (SAR), branched root (BR) length; root dry matter (RDM)	63 MTAs with 7 for SAR, 24 for BR, 4 for TSR, 8 for RDM, and 20 for RD	[102]
spring barley	100 genotypes	visible light imaging	Greenhouse	biomass and related traits (tiller number, tipping time, the calculated inflection point, fresh weight)	21 loci	[104]
barley	1420 lines	RGB imaging	Greenhouse	14 growth traits (e.g. absolute growth rate, relative growth rate, shoot area smoothed, convex hull area integral, caliper length integral)	\	[105]
barley	109 accessions	chlorophyll fluorescence	Laboratory	23 traits, including 19 chlorophyll fluorescence induction parameters	205 markers	[106]
soybean	373 genotypes	visible/near-infrared spectroscopy	Field	photochemical reflectance index (PRI), canopy spectral reflectance	31 SNPs	[110]
soybean	332 genotypes	visible/near-infrared spectroscopy	Field	chlorophyll *a* (eChl_A), chlorophyll *b* (eChl_B), total chlorophyll (eChl_T) content, chlorophyll *a*/*b* ratio (eChl_R)	14 loci for eChl_A, 7 loci for eChl_B, 10 loci for eChl_R, 27 putative loci for total chlorophyll content	[111]
soybean	189 genotypes	chlorophyll fluorescence	Field	21 fluorescence phenotypes	288 SNPs	[112]
soybean	5555 lines	ground-based and unmanned aerial system-based RGB imaging	Field	canopy coverage	a large QTL on chromosome 19	[113]
soybean	200 accessions	unmanned aircraft system, digital imaging	Field	dark green color index (DGCI)	45 SNPs	[114]
soybean	341 accessions	Photosynthetic System II imaging, visible and near-infrared hyperspectral imaging	Greenhouse	NDVI, chlorophyll index (CHL)	38 QTLs for NDVI, 32 QTLs for CHL	[115]
spinach	284 accessions	unmanned aircraft system, RGB imaging	Field	9 plant growth traits: canopy cover, canopy volume, excess greenness index; days after sowing to maximum seasonal values of canopy cover, canopy volume and excess greenness index; days after sowing for manually collected plant bolting stages: early bolting. pollination, and kernel filling	99 SNPs	[9]
cotton	200 accessions	RGB imaging	Greenhouse	119 image-based digital traits (56 morphological traits and 63 texture traits)	390 loci	[116]
sorghum	381 accessions	near-infrared spectroscopy	Laboratory	total phenol, proanthocyanidin, and 3-deoxyanthocyanidin concentrations	\	[14]
arabidopsis	382 accessions	visible light imaging, fluorescent imaging	Laboratory	seed width, seed length, seed area, projected leaf area, relative growth rates	238 MTAs	[118]
arabidopsis	96 accessions	visible light/RGB imaging	Laboratory	75 lesion traits for each infected leaf (e.g. the number of pixels for specific hues within the lesion, lesion perimeter, proportion of the leaf the lesion occupied)	7940 genes	[119]

(*continued on next page*)

Table 1 (*continued*)

Crop	Population size	Technique	Environment	Traits	Number of associated loci/SNP/QTL/candidate genes	Reference
arabidopsis	110 and 70 ecotypes	visible light/RGB imaging	Laboratory	disease symptoms (e.g. the percent of tissue that is chlorotic yellow)	\	[120]
rapeseed	477 genotypes	visible light imaging, fluorescence imaging	Greenhouse	estimated biovolume, projected leaf area, early plant height, colour uniformity	787 MTAs	[117]
rapeseed	248 accessions	visible light/RGB imaging	Laboratory	9 seed germination and vigor traits	18 candidate genes	[121]
rapeseed	238 inbred lines	near-infrared reflectance spectroscopy, nuclear magnetic resonance	Laboratory	seed traits: erucic acid content (EAC), glucosinolate content (GSC), oil content (SOC)	6 loci for EAC, 49 loci for GSC, 17 loci for SOC	[122]

MTA: marker-trait association

and GWAS for the detection of genetic regions associated with lodging identified a significant marker on chromosome 2A [98]. Likewise, unmanned aerial vehicles coupled with multi-spectral imaging was also used in normalized difference vegetation index (NDVI) measurement. Forty-six QTLs associated with NDVI were detected in GWAS [99]. Furthermore, light detection and ranging (LIDAR) contributed in wheat GWAS, which was used to investigate the genetic response to temperature fluctuations during stem elongation [100]. Aerial systems are expected to substantially facilitate canopy trait estimation activities, such as lodging and canopy coverage, which will facilitate the discovery of novel genetic loci related to such traits. RGB imaging combined with GWAS has also been used in detecting genetic architectures related to disease resistance. Yates et al [101] used flatbed scanners to acquire the leaf traits associated with septoria tritici blotch infection. Twenty-six chromosome intervals were idenfied as affecting four independent resistance traits in GWAS analysis. For genetic architecture exploration of root traits, Beyer et al. [102] used a scanner and the traditional root analysis platform, WinRHIZO, to evaluate five root traits; GWAS of 20,881 polymorphic sites revealed 63 marker–trait associations for root morphology.

The application of high-throughput phenotyping integrated with GWAS in the study of wheat is rather rare and the field is still in the early stage of development. At present, few traits have been explored, and only a limited number of techniques have been applied. This is largely because wheat is an allohexaploid, so that sequencing costs are relatively high. With advances in wheat gene resequencing technologies, more wheat GWAS will be undertaken. In fact, numerous high-throughput phenotyping techniques are available and have been applied in wheat. For instance, terrestrial 3D laser scanning has been used to estimate increase in canopy height under field conditions [103]. In addition, a dual-mode microwave resonator has been used to evaluate leaf water content and ionic conductivity [43]. Such techniques are expected to be applied in further genetic studies in wheat. All the above-mentioned studies on root traits applied ectopic investigations, which means that the root system was evaluated outside of the natural growth environment. However, such an approach has some disadvantages, including loss of the 3D root structure and underestimation of the contribution of fine roots, as these are prone to fracture during sampling. At present, studying root phenotypes *in situ* is challenging. The application of techniques suitable for *in situ* root phenotyping (such as X-ray CT and magnetic resonance imaging) may allow more accurate visualization and quantification of root growth, in turn, improving the analysis of root morphology.

Barley

To reveal associations between genetic architecture and barley (*Hordeum vulgare* L.) biomass, biomass and related traits were measured for individual plants at multiple time points using visible light imaging [104]. Twenty-one loci were identified, seventeen of which had also been identified in previous studies. A locus on chromosome 7HL that exhibited a significant effect for a long period and that co-located with *HvDIM* was found to be associated with biomass. Dynamic phenotyping has been conducted to evaluate drought stress responses of barley [105]. In the study, fourteen traits, including absolute and relative growth rates, were determined and numerous significant QTLs that co-localized with known genes (*Ppd-H1*, *HvCEN*, *VRN-H1*, *VRN-H2*, and *sdw1/denso*) were identified, illustrating the validity of traits acquired by visible light imaging for GWAS [105]. A QTL on chromosome 4H was possibly involved in biomass increase under both control and drought conditions [105]. Chlorophyll fluorescence response of leaves under drought stress has also been studied [106]. One hundred and sixty-two associations with physiological parameters (gas exchange, leaf water status, and chlorophyll fluorescence induction) were found, sixty-seven of which were annotated to known sequences [106]. Chlorophyll fluorescence reflects the photosynthetic status of plants and, therefore, is useful in the genetic study of plant stress phenotypes.

As mentioned above, high-throughput phenotyping-integrated GWAS of plants is still in its infancy. Therefore, it is not surprising that its application in barley remains sporadic. Although diverse high-throughput phenotyping techniques are available for trait evaluation, they have been applied in only few barley GWAS, and various techniques have not been introduced in barley GWAS to date. For example, light curtain arrays have been utilized for the rapid determination of plant height and leaf area [107]. Hyperspectral imaging has been used to evaluate changes in barley leaves at the cellular level during resistance reactions against powdery mildew [108]. Hyperspectral absorption-reflectance-transmittance has also been used to assess disease severity on leaves [109]. The examples above suggest that the techniques hold promise in uncovering the genetic architectures of barley traits. Successful phenotyping requires not only accurate measurement of plant traits but also high-quality trait data and abundant genetic information. Most studies used a single technique, such as RGB imaging, to measure morphological traits, or hyperspectral imaging, to acquire spectral parameters. Future research should aim at combining various techniques to measure numerous plant traits from which high-quality traits can be selected to conduct in-depth analysis of plant growth from multiple perspectives. Such an approach would certainly facilitate the identification of more loci involved in the regulation of plant traits.

Soybean

Herritt et al. [110] identified genetic loci for a photosynthesis trait in soybean (*Glycine* max (L) Merr.), i.e., the photochemical reflectance index calculated from canopy spectral reflectance, which was measured in the field by visible/near-infrared spectroscopy. Fifteen putative loci exhibited a significant association with the trait and several of them were located close to genes known to be associated with photosynthesis, nonphotochemical quenching and sugar transport processes. Visible/near-infrared spectroscopy and GWAS analysis was also utilized for exploring the genetic basis of chlorophyll traits. Twenty-seven putative loci were designated as associated with total chlorophyll content, four of which were indicated by all extract-based and canopy spectral reflectance-based approaches [111]. In addition, Herritt et al. [112] carried out an impressive GWAS designed to study genetic dissection associated with chlorophyll fluorescence phenotypes. Twenty-one chlorophyll fluorescence phenotypes were captured by a fluorometers. A total of 288 SNPs were detected as significantly related with one or more of the measured chlorophyll fluorescence phenotypes.

Canopy coverage of soybean measured using ground-based and aerial visible light/RGB imaging has been observed to potentially have a genetic correlation with yield, and a QTL on chromosome 19 with a strong positive effect on production was identified by GWAS [113]. Dark green color index (DGCI) captured by aerial images was used in GWAS to explore the genetic architecture of the intensity of greenness. Among the 43 putative loci identified through GWAS, twenty-one loci were coincident with previously reported genetic regions, which were related with nitrogen traits and ureide concentration [114]. Similarly, NDVI and chlorophyll index (CHL) obtained from hyperspectral images were used to study the population genetics underlying the growth and yield of Chinese soybean germplasm population [115].

In soybean, high-throughput phenotyping combined with GWAS has contributed for the discovery of genetic regions for several spectral traits, such as NVDI, CHL, DCCI, which tend to be closely linked with yield and plant growth [114–115]. It is worth mentioning that the study conducted by Wang [115] elucidated that the exploration of genetic architecture of upstream traits, such as NVDI and CHL, may provide additional understanding for controlling target traits. Therefore, high-throughput phenotyping techniques, allowing to acquire and select upstream traits, will facilitate crop functional genomics and improve the potential for crop breeding.

Other species

In spinach (*Spinacia oleracea* L.), aerial RGB imaging has been used for time-course measurements of growth traits throughout the crop cycle. GWAS identified 99 SNPs, several of which were located in transcription factor and stress-response genes that had putative roles in plant development [9]. In cotton (*Gossypium hirsutum* Linn.), RGB imaging combined with GWAS was used to characterize genetic loci for drought resistance. Using 119 image-based digital traits, GWAS allowed the identification of 390 loci, some of which has been reported previously. Remarkably, some promising genes, which may have a negative affect for drought response, were also detected [116]. In sorghum (*Sorghum bicolor* (L.) Moench), near-infrared spectroscopy has been used to estimate the contents of total phenolic, procyanidins, and 3-deoxyanthocyanins in grain of 381 accessions, and novel QTLs involved in polyphenol synthesis were identified, some of which were homologous to flavonoid genes in maize (*Zea mays* L.) *Pr1* and Arabidopsis (*Arabidopsis thaliana*) *TT16* [14]. Visible light imaging and fluorescence imaging have also been used to screen the genetic variation underlying growth dynamics in canola (*Brassica napus* L.) [117] and *Arabidopsis thaliana* [118]. These studies [117,118] both highlighted the importance of conduct time-course analyses of plant traits since they are regulated by specific QTLs at different growth stages. Notably, attempts have been made to apply high-throughput phenotyping combined with GWAS to the mapping of defense response genes in plants [119,120]. Symptoms of *Botrytis* infection [119] and *Pseudomonas syringae* type-3 secreted effectors for effector-triggered immunity elicitation [120] in Arabidopsis (*Arabidopsis thaliana*) have been quantified using visible light and RGB imaging, respectively. GWAS identified 23 candidate genes significantly related to responses to *Botrytis* infection [119] and host genes for *P. syringae* type-3 secreted effector-triggered immunity, demonstrating the suitability of image-based phenotyping coupled with GWAS [120]. In addition, visible light/RGB imaging, near-infrared reflectance spectroscopy, and NMR have been used in GWAS for seed traits such as germination and vigor [121], and glucosinolate and oil contents [122]. Strong correlations between the identified candidate genes and seed quality variation were observed. Detailed information about studies published over the last five years that we found via a literature search is presented in Table 1. These studies confirmed that GWAS integrated with high-throughput phenotyping is a practical and effective strategy for linking complex traits and genes in plants.

As shown in Table 1, the traits obtained by high-throughput phenotyping techniques under laboratory, greenhouse, and field conditions have been applied in GWAS. It is easy to control environmental factors such as temperature, humidity, and light in laboratory and greenhouse conditions. On the contrary, the phenotypic data obtained in the field are susceptible to interference from environmental factors, which may cause large uncertainties in the results. Therefore, to ensure the stability of the results, there are certain requirements of the timing and environment in the field measurement. Besides, laboratory measurements are more suitable for small-scale samples, such as cells, seeds, and individual plants. Greenhouse and field-based techniques are applicable to obtain the traits of individual plants and canopy. UAV-based high-throughput phenotyping devices can be used in the field to capture traits of large populations, such as plant height, lodging, and yield. Whether in the laboratory, greenhouse, or field conditions, high-throughput phenotyping platforms can be fixed or mobile, although a flexible hardware and software phenotyping platform may facilitate the acquisition of plant traits faster and more efficiently.

The merits and shortcomings of the most extensively used high-throughput phenotyping techniques and the traits evaluated in GWAS are shown in Table 2. Visible light/RGB imaging is the most popular technique for the measurement of various traits, from the cell level to the canopy level. Visible light/RGB imaging applies a series of image processing algorithms to identify features in images acquired with a digital camera. A major advantage of the approach is the low equipment requirements. Although visible light/RGB imaging allows the assessment of diverse morphological traits, it is highly dependent on image processing algorithms; proper image processing algorithms are key for accurate and rapid trait recognition. X-ray CT is a form of imaging that allows structural analysis of samples through X-ray irradiation with high penetrability. Besides rice tiller traits, X-ray CT has been applied in the measurement of other plant traits, such as pollen grain traits [123] and wheat grain traits [124]. Although the application of X-ray CT in GWAS is still in its infancy, the study on the genetic structure of tiller traits in rice demonstrated the feasibility of applying X-ray CT to GWAS and the unique advantages of the technique in detecting rice tiller traits. X-ray CT and micro-CT provide high-resolution images of plants, and reconstruction of the 3D structure of a sample facilitates accurate trait measurements. However, their high costs limit their application. Visible near-infrared spectroscopy is convenient for acquiring spectral parameters, and spectral reflectance values can be used to investigate the contents of certain biochemical components through modeling, in a non-invasive manner and on a large scale. Multispectral/hyperspectral imaging combines spectroscopy with

Table 2
Advantages and disadvantages of common used high-throughput phenotyping techniques and their evaluated traits used in GWAS.

Techniques	Traits used for GWAS	Advantages	Disadvantages
Visible light/RGB imaging	morphological traits (shape, color, size-related traits): (1) panicle traits: e.g. panicle length, rachis length, primary branch number; (2) leaf traits: e.g. green leaf area; (3) tassel traits: e.g. tassel weight, tassel length, spike length, branch number; (4) root traits: e.g. total root length, total surface area, convex hull area, adjust depth; (5) canopy traits: canopy coverage, biomass, radiation interception efficiency, radiation use efficiency; (6) cell traits: cell column number and width; (7) seed traits: e.g. germination rate at certain time, volume increase, mean germination time. (8) others: tiller number, projected shoot area, relative growth rate, transpiration rate, transpiration use efficiency, plant compactness, digital biomass, plant height, etc.	low equipment expense, suitable for wide applications	only allow appearance information acquisition; highly depend on image processing algorithms,
X-ray computed tomography	(1) tiller traits: tiller number, size, and shape related parameters, tiller angle; (2) tiller growth traits: absolute growth rate of total tiller area, relative growth rate of total tiller area, absolute growth rate of tiller number, relative growth rate of tiller number. (3) stem vascular bundles traits (micro-CT)	sensitive	high expense
Visible and near-infrared spectroscopy	(1) spectral indices for crop and canopy: e.g. reflectance at specific wavelengths, normalized differential spectral index, differential spectral index, simple ratio index, the modified canopy adjusted ratio index 2, the MERIS terrestrial chlorophyll index, the normalized difference vegetation index, photochemical reflectance index, canopy spectral reflectance, the derived hyperspectral indices; (2) biochemical parameters: e.g. total phenol, proanthocyanidin, and 3-deoxyanthocyanidin concentrations in grain, glucosinolate content in seed.	biochemical component content can be estimated by modeling	point measurement (cannot be represent spatial information), background interference
Multispectral/ hyperspectral imaging	(1) hyperspectral indices: e.g. total reflectance, average reflectance, the derived hyperspectral indices; (2) lodging traits: e.g. visual scores of lodging intensity, severity and lodging index.	rich spatial and spectral information acquisition	complex data/image processing
Chlorophyll fluorescence	chlorophyll fluorescence induction parameters (e.g. Fv/Fm)	high sensitive to plant physiological changes	point measurement
Fluorescence imaging	biovolume estimations	high sensitivity	sensitive to interference, small field of view
Nuclear magnetic resonance	seed oil content	high resolution	high cost

conventional imaging and not only provides 1D spectral information but also 2D spatial information, enabling rapid acquisition of massive data and allowing comprehensive sample analyses. However, the technique also has some limitations. Effective processing of the massive data and redundancy elimination are key to generating accurate representations of traits, and high-performance data-processing methods are crucial. Both chlorophyll fluorescence and fluorescence imaging are based on organic materials emitting unique fluorescence patterns following excitation. Chlorophyll fluorescence is considered a probe for photosynthesis and chlorophyll fluorescence parameters are appropriate tools for evaluating the physiological status associated with photosynthesis. Fluorescence imaging integrates fluorescence and imaging techniques to generate fluorescence images. While such tools can sensitively detect changes in fluorescence, the results are prone to be the influence of external factors such as an exterior light source. NMR uses radio frequency pulses to capture signals reflecting the proton content to determine the characteristics of specific components in samples. NMR has demonstrates particularly high resolution in the detection of certain key metabolites in seeds [125]; however, its application is also limited by high equipment costs. Various other high-throughput phenotyping techniques that are not discussed in this review, such as infrared thermography and optical coherence tomography, require further investigations, in combination with GWAS.

Owing to the advantages of automated data acquisition, minimal sample preparation, and non-invasiveness, the above high-throughput phenotyping techniques have been exploited to assess numerous plant traits in GWAS. For instance, traits related to morphology and texture can be acquired by imaging techniques, and biochemical components can be estimated by building models based on near-infrared spectral reflectance or other types of signals. However, it is worth noting that high-throughput phenotyping techniques still cannot completely replace manual measurements. Modern analytical tools make it feasible to automatically or semi-automatically measure plant phenotypes from the microscale to the macroscale. However, at present, some of the tools require complex sample preparation, which hamper high-throughput phenotyping. For example, high-throughput phenotyping at the cellular or subcellular scale for certain metabolites and indicators that are difficult to measure using currently available instruments has been rarely reported since complex sample preparation procedures and high-performance instruments are required. With further analytical and instrumental developments, high-throughput phenotyping of such attributes could be realized.

Challenges and future prospects

Although numerous studies have demonstrated the potential applications and roles of high-throughput phenotyping in plant research, relatively few studies have integrated high-throughput phenotyping and GWAS. The major factors limiting the application of high-throughput phenotyping and GWAS are challenges in acquisition of genetic data, accurate characterization of plant traits, and lack of adequate individuals with the required skills. Genetic data are mainly acquired from gene banks or by resequencing. Whole-genome resequencing is laborious, time-consuming, and costly. In addition, current gene banks are few and the available data may not match the actual samples. Accurate trait acquisition is essential for GWAS; however, current high-throughput phenotyping techniques applied in GWAS are still generally flawed. Various high-throughput phenotyping techniques, including the most commonly used visible light/RGB imaging, hyperspectral imaging, and X-ray CT, strongly depend on data/image processing algorithms. Currently, available signal analysis algorithms have various shortcomings, such as low efficiency and imperfect and inaccurate feature extraction; therefore, they require improvements. High-resolution and high-sensitivity equipment used for hyperspectral imaging, X-ray CT, and fluorescence imaging are costly and are therefore not used extensively. Unmanned aerial vehicles for near-surface high-throughput phenotyping are suitable for obtaining canopy phenotypic data due to flexibility and extensive spatial coverage. However, high cost, insufficient payload, and the complex techniques required to process massive remote sensing data limit their adoption. More promising approaches, such as infrared thermal imaging and optical coherence tomography, should be introduced into GWAS.

At present, some studies focused more on how to acquire phenotypic data of plants using high-throughput phenotyping techniques and how to deal with the acquired data for phenotype analysis [22,35,37,38,54,56,126]. Some other studies presented in this review used high-throughput phenotyping techniques to acquire phenotypic data for genomic analysis [16–17]. For both kinds of researches, a common phenomenon could be observed that the acquired phenotypic data were used for specific objectives, and may not be accessible over generations. Phenotypic data are significantly affected by genotype and environment. Changes in the environment can cause changes in phenotypic data, the unknown environmental factors will cuase uncertainty of phenotypic data. Indeed, it is possible to generate phenotypic data databases by high-throughput phenotyping techniques, only if enough environmental situations are considered, and it is quite difficult to achieve this goal. Recent studies focused on linking the phenotypic data with certain phenotypes and with certain genotyping data that may have relationships with specific phenotypes. The expectation that phenotypic data should be accessible over generations to preserve non-repeatable experiments in the context of constantly changing environments has not been fully considered. However, although the acquired phenotypic data may not be re-used directly due to the changing environmental factors, specific traits associated with certain genes reflected in the current phenotypic data could still be used over generations for the same genotypes. These traits can be acquired and analyzed for the next generations.

Various high-throughput phenotyping techniques can obtain multiple types of plant phenotypic data from different aspects, which significantly extend the content of plant phenotypic data. These data are presented in heterogeneous data formats, causing difficulties in data analysis and management. Data analysis approaches are important to deal with these heterogeneous data formats and transform these data formats into the same format for data representation. Different data sources representing various types of phenotypic data can be further fused to comprehensively understand the relationship between phenotypic data, phenotypes, and genotypes. The high data volume is also a problem for analyzing phenotypic data obtained by these techniques. Multiple data analysis approaches (for example, deep learning [57–77]) have been developed to deal with phenotypic data obtained from various techniques. However, how to make full use of massive data is challenging for phenotypic data analysis. Although high-throughput plant phenotyping has been widely studied in recent studies, very few phenotypic datasets have been provided for free access. Indeed, it is quite hard to acquire phenotypic data, they are very valuable, and the researchers have paid a great deal of effort. However, with more public available phenotypic datasets, non-repeatable experiments in the context of constantly changing environments can be preserved. Many more efforts are needed to enrich the sources of public available phenotypic datasets. There are still shortcomings in the analysis, management, and effective preservation of phenotypic data, limiting the use of phenotypic data.

The combination of high-throughput phenotyping and GWAS could reveal the genetic variation underlying complex traits in various plant species. The current status of application of GWAS integration with high-throughput phenotyping in plants is illustrated in Fig. 3. It is obvious that current research is mainly focused on model species such as rice and maize, whereas studies on other plants are rare. Other important plant traits, such as heading date, and other species are expected to be investigated in future studies.

This review focused on the application of high-throughput phenotyping techniques to explore genetic architectures of plant traits. Such techniques can also be applied in GWAS in humans, animals, bacteria, and in other fields. For example, magnetic resonance neuroimaging data have been used to analyze genetic information related to brain sulcus widening [127]. Microscope-RGB imaging has potential applications in the identification of candidate genes associated with morphological traits of cells [95].

To enhance the use of high-throughput phenotyping and GWAS for the exploration of the genetic structures underlying complex traits and to carry out more relevant studies, the following aspects can be considered:

(1) Investment in research and development in high-efficiency and accurate population genotypic data acquisition technologies should be increased.
(2) Ample genotypic data on several important crops are available in public databases (https://bigd.big.ac.cn/gvm/home). Collecting plant materials with known genotypes for high-throughput phenotyping is a potential strategy of minimizing associated costs.
(3) Low-cost high-throughput phenotyping techniques should be developed to encourage other applications and adoption.
(4) The establishment and improvement of the public phenotypic database are of great significance to effectively solve insufficient metadata, heterogeneous data formats, or resource problems in data provision. We encourage and recommend the publication of metadata, which should be structured according to the FAIR principles [128] and be clearly provided with all the detailed information, such as data format and the environmental conditions. The guidelines for governing the description of phenotypic data has been proposed, which offered a document of Minimum Information About a Plant Phenotyping Experiment (MIAPPE) and recommended to use ISA-Tab formatting for organizing metadata set [129]. Moreover, unified standard formats, which can be used universally, should be developed for phenotypic data generated by different techniques. To support data reuse, the maintenance and continuously update of research data management plans within project live span and the selection of data deposition sites that ensure long-term data availability under FAIR criteria are equally important. Some progress has been made in developing publicly available phenotypic database, such as PHENOPSIS DB (http://bioweb.supagro.inra.fr/phenopsis/) for *Arabidopsis thaliana* phenotypic data [130], which can be used as a template for further advancing the development of more similar databases. Despite there has been some attempts in data management, researchers should band together to pave the way for data reuse.
(5) Continuous development of multivariate data/image analysis algorithms is essential. For instance, deep learning has exhibited excellent performance in the processing of voluminous data and image processing, owing to its unique strengths in the form of self-learning ability and efficiency

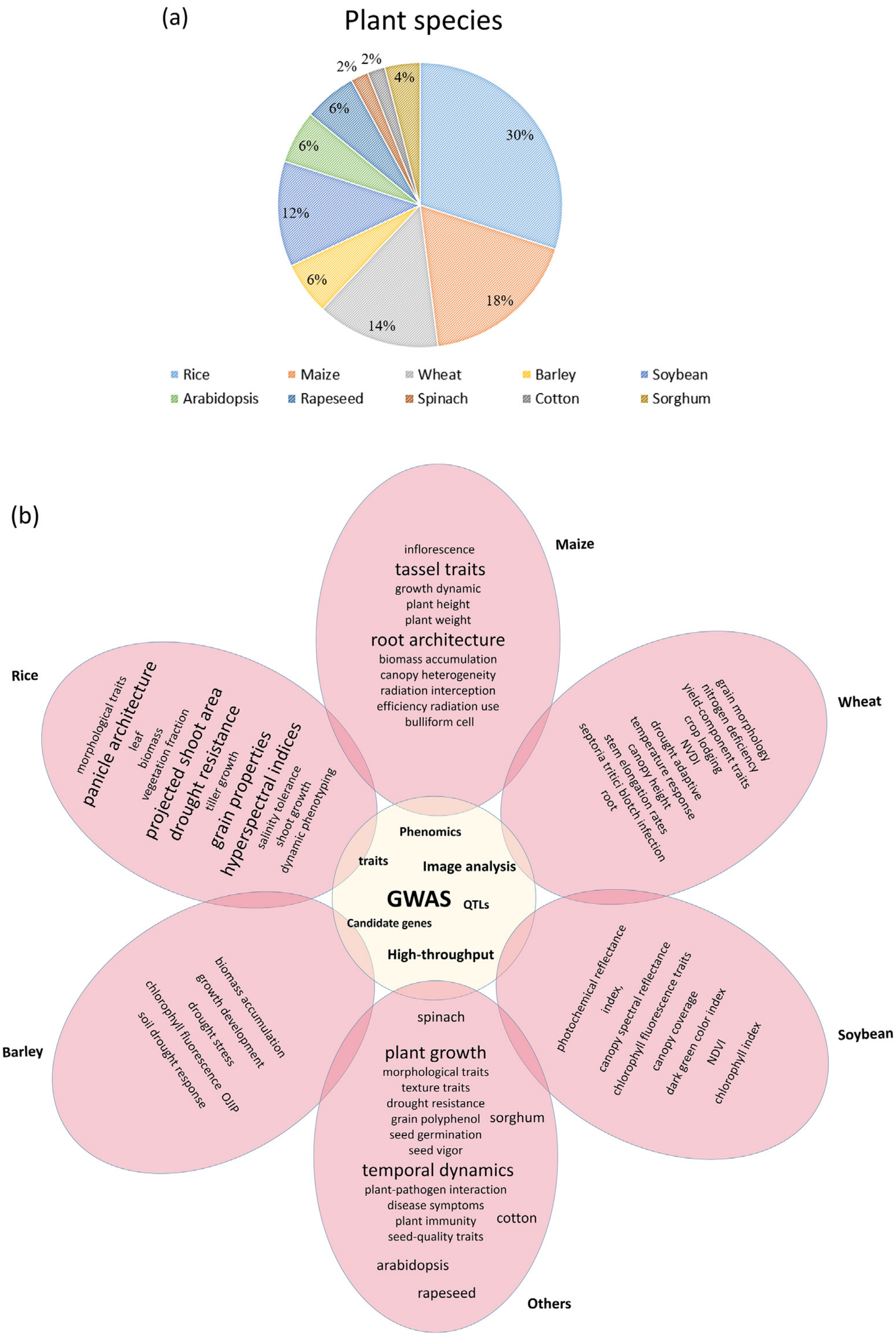


Fig. 3. The current status of GWAS equipped with high-throughput phenotyping in plants. (a) The proportion of studies using GWAS combined with high-throughput phenotyping in plant species (based on searched articles); (b) The main focus of searched articles for various plants. The bigger font size corresponding to the more high attention.

in big data analysis. The application of deep learning in plant trait data extraction will definitely become a major topic in future research.

(6) The combination of various high-throughput phenotyping techniques would facilitate more comprehensive evaluation of plant traits.

(7) Studies applying various high-throughput phenotyping techniques in combination with GWAS in more diverse plant species and traits are urgently required. More high-throughput phenotyping techniques are expected with further developments in analysis techniques and equipment.

Conclusions

This review provides an overview of the application of high-throughput phenotyping techniques and GWAS in plants. The highlighted techniques, including visible light/RGB imaging, X-ray CT, visible and near-infrared spectroscopy, multispectral/hyperspectral imaging, chlorophyll fluorescence, fluorescence imaging, and NMR, have been applied to obtain plant trait data, which are subsequently applied in GWAS. Recent studies suggest that traits obtained by high-throughput phenotyping perform similarly or even better in GWAS than those obtained by traditional, manual methods. Moreover, traits can be linked to known gene loci using high-throughput phenotyping. High-throughput phenotyping, which facilitates non-contact and dynamic measurements, has the potential to offer high-quality trait data for GWAS and, in turn, to enhance the unraveling of genetic structures of complex plant traits. Bottlenecks and challenges in the further development of high-throughput phenotyping combined with GWAS, as well as future prospects, are also discussed. The combination of high-throughput phenotyping and GWAS in linking phenotypes and genes has broad applications and is not limited to the techniques and plant traits mentioned in the present review. There are many more phenotypes to be investigated and techniques to be applied. Notably, currently, high-throughput phenotyping cannot entirely replace manual measurements, especially in the quantification of certain metabolites and indicators that are challenging to measure. We expect that high-throughput phenotyping integrated with GWAS will allow the unravelling of coding information in plant genomes and could promote plant breeding and modern genomics studies and applications.

Compliance with Ethics Requirements

This article does not contain any studies with human or animal subjects.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research has been supported by National Key R&D program of China (2018YFD0101002), National Natural Science Foundation of China (61705195), Postdoctoral Science Foundation of Zhejiang province, China (zj2019091), the Planned Science and Technology Project of Guangdong Province, China (2019B020216001), the Project of Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme, China (2016).

References

[1] Johannsen W. The genotype conception of heredity. Int J Epidemiol 2014;43:989–1000. doi: https://doi.org/10.1093/ije/dyu063.

[2] Fiorani F, Schurr U. Future scenarios for plant phenotyping. Annu Rev Plant Biol 2013;64:267–91. doi: https://doi.org/10.1146/annurev-arplant-050312-120137.

[3] Rebolledo MC, Pena AL, Duitama J, Cruz DF, Dingkuhn M, Grenier C, et al. Combining image analysis, genome wide association studies and different field trials to reveal stable genetic regions related to panicle architecture and the number of spikelets per panicle in rice. Front Plant Sci 2016;7:1384. doi: https://doi.org/10.3389/fpls.2016.01384.

[4] Honsdorf N, March TJ, Berger B, Tester M, Pillen K. High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. PLoS ONE 2013;9:. doi: https://doi.org/10.1371/journal.pone.0097047e97047.

[5] Li WT, Liu C, Liu YX, Pu ZE, Dai SF, Wang JR, et al. Meta-analysis of QTL associated with tolerance to abiotic stresses in barley. Euphytica 2013;189:31–49. doi: https://doi.org/10.1007/s10681-012-0683-3.

[6] Li Y, Xiao J, Chen L, Huang X, Cheng Z, Han B, et al. Rice functional genomics research: past decade and future. Mol Plant 2018;11:359–80. doi: https://doi.org/10.1016/j.molp.2018.01.007.

[7] Werner T. Next generation sequencing in functional genomics. Brief Bioinform 2010;11:499–511. doi: 0.1093/bib/bbq018.

[8] Clarke WE, Higgins EE, Plieske J, Wieseke R, Sidebottom C, Khedikar Y, et al. A high-density SNP genotyping array for *Brassica napus* and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. Theor Appl Genet 2016;129:1887–99. doi: https://doi.org/10.1007/s00122-016-2746-7.

[9] Awika HO, Bedre R, Yeom J, Marconi TG, Enciso J, Mandadi KK, et al. Developing growth-associated molecular markers via high-throughput phenotyping in spinach. Plant Genome 2019;12:1–19. doi: https://doi.org/10.3835/plantgenome2019.03.0027.

[10] Muraya MM, Chu J, Zhao Y, Junker A, Klukas C, Reif JC, et al. Genetic variation of growth dynamics in maize (*Zea mays* L.) revealed through automated non-invasive phenotyping. Plant J 2017;89:366–80. doi: https://doi.org/10.1111/tpj.13390.

[11] Chen TW, Cabrera-Bosquet L, Alvarez Prado S, Perez R, Artzet S, Pradal C, et al. Genetic and environmental dissection of biomass accumulation in multi-genotype maize canopies. J Exp Bot 2019;70:2523–34. doi: https://doi.org/10.1093/jxb/ery309.

[12] Jiang L, Sun L, Ye M, Wang J, Wang Y, Bogard M, et al. Functional mapping of N deficiency-induced response in wheat yield-component traits by implementing high-throughput phenotyping. Plant J 2019;97:1105–19. doi: https://doi.org/10.1111/tpj.14186.

[13] Feng H, Guo Z, Yang W, Huang C, Chen G, Fang W, et al. An integrated hyperspectral imaging and genome-wide association analysis platform provides spectral and genetic insights into the natural variation in rice. Sci Rep 2017;7:4401. doi: https://doi.org/10.1038/s41598-017-04668-8.

[14] Rhodes DH, Hoffmann LJr, Rooney WL, Ramu P, Morris GP, Kresovich S. Genome-wide association study of grain polyphenol concentrations in global sorghum [*Sorghum bicolor* (L.) Moench] germplasm. J Agric Food Chem 2014;62:10916–27. doi: https://doi.org/10.1021/jf503651t.

[15] Rahaman MM, Chen D, Gillani Z, Klukas C, Chen M. Advanced phenotyping and phenotype data analysis for the study of plant growth and development. Front Plant Sci 2015;6:619. doi: https://doi.org/10.3389/fpls.2015.00619.

[16] Mir RR, Reynolds M, Pinto F, Khan MA, Bhat MA. High-throughput phenotyping for crop improvement in the genomics era. Plant Sci 2019;282:60–72. doi: https://doi.org/10.1016/j.plantsci.2019.01.007.

[17] Moreira FF, Oliveira HR, Volenec JJ, Rainey KM, Brito LF. Integrating high-throughput phenotyping and statistical genomic methods to genetically improve longitudinal traits in crops. Front Plant Sci 2020;11:681. doi: https://doi.org/10.3389/fpls.2020.00681.

[18] Li B. Identification of genes conferring plant salt tolerance using GWAS: current success and perspectives. Plant Cell Physiol 2020:pcaa073. doi: https://doi.org/10.1093/pcp/pcaa073.

[19] Han B, Huang X. Sequencing-based genome-wide association study in rice. Curr Opin Plant Biol 2013;16:133–8. doi: https://doi.org/10.1016/j.pbi.2013.03.006.

[20] Wang Q, Tang J, Han B, Huang X. Advances in genome-wide association studies of complex traits in rice. Theor Appl Genet 2020;133:1415–25. doi: https://doi.org/10.1007/s00122-019-03473-3.

[21] Alqudah AM, Sallam A, Baenziger PS, Boerner A. GWAS: fast-forwarding gene identification and characterization in temperate cereals: lessons from barley - a review. J Adv Res 2020;22:119–35. doi: https://doi.org/10.1016/j.jare.2019.10.013.

[22] Yang W, Feng H, Zhang X, Zhang J, Doonan JH, Batchelor WD, et al. Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. Mol Plant 2020;13:187–214. doi: https://doi.org/10.1016/j.molp.2020.01.008.

[23] Mele G, Gargiulo L. Automatic cell identification and counting of leaf epidermis for plant phenotyping. MethodsX 2020;7:. doi: https://doi.org/10.1016/j.mex.2020.100860100860.

[24] Faulkner C, Zhou J, Evrard A, Bourdais G, Maclean D, Haweker H, et al. An automated quantitative image analysis tool for the identification of microtubule patterns in plants. Traffic 2017;18:683–93. doi: https://doi.org/10.1111/tra.12505.

[25] Gallegos JE, Adames NR, Rogers MF, Kraikivski P, Ibele A, Nurzynski-Loth K, et al. Genetic interactions derived from high-throughput phenotyping of 6589 yeast cell cycle mutants. Npj Syst Biol Appl 2020;6:11. doi: https://doi.org/10.1038/s41540-020-0134-z.

[26] Zhang Y, Wang JL, Du JJ, Zhao YX, Lu XJ, Wen WL, et al. Dissecting the phenotypic components and genetic architecture of maize stem vascular bundles using high-throughput phenotypic analysis. Plant Biotechnol J 2020. , https://pubmed.ncbi.nlm.nih.gov/32569428/.

[27] Baek J, Lee E, Kim N, Kim SL, Choi I, Ji H, et al. High throughput phenotyping for various traits on soybean seeds using image analysis. Sensors (Basel) 2020;20:248. doi: https://doi.org/10.3390/s20010248.

[28] Zhang CY, Si YS, Lamkey J, Boydston RA, Garland-Campbell KA, Sankaran S. High-throughput phenotyping of seed/seedling evaluation using digital

image analysis. Agronomy-Basel 2018;8:63. doi: https://doi.org/10.3390/agronomy8050063.

[29] Ligterink W, Hilhorst HWM. High-throughput scoring of seed germination. In: KleineVehn J, Sauer M, editors. Plant hormones: Methods and protocols. Netherlands: Wageningen; 2017. p. 57–72.

[30] Ma ZH, Mao YH, Gong L, Liu CL. Smartphone-based visual measurement and portable instrumentation for crop seed phenotyping. Ifac Papersonline 2016;49:259–64. doi: https://doi.org/10.1016/j.ifacol.2016.10.048.

[31] De Medeiros AD, Da Silva LJ, Pereira MD, Oliveira AMS, Dias DCFS. High-throughput phenotyping of brachiaria grass seeds using free access tool for analyzing x-ray images. An Acad Bras Cienc 2020;92:. doi: https://doi.org/10.1590/0001-3765202020190209e20190209.

[32] Jasinski S, Lecureuil A, Durandet M, Bernard-Moulin P, Guerche P. *Arabidopsis* seed content QTL mapping using high-throughput phenotyping: The assets of near infrared spectroscopy. Front Plant Sci 2016;7:1682. doi: https://doi.org/10.3389/fpls.2016.01682.

[33] Anderson JV, Wittenberg A, Li H, Berti MT. High throughput phenotyping of *Camelina sativa* seeds for crude protein, total oil, and fatty acids profile by near infrared spectroscopy. Ind Crops Prod 2019;137:501–7. doi: https://doi.org/10.1016/j.indcrop.2019.04.075.

[34] Melchinger AE, Boehm J, Utz HF, Mueller J, Munder S, Mauch FJ. High-throughput precision phenotyping of the oil content of single seeds of various oilseed crops. Crop Sci 2018;58:670–8. doi: https://doi.org/10.2135/cropsci2017.07.0429.

[35] Zhang Y, Zhang N. Imaging technologies for plant high-throughput phenotyping: a review. Front Agric Sci Eng 2018;5:406–19. doi: https://doi.org/10.15302/j-fase-2018242.

[36] Yang W, Duan L, Chen G, Xiong L, Liu Q. Plant phenomics and high-throughput phenotyping: Accelerating rice functional genomics using multidisciplinary technologies. Curr Opin Plant Biol 2013;16:180–7. doi: https://doi.org/10.1016/j.pbi.2013.03.005.

[37] Shakoor N, Lee S, Mockler TC. High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. Curr Opin Plant Biol 2017;38:184–92. doi: https://doi.org/10.1016/j.pbi.2017.05.006.

[38] Liu H, Bruning B, Garnett T, Berger B. Hyperspectral imaging and 3D technologies for plant phenotyping: from satellite to close-range sensing. Comput Electron Agric 2020;175:. doi: https://doi.org/10.1016/j.compag.2020.105621105621.

[39] Jang G, Kim J, Yu J-K, Kim H-J, Kim Y, Kim D-W, et al. Review: Cost-effective unmanned aerial vehicle (UAV) platform for field plant breeding application. Remote Sensing 2020;12:998. doi: https://doi.org/10.3390/rs12060998.

[40] Yang G, Liu J, Zhao C, Li Z, Huang Y, Yu H, et al. Unmanned aerial vehicle remote sensing for field-based crop phenotyping: current status and perspectives. Front Plant Sci 2017;8:1111. doi: https://doi.org/10.3389/fpls.2017.01111.

[41] Brezovcsik K, Veres S, Molnar J, Fenyvesi A, Szucs Z. Comparison of manganese uptake and transport of maize seedlings by mini-pet camera. Appl Radiat Isot 2020;160:. doi: https://doi.org/10.1016/j.apradiso.2020.109127109127.

[42] Qu W, Robert CAM, Erb M, Hibbard BE, Paven M, Gleede T, et al. Dynamic precision phenotyping reveals mechanism of crop tolerance to root herbivory. Plant Physiol 2016;172:776–88. doi: https://doi.org/10.1104/pp.16.00735.

[43] Dadshani S, Kurakin A, Amanov S, Hein B, Rongen H, Cranstone S, et al. Non-invasive assessment of leaf water status using a dual-mode microwave resonator. Plant Methods 2015;11:8. doi: https://doi.org/10.1186/s13007-015-0054-x.

[44] Sydoruk VA, Fiorani F, Jahnke S, Krause H-J. Design and characterization of microwave cavity resonators for noninvasive monitoring of plant water distribution. IEEE Trans Microw Theory Tech 2016;64:2894–904. doi: https://doi.org/10.1109/tmtt.2016.2594218.

[45] Herrero-Huerta M, Rodriguez-Gonzalvez P, Rainey KM. Yield prediction by machine learning from UAS-based mulit-sensor data fusion in soybean. Plant Methods 2020;16:78. doi: https://doi.org/10.1186/s13007-020-00620-6.

[46] Maimaitijiang M, Sagan V, Sidike P, Hartling S, Esposito F, Fritschi FB. Soybean yield prediction from UAV using multimodal data fusion and deep learning. Remote Sens Environ 2020;237:. doi: https://doi.org/10.1016/j.rse.2019.111599111599.

[47] Maimaitijiang M, Ghulam A, Sidike P, Hartling S, Maimaitiyiming M, Peterson K, et al. Unmanned aerial system (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. ISPRS J Photogramm Remote Sens 2017;134:43–58. doi: https://doi.org/10.1016/j.isprsjprs.2017.10.011.

[48] Bai G, Ge YF, Hussain W, Baenziger PS, Graef G. A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. Comput Electron Agric 2016;128:181–92. doi: https://doi.org/10.1016/j.compag.2016.08.021.

[49] Rischbeck P, Elsayed S, Mistele B, Barmeier G, Heil K, Schmidhalter U. Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley. Eur J Agron 2016;78:44–59. doi: https://doi.org/10.1016/j.eja.2016.04.013.

[50] Tang Y, Dananjayan S, Hou CJ, Guo QW, Luo SM, He Y. A survey on the 5g network and its impact on agriculture: challenges and opportunities. Comput Electron Agric 2021;180:. doi: https://doi.org/10.1016/j.compag.2020.105895105895.

[51] Campbell MT, Du Q, Liu K, Brien CJ, Berger B, Zhang C, et al. A comprehensive image-based phenomic analysis reveals the complex genetic architecture of shoot growth dynamics in rice (*Oryza sativa*). Plant Genome 2017;10. doi: https://doi.org/10.3835/plantgenome2016.07.0064.

[52] Campbell MT, Momen M, Walia H, Morota G. Leveraging breeding values obtained from random regression models for genetic inference of longitudinal traits. Plant Genome 2019;12. doi: https://doi.org/10.1101/435685.

[53] Bao Y, Tang L, Breitzman MW, Salas Fernandez MG, Schnable PS. Field-based robotic phenotyping of sorghum plant architecture using stereo vision. J Field Robot 2019;36:397–415. doi: https://doi.org/10.1002/rob.21830.

[54] Mochida K, Koda S, Inoue K, Hirayama T, Tanaka S, Nishii R, et al. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective. GigaScience 2019;8:giy153. doi: https://doi.org/10.1093/gigascience/giy153.

[55] Zhou L, Zhang C, Liu F, Qiu ZJ, He Y. Application of deep learning in food: a review. Compr Rev Food Sci Food Saf 2019;18:1793–811. doi: https://doi.org/10.1111/1541-4337.12492.

[56] Singh AK, Ganapathysubramanian B, Sarkar S, Singh A. Deep learning for plant stress phenotyping: trends and future perspectives. Trends Plant Sci 2018;23:883–98. doi: https://doi.org/10.1016/j.tplants.2018.07.004.

[57] Rehman TU, Ma D, Wang L, Zhang L, Jin J. Predictive spectral analysis using an end-to-end deep model from hyperspectral images for high-throughput plant phenotyping. Comput Electron Agric 2020;177:. doi: https://doi.org/10.1016/j.compag.2020.105713105713.

[58] Wang X, Xuan H, Evers B, Shrestha S, Pless R, Poland J. High-throughput phenotyping with deep learning gives insight into the genetic architecture of flowering time in wheat. GigaScience 2019;8:giz120. doi: https://doi.org/10.1093/gigascience/giz120.

[59] Yang W, Yang C, Hao ZY, Xie CQ, Li MZ. Diagnosis of plant cold damage based on hyperspectral imaging and convolutional neural network. IEEE Access 2019;7:118239–48. doi: https://doi.org/10.1109/access.2019.2936892.

[60] Lin Z, Guo W. Sorghum panicle detection and counting using unmanned aerial system images and deep learning. Front Plant Sci 2020;11:. doi: https://doi.org/10.3389/fpls.2020.534853534853.

[61] Malambo L, Popescu S, Ku N-W, Rooney W, Zhou T, Moore S. A deep learning semantic segmentation-based approach for field-level sorghum panicle counting. Remote Sensing 2019;11:2939. doi: https://doi.org/10.3390/rs11242939.

[62] Misra T, Arora A, Marwaha S, Chinnusamy V, Rao AR, Jain R, et al. SpikeSegNet-a deep learning approach utilizing encoder-decoder network with hourglass for spike segmentation and counting in wheat plant from visual imaging. Plant Methods 2020;16:40. doi: https://doi.org/10.1186/s13007-020-00582-9.

[63] Sadeghi-Tehran P, Virlet N, Ampe EM, Reyns P, Hawkesford MJ. *Deepcount*: in-field automatic quantification of wheat spikes using simple linear iterative clustering and deep convolutional neural networks. Front Plant Sci 2019;10:1176. doi: https://doi.org/10.3389/fpls.2019.01176.

[64] Hasan MM, Chopin JP, Laga H, Miklavcic SJ. Detection and analysis of wheat spikes using convolutional neural networks. Plant Methods 2018;14:100. doi: https://doi.org/10.1186/s13007-018-0366-8.

[65] Samiei S, Rasti P, Vu JL, Buitink J, Rousseau D. Deep learning-based detection of seedling development. Plant Methods 2020;16:103. doi: https://doi.org/10.1186/s13007-020-00647-9.

[66] Yang Z, Gao S, Xiao F, Li G, Ding Y, Guo Q, et al. Leaf to panicle ratio (LPR): a new physiological trait indicative of source and sink relation in japonica rice based on deep learning. Plant Methods 2020;16:117. doi: https://doi.org/10.1186/s13007-020-00660-y.

[67] Moghimi A, Yang C, Anderson JA. Aerial hyperspectral imagery and deep neural networks for high-throughput yield phenotyping in wheat. Comput Electron Agric 2020;172:. doi: https://doi.org/10.1016/j.compag.2020.105299105299.

[68] Khan Z, Rahimi-Eichi V, Haefele S, Garnett T, Miklavcic SJ. Estimation of vegetation indices for high-throughput phenotyping of wheat using aerial imaging. Plant Methods 2018;14:20. doi: https://doi.org/10.1186/s13007-018-0287-6.

[69] Hamidinekoo A, Garzon-Martinez GA, Ghahremani M, Corke FMK, Zwiggelaar R, Doonan JH, et al. A convolutional neural network based quantification of fruit number in *Arabidopsis*. GigaScience 2020;9:giaa012. doi: https://doi.org/10.1093/gigascience/giaa012.

[70] Dobos O, Horvath P, Nagy F, Danka T, Viczian A. A deep learning-based approach for high-throughput hypocotyl phenotyping. Plant Physiol 2019;181:1415–24. doi: https://doi.org/10.1104/pp.19.00728.

[71] Stewart EL, Wiesner-Hanks T, Kaczmar N, Dechant C, Wu H, Lipson H, et al. Quantitative phenotyping of northern leaf blight in UAV images using deep learning. Remote Sensing 2019;11:2209. doi: https://doi.org/10.3390/rs11192209.

[72] Dechant C, Wiesner-Hanks T, Chen SY, Stewart EL, Yosinski J, Gore MA, et al. Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. Phytopathology 2017;107:1426–32. doi: https://doi.org/10.1094/phyto-11-16-0417-r.

[73] Qiu RC, Yang C, Moghimi A, Zhang M, Steffenson BJ, Hirsch CD. Detection of fusarium head blight in wheat using a deep neural network and color imaging. Remote Sensing 2019;11:2658. doi: https://doi.org/10.3390/rs11222658.

[74] Giuffrida MV, Doerner P, Tsaftaris SA. Pheno-deep counter: A unified and versatile deep learning architecture for leaf counting. Plant J 2018;96:880–90. doi: https://doi.org/10.1111/tpj.14064.

[75] Namin ST, Esmaeilzadeh M, Najafi M, Brown TB, Borevitz JO. Deep phenotyping: Deep learning for temporal phenotype/genotype classification. Plant Methods 2018;14:66. doi: https://doi.org/10.1186/s13007-018-0333-4.

[76] Madec S, Jin XL, Lu H, De Solan B, Liu SY, Duyme F, et al. Ear density estimation from high resolution rgb imagery using deep learning technique. Agric For Meteorol 2019;264:225–34. doi: https://doi.org/10.1016/j.agrformet.2018.10.013.

[77] Zhou CQ, Hu J, Xu ZF, Yue JB, Ye HB, Yang GJ. A monitoring system for the segmentation and grading of broccoli head based on deep learning and neural networks. Front Plant Sci 2020;11:402. doi: https://doi.org/10.3389/fpls.2020.00402.

[78] Xiao Y, Liu H, Wu L, Warburton M, Yan J. Genome-wide association studies in maize: praise and stargaze. Mol Plant 2017;10:359–74. doi: https://doi.org/10.1016/j.molp.2016.12.008.

[79] Yang W, Guo Z, Huang C, Duan L, Chen G, Jiang N, et al. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. Nat Commun 2014;5:5087. doi: https://doi.org/10.1038/ncomms6087.

[80] Crowell S, Korniliev P, Falcao A, Ismail A, Gregorio G, Mezey J, et al. Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. Nat Commun 2016;7:10527. doi: https://doi.org/10.1038/ncomms10527.

[81] Yang W, Guo Z, Huang C, Wang K, Jiang N, Feng H, et al. Genome-wide association study of rice (*Oryza sativa* L.) leaf traits with a high-throughput leaf scorer. J Exp Bot 2015;66:5605–15. doi: https://doi.org/10.1093/jxb/erv100.

[82] Ogawa D, Sakamoto T, Tsunematsu H, Kanno N, Nonoue Y, Yonemaru JI. Haplotype analysis from unmanned aerial vehicle imagery of rice magic population for the trait dissection of biomass and plant architecture. J Exp Bot 2021;72:2371–82. doi: https://doi.org/10.1093/jxb/eraa605.

[83] Wu D, Guo Z, Ye J, Feng H, Liu J, Chen G, et al. Combining high-throughput micro-CT-RGB phenotyping and genome-wide association study to dissect the genetic architecture of tiller growth in rice. J Exp Bot 2018;70:545–61. doi: https://doi.org/10.1093/jxb/ery373.

[84] Sun D, Cen H, Weng H, Wan L, Abdalla A, El-Manawy AI, et al. Using hyperspectral analysis as a potential high throughput phenotyping tool in GWAS for protein content of rice quality. Plant Methods 2019;15:54. doi: https://doi.org/10.1186/s13007-019-0432-x.

[85] Barnaby JY, Huggins TD, Lee H, Mcclung AM, Pinson SRM, Oh M, et al. Vis/nir hyperspectral imaging distinguishes sub-population, production environment, and physicochemical grain properties in rice. Sci Rep 2020;10:9284. doi: https://doi.org/10.1038/s41598-020-65999-7.

[86] Al-Tamimi N, Brien C, Oakey H, Berger B, Saade S, Ho YS, et al. Salinity tolerance loci revealed in rice using high-throughput non-invasive phenotyping. Nat Commun 2016;7:13342. doi: https://doi.org/10.1038/ncomms13342.

[87] Campbell MT, Knecht AC, Berger B, Brien CJ, Wang D, Walia H. Integrating image-based phenomics and association analysis to dissect the genetic architecture of temporal salinity responses in rice. Plant Physiol 2015;168:1476–89. doi: https://doi.org/10.1104/pp.15.00450.

[88] Guo Z, Yang W, Chang Y, Ma X, Tu H, Xiong F, et al. Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. Mol Plant 2018;11:789–805. doi: https://doi.org/10.1016/j.molp.2018.03.018.

[89] Campbell MT, Grondin A, Walia H, Morota G. Leveraging genome-enabled growth models to study shoot growth responses to water deficit in rice. J Exp Bot 2020;71:5669–79. doi: https://doi.org/10.1093/jxb/eraa280.

[90] Gage JL, White MR, Edwards JW, Kaeppler S, de Leon N. Selection signatures underlying dramatic male inflorescence transformation during modern hybrid maize breeding. Genetics 2018;210:1125–38. doi: https://doi.org/10.1534/genetics.118.301487.

[91] Gage JL, de Leon N, Clayton MK. Comparing genome-wide association study results from different measurements of an underlying phenotype. G3 (Bethesda) 2018;8:3715–22. doi: https://doi.org/10.1534/g3.118.200700.

[92] Pace J, Lee N, Naik HS, Ganapathysubramanian B, Lübberstedt T. Analysis of maize (*Zea mays* L.) seedling roots with the high throughput image analysis tool *ARIA* (automatic root image analysis). PLoS ONE 2014;9:. doi: https://doi.org/10.1371/journal.pone.0108255e108255.

[93] Zheng Z, Hey S, Jubery T, Liu H, Yang Y, Coffey L, et al. Shared genetic control of root system architecture between *zea mays* and *sorghum* bicolor. Plant Physiol 2020;182:977–91. doi: https://doi.org/10.1104/pp.19.00752.

[94] Wang X, Zhang R, Song W, Han L, Liu X, Sun X, et al. Dynamic plant height QTL revealed in maize through remote sensing phenotyping using a high-throughput unmanned aerial vehicle (UAV). Sci Rep 2019;9:3458. doi: https://doi.org/10.1038/s41598-019-39448-z.

[95] Qiao P, Lin M, Vasquez M, Matschi S, Chamness J, Baseggio M, et al. Machine learning enables high-throughput phenotyping for analyses of the genetic architecture of bulliform cell patterning in maize. G3 (Bethesda) 2019;9:4235–43. doi: https://doi.org/10.1534/g3.119.400757.

[96] Zhang Y, Wang J, Du J, Zhao Y, Lu X, Wen W, et al. Dissecting the phenotypic components and genetic architecture of maize stem vascular bundles using high-throughput phenotypic analysis. Plant Biotechnol J 2020. doi: https://doi.org/10.1111/pbi.13437.

[97] Rasheed A, Xia X, Ogbonnaya F, Mahmood T, Zhang Z, Mujeeb-Kazi A, et al. Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. BMC Plant Biol 2014;14:128. doi: https://doi.org/10.1186/1471-2229-14-128.

[98] Singh D, Wang X, Kumar U, Gao L, Noor M, Imtiaz M, et al. High-throughput phenotyping enabled genetic dissection of crop lodging in wheat. Front Plant Sci 2019;10:394. doi: https://doi.org/10.3389/fpls.2019.00394.

[99] Condorelli GE, Maccaferri M, Newcomb M, Andrade-Sanchez P, White JW, French AN, et al. Comparative aerial and ground based high throughput phenotyping for the genetic dissection of NDVI as a proxy for drought adaptive traits in Durum wheat. Front Plant Sci 2018;9:893. doi: https://doi.org/10.3389/fpls.2018.00893.

[100] Kronenberg L, Yates S, Boer MP, Kirchgessner N, Walter A, Hund A. Temperature response of wheat affects final height and the timing of stem elongation under field conditions. J Exp Bot 2021;72:700–17. doi: https://doi.org/10.1093/jxb/eraa471.

[101] Yates S, Mikaberidze A, Krattinger SG, Abrouk M, Hund A, Yu K, et al. Precision phenotyping reveals novel loci for quantitative resistance to septoria tritici blotch. Plant Phenomics 2019;2019:3285904. doi: https://doi.org/10.34133/2019/3285904.

[102] Beyer S, Daba S, Tyagi P, Bockelman H, Brown-Guedira G, Iwgsc, et al. Loci and candidate genes controlling root traits in wheat seedlings-a wheat root GWAS. Funct Integr Genomics 2019;19:91–107. doi: https://doi.org/10.1007/s10142-018-0630-z.

[103] Friedli M, Kirchgessner N, Grieder C, Liebisch F, Mannale M, Walter A. Terrestrial 3D laser scanning to track the increase in canopy height of both monocot and dicot crop species under field conditions. Plant Methods 2016;12:9. doi: https://doi.org/10.1186/s13007-016-0109-7.

[104] Neumann K, Zhao Y, Chu J, Keilwagen J, Reif JC, Kilian B, et al. Genetic architecture and temporal patterns of biomass accumulation in spring barley revealed by image analysis. BMC Plant Biol 2017;17:137. doi: https://doi.org/10.1186/s12870-017-1085-4.

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

[106] Rapacz M, Wojcik-Jagla M, Fiust A, Kalaji HM, Koscielniak J. Genome-wide associations of chlorophyll fluorescence OJIP transient parameters connected with soil drought response in barley. Front Plant Sci 2019;10:78. doi: https://doi.org/10.3389/fpls.2019.00078.

[107] Fanourakis D, Briese C, Max JFJ, Kleinen S, Putz A, Fiorani F, et al. Rapid determination of leaf area and plant height by using light curtain arrays in four species with contrasting shoot architecture. Plant Methods 2014;10:9. doi: https://doi.org/10.1186/1746-4811-10-9.

[108] Kuska M, Wahabzada M, Leucker M, Dehne H-W, Kersting K, Oerke E-C, et al. Hyperspectral phenotyping on the microscopic scale: towards automated characterization of plant-pathogen interactions. Plant Methods 2015;11:28. doi: https://doi.org/10.1186/s13007-015-0073-7.

[109] Bergstraesser S, Fanourakis D, Schmittgen S, Cendrero-Mateo MP, Jansen M, Scharr H, et al. HyperART: non-invasive quantification of leaf traits using hyperspectral absorption-reflectance-transmittance imaging. Plant Methods 2015;11:1. doi: https://doi.org/10.1186/s13007-015-0043-0.

[110] Herritt M, Dhanapal AP, Fritschi FB. Identification of genomic loci associated with the photochemical reflectance index by genome-wide association study in soybean. Plant Genome 2016;9. doi: https://doi.org/10.3835/plantgenome2015.08.0072.

[111] Dhanapal AP, Ray JD, Singh SK, Hoyos-Villegas V, Smith JR, Purcell LC, et al. Genome-wide association mapping of soybean chlorophyll traits based on canopy spectral reflectance and leaf extracts. BMC Plant Biol 2016;16:174. doi: https://doi.org/10.1186/s12870-016-0861-x.

[112] Herritt M, Dhanapal AP, Purcell LC, Fritschi FB. Identification of genomic loci associated with 21chlorophyll fluorescence phenotypes by genome-wide association analysis in soybean. BMC Plant Biol 2018;18:312. doi: https://doi.org/10.1186/s12870-018-1517-9.

[113] Xavier A, Hall B, Hearst AA, Cherkauer KA, Rainey KM. Genetic architecture of phenomic-enabled canopy coverage in *glycine max*. Genetics 2017;206:1081–9. doi: https://doi.org/10.1534/genetics.116.198713.

[114] Kaler AS, Abdel-Haleem H, Fritschi FB, Gillman JD, Ray JD, Smith JR, et al. Genome-wide association mapping of dark green color index using a diverse panel of soybean accessions. Sci Rep 2020;10:5166. doi: https://doi.org/10.1038/s41598-020-62034-7.

[115] Wang L, Liu F, Hao X, Wang W, Xing G, Luo J, et al. Identification of the QTL-allele system underlying two high-throughput physiological traits in the chinese soybean germplasm population. Front Genet 2021;12:. doi: https://doi.org/10.3389/fgene.2021.600444600444.

[116] Li B, Chen L, Sun W, Wu D, Wang M, Yu Y, et al. Phenomics-based gwas analysis reveals the genetic architecture for drought resistance in cotton. Plant Biotechnol J 2020;18:2533–44. doi: https://doi.org/10.1111/pbi.13431.

[117] Knoch D, Abbadi A, Grandke F, Meyer RC, Samans B, Werner CR, et al. Strong temporal dynamics of QTL action on plant growth progression revealed through high-throughput phenotyping in canola. Plant Biotechnol J 2020;18:68–82. doi: https://doi.org/10.1111/pbi.13171.

[118] Meyer RC, Weigelt-Fischer K, Knoch D, Heuermann M, Zhao Y, Altmann T. Temporal dynamics of QTL effects on vegetative growth in *Arabidopsis thaliana*. J Exp Bot 2021;72:476–90. doi: https://doi.org/10.1093/jxb/eraa490.

[119] Fordyce RF, Soltis NE, Caseys C, Gwinner R, Corwin JA, Atwell S, et al. Digital imaging combined with genome-wide association mapping links loci to plant-pathogen interaction traits. Plant Physiol 2018;178:1406–22. doi: https://doi.org/10.1104/pp.18.00851.
[120] Martel A, Lo T, Desveaux D, Guttman DS. A high-throughput, seedling screen for plant immunity. Mol Plant Microbe Interact 2020;33:394–401. doi: https://doi.org/10.1094/MPMI-10-19-0295-TA.
[121] Hatzig SV, Frisch M, Breuer F, Nesi N, Ducoumau S, Wagner M-H, et al. Genome-wide association mapping unravels the genetic control of seed germination and vigor in *Brassica napus*. Front Plant Sci 2015;6:221. doi: https://doi.org/10.3389/fpls.2015.00221.
[122] Wang B, Wu Z, Li Z, Zhang Q, Hu J, Xiao Y, et al. Dissection of the genetic architecture of three seed-quality traits and consequences for breeding in *Brassica napus*. Plant Biotechnol J 2018;16:1336–48. doi: https://doi.org/10.1111/pbi.12873.
[123] Staedler YM, Kreisberger T, Manafzadeh S, Chartier M, Handschuh S, Pamperl S, et al. Novel computed tomography-based tools reliably quantify plant reproductive investment. J Exp Bot 2018;69:525–35. doi: https://doi.org/10.1093/jxb/erx405.
[124] Hughes N, Askew K, Scotson CP, Williams K, Sauze C, Corke F, et al. Non-destructive, high-content analysis of wheat grain traits using X-ray micro computed tomography. Plant Methods 2017;13:76. doi: https://doi.org/10.1186/s13007-017-0229-8.
[125] Borisjuk L, Rolletschek H, Fuchs J, Melkus G, Neuberger T. Low and high field magnetic resonance for in vivo analysis of seeds. Materials (Basel) 2011;4:1426–39. doi: https://doi.org/10.3390/ma4081426.
[126] Qiu RC, Wei S, Zhang M, Li H, Sun H, Liu G, et al. Sensors for measuring plant phenotyping: a review. Int J Agric Biol Eng 2018;11:1–17. doi: https://doi.org/10.25165/j.ijabe.20181102.2696.
[127] Le Guen Y, Philippe C, Riviere D, Lemaitre H, Grigis A, Fischer C, et al. eQTL of *KCNK*$_2$ regionally influences the brain sulcal widening: evidence from 15,597 UK Biobank participants with neuroimaging data. Brain Struct Funct 2019;224:847–57. doi: https://doi.org/10.1007/s00429-018-1808-9.
[128] Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. Comment: The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016;3:. doi: https://doi.org/10.1038/sdata.2016.18160018.
[129] Cwiek-Kupczynska H, Altmann T, Arend D, Arnaud E, Chen D, Cornut G, et al. Measures for interoperability of phenotypic data: Minimum information requirements and formatting. Plant Methods 2016;12:44. doi: https://doi.org/10.1186/s13007-016-0144-4.
[130] Fabre J, Dauzat M, Negre V, Wuyts N, Tireau A, Gennari E, et al. PHENOPSIS DB: an information system for *Arabidopsis thaliana* phenotypic data in an environmental context. BMC Plant Biol 2011;11:77. doi: https://doi.org/10.1186/1471-2229-11-77.

QINLIN XIAO received the B.S. degree in Food quality and safety from Southwest University, Chongqing, China, in 2018. She is currently a Ph.D. candidate in College of Biosystems Engineering and Food Science in Zhejiang University. Her research interests include the spectroscopy and spectral imaging analysis and the application in the agriculture and food.

XIULIN BAI received the B.S. degree in Agricultural mechanization and automation from Guangxi University, Nanning, China, in 2018. She is currently a Ph.D. candidate in College of Biosystems Engineering and Food Science in Zhejiang University. Her research interests include the spectroscopy and spectral imaging analysis and the application in the agriculture and food.

CHU ZHANG received the B.S. degree in Electronic and Information Engineering from Northwest AandF University, Yangling, China, in 2011, and the Ph.D. degree in agricultural electrification and automation from College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China, in 2016. He is currently a postdoctor with the College of Biosystems Engineering and Food Science, Zhejiang University. His research interests include the spectroscopy and spectral imaging analysis and the application in the agriculture and food.

YONG HE received the B.S. and M.S. degrees in Agricultural Engineering from Zhejiang Agricultural University, Hangzhou, China, in 1984 and 1987, respectively. He received Ph.D. degree in Biosystems Engineering from Zhejiang University, Hangzhou, China, in 1998. He is currently a professor in Zhejiang University. He is the dean of College of Biosystems Engineering and Food Science, Zhejiang University; Director of the Key Laboratory of spectroscopy, Ministry of Agricultural and Rural Affairs; the national prestigious teacher; one of the hundred thousands of national talents. He was selected as Clarivate Analytics Global Highly Cited Researchers in 2016-2018. He is the Editor-in-Chief of Computers and Electronics in Agriculture and editorial board member of Food and Bioprocess Technology.