DRAFT THEMATIC STUDY

for

THE THIRD REPORT ON THE STATE OF THE WORLD'S PLANT GENETIC RESOURCES FOR FOOD AND AGRICULTURE

(Not for Quotation or Citation)

Comments must be submitted by filling in the designated table

PHENOTYPING AND GENOTYPING FOR THE CONSERVATION AND SUSTAINABLE USE OF PLANT GENETIC RESOURCES FOR FOOD AND AGRICULTURE

Rajeev K. Varshney, State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Food Futures Institute, Murdoch University, Murdoch, WA 6150, Australia

Abhishek Bohra, State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Food Futures Institute, Murdoch University, Murdoch, WA 6150, Australia

Michael Abberton, Genetic Resources Center, International Institute of Tropical Agriculture (IITA), Ibadan 200001, Nigeria

Yunbi Xu, Institute of Crop Science (ICS), Chinese Academy of Agricultural Sciences (CAAS), Beijing 100081, China

Shoba Sivasankar, International Atomic Energy Agency (IAEA), Vienna International Centre, 100 1400 Vienna, Austria

Sarah Hearne, International Maize and Wheat Improvement Center (CIMMYT), Texcoco CP 56237, Mexico

Henry T. Nguyen, Division of Plant Sciences, University of Missouri, Columbia, Missouri 65211, USA

Contents

Abbreviations

APPF	Australian Plant Phenomics Facility
CBD	Convention on Biological Diversity
CIMMYT	International Centre for Maize and Wheat Improvement
CT	Computed Tomography
DArT-Seq	Diversity Arrays Technology Sequencing
DPPN	German Plant Phenotyping Network
DSI	Digital Sequence Information
EPPN	European Plant Phenotyping Network
EURISCO	European Search Catalogue for Plant Genetic Resources
FIGS	Focused Identification of Germplasm Strategy
GBS	Genotyping by Sequencing
GWAS	Genome-Wide Association Studies
HCRF	Hemispherical Conical Reflectance Factor
HTPP	High-throughput Phenotyping Platform
IPK	Leibniz-Institute of Plant Genetics and Crop Plant Research
ITPGRFA	International Treaty on Plant Genetic Resources for Food and Agriculture
LiDAR	Light Detection and Ranging
MAP	Manned Aerial Platform
MCPD	Multi-Crop Passport Descriptor
NAPPN	North American Plant Phenotyping Network
NDVI	Normalized Difference Vegetation Index
NPPN	Nordic Plant Phenotyping Network
PGRFA	Plant Genetic Resources for Food and Agriculture
QTL	Quantitative Trait Loci
REST-Seq	Restriction Fragment Sequencing
RESCAN	Restriction Enzyme Site Comparative Analysis
RGB	Red, Green, and Blue
RSA	Root System Architecture
SBG	Sequence Based Genotyping
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
TILLING	Targeting Induced Local Lesions in Genomes
WGRS	Whole Genome Resequencing
WUE	Water Use Efficiency

Abstract

Conserving and utilizing diverse plant genetic resources is crucial for achieving global food security targets. Accurate phenotyping and genotyping of large sets of plant genetic resources for food and agriculture (PGRFA) can help characterize the inherent and enormous diversity of these resources for their better management and efficient utilization in crop improvement. Modern DNA sequencing and genotyping technologies enabling simultaneous identification and mapping of thousands of genetic markers in hundreds of samples have facilitated detailed characterization of large PGRFA collections and breeding populations across the globe. This has been exemplified by the detailed genetic profiling of germplasm accessions, breeding lines and varieties belonging to, not only major crops, but also less studied and so-called orphan crops. Hence, these large-scale sequencing and genotyping efforts have rendered novel genetic diversity accessible for research and breeding of crop species. Concurrent developments in plant phenotyping have also greatly strengthened trait mapping, gene discovery, crop breeding and germplasm management efforts. However, the high-volume and complexity of phenotypic and genotypic datasets presents unprecedented challenges with respect to analysis of large data sets, storage, access and sharing. This thematic background study highlights the major trends in genotyping and phenotyping of PGRFA since 2010. The key breakthroughs related to high-density genotyping and large-scale phenotyping of both natural diversity panels and artificially-created mapping populations are discussed. This review also underscores the future requirements for the efficient screening and management of PGRFA to accelerate their sustainable use.

1. Introduction

Measurement methods that describe variations in plant phenotypes are the cornerstone of plant breeding. Plant phenotypes are determined by genetic, environment and genotype by environment (G×E) interactions. Phenotyping refers to evaluation of "structural, physiological, and performance-related traits" of plants (Großkinsky et al. 2015). Crop phenotypic performance involves a complex interaction between genotypes and environmental factors, which include climate, soil factors, abiotic/biotic factors, and crop management methods.

Over the past two decades, the characterization and evaluation of PGRFA was conducted using mainly field-based measurements focused on visually observable morphological traits with high heritability. Manual measurements that formed the basis of plant evaluations were cumbersome, time-intensive, destructive and often subjective. Harnessing the immense potential of PGRFA relies, to a significant extent, on the accuracy and precision of the methods and tools being used for their phenotypic assessment. To this end, non-invasive phenotyping systems driven largely by the image and sensor technologies facilitating accurate and precise measurements in both controlled and field conditions, are contributing to overcome the bottlenecks of traditional phenotyping methods.

The idea of using genotype as a screening tool for genetic variations and a surrogate for variations in external phenotypes has revolutionized crop improvement. Genotyping refers to a procedure that assays individual's DNAs to examine their differences at genetic levels. DNA marker technologies have given great impetus to efforts aimed at genetic screening. Ease of assaying DNA at reduced cost and greater scales drove the innovations in DNA marker systems, leading a recent shift from low-throughput PCR-based systems.

Since 2010, DNA marker technologies have witnessed remarkable improvements in terms of their throughput, accuracy, reproducibility, ease of assaying and cost effectiveness. Concurrent with evolving technologies in general, and in the DNA sequencing landscape in particular, methods for identifying and mapping genetic markers in the genome have also undergone a paradigm shift to support genetic mapping and breeding efforts. The advent of next generation sequencing (NGS) and third generation sequencing (TGS) technologies, with their constant refinements in their throughput, scalability, resolution,

time and cost, has resulted in another rapid increase in both the quantity and quality of genome assemblies reported for land plants.

Alongside the development of high-density genotyping technologies, the notable improvements in plant phenotyping capacities, especially in terms of mobility, affordability, throughput, accuracy, scalability, data storage and analytics, have significantly enhanced the ability to bridge the gap between genetic information and the phenotypic manifestations of that information (Zhu et al. 2021). Moreover, the acquisition of high-quality phenotyping data from large samples is of paramount significance to realize the full potential of cutting-edge genotyping technologies for crop improvement. This thematic background study highlights the major trends in the adoption of modern genotyping and phenotyping platforms to facilitate sustainable use of plant genetic resources for food and agriculture (PGRFA).

2. Advances in large-scale phenotyping of PGRFA

The latest developments in plant phenotyping, driven largely by the evolving sensor and imaging technologies, have helped relieve the long-standing "phenotyping bottlenecks" in plant research and breeding. PGRFA characterization has greatly benefited from the increased scale and accuracy of advanced plant phenotyping methods and technologies that can record morphological, physiological and biochemical changes at relatively high resolutions (Li et al. 2021). Recent developments in robotics, artificial intelligence, advanced imaging and data analysis have paved the way for the increased adoption of PGRFA phenotyping in an automated/semi-automated, cost-efficient, non-destructive and non-invasive fashion (Table 1).

2.1 High-throughput Phenotyping Platforms

The development of high-throughput plant phenotyping platforms (HTPPs), with higher accuracy and precision, represent a breakthrough in plant biology and genetic research. Increased understanding of diversity within and among accessions by using high-quality molecular data and accurate phenotypic values facilitates the selection and use of valuable genetic variation in crop improvement as well as transforming genebank management practises.

HTPPs involve the accurate acquisition and analysis of numerous multi-dimensional phenotypes during various plant growth stages at the organism level, including the cell,

tissue, organ, individual plant, plot, and field levels (Song et al., 2021). HTPPs employ multiple advanced sensors (Bohra et al. 2021) and are equipped with data collection algorithms to perform large-scale phenotyping tasks. The sensors operate on various technologies that capture images as reflectance, multispectral, hyperspectral, red green and blue imaging (RGB), light detection and ranging (LiDAR), and infrared imaging.

One example of a High-throughput Phenotyping Platform is TraitMill, developed by the private company CropDesign, and was among the first automated platforms that allowed the high-throughput evaluation of yield parameters on transgenic rice lines (Reuzeau et al. 2010). In the public sector, the establishment of the Australian Plant Phenomics Facility (APPF) pioneered state-of-the-art phenomics technologies involving hyperspectral, choloryphyll fluorescence and RGB imaging using the platforms PlantScan and TrayScan, developed by Photon Systems Instruments at the High Resolution Plant Phenomics Center in Canberra, Australia as part of the Commonwealth Scientific and Industrial Research Organisation (CSIRO).

2.2 HTPP application and advancements in multi-trait capturing using automated sensors and computations

Research in different crops has demonstrated the use of sensors in large-scale plant phenotyping, including maize and sorghum (Shafiekhani et al. 2017), wheat (Deery et al. 2019; Madec et al. 2017), cotton (Andrade-Sanchez et al. 2014), and soybean (Sagan et al. 2019). Sensor systems such as the conveyor belt system, operating on the 'plant-to-sensor' mode, captures images of the crop by keeping plants on a moving belt that passes in front of a camera stand, ensuring consistent imaging (Li et al. 2021). For example, the conveyor belt system was used to study the effects of zinc and an arbuscular mycorrhizal fungus on tomato (Brien et al. 2020). The LemnaTec Field Scanalyzer also uses the 'sensor-to-plant' mode and has been employed to accurately monitor crop performance via measuring plant height, spike number, and the normalized difference vegetation index (Virlet et al. 2017). Several studies using the Scanalyzer 3D have also demonstrated the growing potential of imaging techniques for accurate plant phenotyping (Nguyen et al. 2019).

The Plant Accelerator in Adelaide, Australia, is part of the APPF and is a cutting-edge plant phenotyping system equipped with multiple sensors that has been used in phenotyping different crops, including rice (Mazis et al. 2020) and chickpea (Atenio et al. 2017). Other examples employing high-throughput image analysis include the screening of

biomass accumulation in barley cultivars under drought stress (Neumann et al. 2015) and the monitoring of water-use dynamics for quantifying leaf water content in maize (Ge et al. 2016).

Obtaining visual representations of water and other elements within a plant could help in optimizing the breeding and management strategies in relation to fertilization and irrigation, breeding trial conductance and stress conditions like drought. The application of hyperspectral imaging in wheat has led researchers to develop distribution maps for water and nitrogen content within plants (Bruning et al. 2019). Similarly**,** digital features extracted from image data were used for the quantitative measurements of the drought response in rice and the analysis was extendable to other gramineous crops, such as maize (Duan et al. 2018). However, a crucial factor in the adoption of HTPPs is the degree of correlation between the traits measured by sensor systems and the resulting plant performance in the field (Bai et al. 2019).

Computed tomography (CT) is another method used to provide a thorough evaluation of the external and internal features encompassing plant morphology and anatomy. In wheat, 3D morphological images created by X-ray micro-CT were used to accurately identify spike kernels (Liu et al. 2020). Advanced procedures and tools, such as RootAnalyzer (Chopin et al. 2015) and RootScan (Burton et al. 2013), have also uncovered the diversity of root traits through the analysis of cross-section images in wheat and maize, respectively.

Portable instruments are widely used to measure crop phenotypic traits owing to the ease in operation, portability and low cost (Yang et al. 2020). However, their applications for high-throughput phenotyping are limited by detection scale and efficiency. Mobile phenotyping systems with multiple sensors have been successfully used to collect plot-level, blur-free phenotypic data in soybean and wheat (Bai et al. 2016).

Monitoring of plant growth and yield in sorghum during the 1998 growing season by a Cessna 206 aircraft was among the first examples of airborne digital imagery in plant phenotyping (Yang et al. 2000). Aerial plant phenotyping systems with 3D imaging generated accurate data in potatoes, capturing plant height and canopy coverage (Xie et al. 2022). Similarly, thermal imaging data from a manned aerial platform (MAP) resulted in the acquisition of data on canopy temperature covering thousands of research plots, thus, demonstrating the potential airborne thermography for the rapid measurement of crop phenotypic traits (Deery et al. 2016).

Applications of crop phenomics or high-throughput crop phenotyping in different species were reviewed by Yang et al. (2020). The author reviewed related articles in Web of Science, which the following order: wheat (30 percent), rice (13 percent), maize (17 percent), *Arabidopsis* (13 percent), barley (7 percent), and others (20 percent). A bibliometric analysis by Costa et al. (2019) based on the publications on plant phenotyping research made over the last 20 years (1997-2017) showed that majority of the analysed papers (41.8 percent) were authored by the EU researchers followed by the USA (15.4 percent), Australia (6 percent) and India (5.6 percent).

A dramatic rise in the use of modern plant phenotyping, with more than 90 percent of the publications available since 2010. Plant phenotyping systems were expanded from *Arabidopsis*, a model plant, to real crops over the past two decades. Importantly, the analysis of the trend highlighted the pressing need to expand the knowledge, infrastructures and expertise and international collaborations that provides greater opportunities for participation of the researchers, organizations and countries from the developing world.

2.3 Global, regional and local phenotyping networks

Collaborations among the public and private scientific sectors have resulted in the formation of plant phenotyping networks that use advanced technologies to monitor plant growth and development at different levels. One such plant phenotyping network is the International Plant Phenotyping Network (IPPN) that involves various regional and national partners: EMPHASIS , European Plant Phenotyping Network (EPPN), North American Plant Phenotyping Network (NAPPN), Nordic Plant Phenotyping Network (NPPN), PhenomUK, German Plant Phenotyping Network (DPPN), and the PHENOME Networks (https://www.plant-phenotyping.org/; https://phenome-networks.com/en/). Other automated phenotyping facilities include the Bellwether phenotyping platform at the Donald Danforth Plant Science Center, USA and the high-throughput rice phenotyping platform (HRPF; Yang et al. 2014).

3. Advances in genotyping of PGRFA

Recent advances in genotyping have facilitated rapid generation of high-throughput marker data in a cost-effective manner. The growing affordability to advanced genotyping technologies has contributed to DNA analysis of large samples, leading to their comprehensive genetic characterisation. Enhanced ability to describe genetic differences

in diverse collections and artificial populations has accelerated population genetics studies and trait mapping and germplasm management efforts.

3.1 High-density single nucleotide polymorphism arrays

The increasing availability of reference genome sequences over the past decade, in combination with resequencing of multiple genomes, has given a great impetus to the discovery and application of single nucleotide polymorphism (SNP) markers in plant genetic research and molecular breeding (Mammadov et al., 2012). The need to genotype SNPs in large samples for genetic diversity, trait mapping and gene discovery has led to the establishment of a variety of high-throughput genotyping platforms with varying levels of multiplexing, such as Illumina GoldenGate, Illumina Infinium BeadChips, and more recently, the Kompetitive allele specific PCR (KASP™) assay (Mir et al., 2013).

The identification of large-scale high-quality SNP datasets, following the resequencing of several genomes in different crop species, has facilitated the development of fixed SNP arrays for genetic diversity analysis, the construction of high and ultra-high density genetic maps and the identification of genome-wide marker-trait associations for important agronomic traits (Thomson et al., 2017). A review by Rasheed et al. (2017) discussed advances and challenges associated with the development and application of diverse high-throughput genotyping platforms in different crop species. The high cost associated with the development of crop-specific SNP arrays, along with ascertainment bias towards variants present in the population, which are involved in array development limits, have led to the use of fixed arrays for population genetic studies (Geibel et al. 2021).

Recent examples of the establishment of genome-wide SNP arrays in crop plants include 50 thousand in rice (Chen et al. 2014, Singh et al. 2015), 660 thousand (Sun et al. 2020) and 820 thousand (Winfield et al. 2016) in wheat, 600 thousand in maize (Unterseer et al. 2014), 50 thousand in barley (Bayer et al. 2017), 480 thousand in apple (Bianco et al. 2016), 60 thousand in Brassica (Clarke et al. 2016), 63 thousand (Hulse-Kemp et al. 2015) and 80 thousand (Cai et al. 2017) in cotton, 50 thousand in chickpea (Roorkiwal et al. 2018), 50 thousand in cowpea (Muñoz-Amatriaín et al. 2017), and 58 thousand in groundnut (Pandey et al. 2017).

3.2 Low-depth whole genome resequencing

Advances in DNA sequencing technologies have paved the way for the adoption of sequence-based methods for plant genotyping. The sequence-based methods enable the

simultaneous discovery and the mapping of large numbers of genetic markers within a short period of time. For instance, Huang et al. (2009) pioneered the research on demonstrating the utility of sequence-based methods in the genetic mapping of a biparental population. The study showed that sequencing-based genotyping is 20 times faster in data collection relative to the traditional PCR-based DNA marker systems and was 35 times more precise in determining recombination breakpoints. Since then, WGRS has been employed for understanding genetic diversity, evolutionary and domestication history as well as functional gene discovery across several crop species (Varshney et al., 2017a, 2017b, 2019). Additionally, WGRS of diverse accessions has allowed high-resolution identification of gene-trait associations underlying important plant phenotypes through genome-wide association studies GWAS on rice, chickpea, soybean, pigeon pea, common bean, linseed and *Brassica* (see Bohra et al. 2020 and references therein).

Several other methods that use NGS for genotyping have been proposed in recent years, which include diversity array technology sequencing (DArT-Seq), sequence-based genotyping (SBG), restriction fragment sequencing (REST-Seq), restriction enzyme site comparative analysis (RESCAN), and specific-locus amplified fragment sequencing (SLAF-Seq). Of these GBS, DArT-Seq and SLAF-Seq have dominated the genetic research in plants over the last decade (Rasheed et al. 2017).

3.3 Moderate to low-density genotyping assays for breeding applications

For the purpose of molecular breeding, customized genotyping platforms that accommodate a finite number of SNP markers can offer significant cost savings over using high-density genotyping platforms. One of such genotyping platforms, the KASP™, represents a significant improvement in allele-specific PCR and has garnered wide recognition and application amongst plant scientists for tracking and stacking genomic loci. The flexibility and efficiency of the KASP™ platform has allowed its widespread use for varietal identification (Tang et al. 2022) and genomics-assisted breeding (Saxena et al. 2021).

4. Application of phenotyping and genotyping for the conservation and sustainable use of PGFRA

The advances in phenomics and genomics have enhanced our capacity to characterize and use the germplasm collections for trait discovery and varietal improvement. Ready access to PGRFA, particularly CWR and landraces that contain valuable genetic diversity,

is of vital importance to food security and climate adaptation of future crops. This was evident from elucidation of genetic contributions of GKP 10017, an accession of wild groundnut (*Arachis cardenasii*), towards 251 lines and cultivars in 30 countries in Africa, Asia, Oceania and the Americas. The genetic influence of GKP 10017 improved the tolerance level of groundnut lines and varieties against various diseases and pest including root knot nematode, late leaf spot and rust (Bertioli et al. 2021). In this study establishment of the chromosome-scale genome assembly of GKP 10017 in combination of diagnostic SNPs and pedigree research facilitated a systematic analysis of the genetic exchanges following international collaborations.

4.1 Genotyping applied to the conservation and management of PGRFA in genebanks

Genetic marker technologies have been applied for the estimation of diversity in PGRFA collections (Varshney et al. 2021a; Bohra et al. 2022a, b). As previously mentioned, advances in DNA sequencing technologies have led to a dramatic shift in the scale of molecular characterization of germplasm collections and breeding populations. In Table 2, selected DNA sequencing and genotyping technologies have been summarized, which facilitated the large-scale profiling of germplasm collections in different crops. For example, Milner et al. (2019) analysed collections at the Leibniz-Institute of Plant Genetics and Crop Plant Research (IPK)'s *ex situ* genebank that encompasses cultivars, landraces and crop wild relatives by using more than 100,000 SNP markers that were obtained by the GBS technique. Similarly, Singh et al. (2019) employed DArT-Seq and GBS platforms to characterize 1,143 *Aegilops tauschii* accessions from the Wheat Genetics Resource Center (WGRC) in the United States of America, the International Maize and Wheat Improvement Center (CIMMYT) in Mexico and Punjab Agricultural University (PAU) in India.

The past decade has witnessed an increased application of genome sequencing for the in-depth genetic characterization of reference genotypes and large crop germplasm collections (Box 1). The latest examples using WGRS include cassava [241 accessions (Ramu et al. 2017) and 388 accessions (Hu et al. 2021)], sunflower (493 accessions, Hübner et al. 2019), common bean (683 accessions, Wu et al. 2020), chickpea (429 accessions, Varshney et al. 2019), pigeon pea (292 accessions, Varshney et al., 2017a), pearl millet (994 accessions, Varshney et al. 2017b) and lettuce (445 accessions, Wei et al. 2021). Besides WGRS, high-density genotyping platforms (GBS, DArT-Seq, SNP arrays, etc.) have also been employed for the characterization of diverse germplasm across different crop species,

as exemplified by the 80 000 accessions of wheat (Sansaloni et al. 2020), 2,815 accessions of maize (Romay et al. 2013), 14 000 accessions of soybean (Bandillo et al. 2015), 1 628 accessions of sorghum (Girma et al. 2020), 10 038 accessions of pepper (Tripodi et al. 2021) and 271 accessions of cassava (Bredeson et al. 2016).

Pairwise identity-by-state (IBS) comparisons following high-density genetic profiling of germplasm collections can also help identification of near-identical samples in large collections. This approach was applied to the study of genetic diversity, intra- and inter-populations, of two durum wheat landraces "Russello" and "Timilia" from Sicily (Taranto et al. 2022). Results showed that the Russello landrace was in fact two genetic groups, differing in important traits related to gluten quality and adaptation. This study highlighted that a more in-depth assessment of genetic diversity should be undertaken in future breeding programs. It also indicated that intra-population genetic diversity should be considered for conservation and documentation efforts.

Sequencing and genotyping germplasm can significantly contribute towards the identification of gaps in germplasm collections and mislabelling of biological status in genebank historical records. This was exemplified by the genome analysis of chickpea accessions that elucidated a wild-specific allele of the *SHATTERPROOF2* homolog, which assisted in the identification and correction of the mislabelling of the chickpea accession ICC 16369 (Varshney et al. 2021b).

Furthermore, the development of molecular passport data is used to augment collection management efforts by identifying different conservation gaps and duplicated accessions as well as correcting historical passport records. For instance, the analysis of 1 143 *Aegilops tauschii* accessions from WGRC, CIMMYT and PAU using genome-wide SNPs has allowed the identification of 564 accessions as unique, thus, revealing more than 50 percent redundancy levels among the *Ae. Tauschii* collections that were hosted in the three genebanks (Singh et al. 2019). Similarly, high-density genotyping of the IPK barley collections data helped in correcting the biological status of Ethiopian accessions from 'wild' to 'landrace' in the genebank information system (Milner et al. 2019).

4.2. Genotyping applied to the characterization and evaluation for sustainable use of PGRFA

The availability of the high-density sequencing and genotyping data on large germplasm collections will guide strategies for creation of the diversity sets for their enhanced use in plant breeding programs. Based on the genome-wide marker information, genomic

prediction could support optimizing selection strategies for choosing worthy germplasm accessions in the absence of phenotypic information. For example, by generating GBS data on 962 sorghum accessions, Yu et al. (2016) investigated the efficacy of genomic prediction as a novel and cost-effective strategy for "turbocharging" genebanks. A training set of 299 accessions with information on both its phenotype and genotype was used to assess this approach. Likewise in wheat, the genomic prediction involving GBS-analysed 7 745 germplasm accessions and 325 cultivars, which enabled the identification of suitable parents and donors for grain yield and resistance to yellow rust (Schulthess et al. 2022). Hence, availability of detailed information related to genotype and phenotype of the accessions would unleash the true breeding potential of each seed stored in the genebank for future use.

Molecular characterization was applied to several subsets of different crop species in Italy. Researchers at the Research Centre for Olive, Fruit and Citrus Crops undertook the molecular assessment of about 400 peach accessions using IPSC 9K SNP array (Micheletti et al. 2015, Verde et al. 2012). The entire peach collection (about 900 accessions) had also been characterized using SSR markers. Furthermore, about 400 apple local accessions of Central Italy were characterized with 20K SNP arrays. Other studies have also focused on wheat, where a subset of 200 bread wheat accessions was analyzed by SNPs and phenotyped for relevant agronomic and qualitative traits (Lazzaro et al. 2019, Ormoli et al. 2015, Talini et al. 2020).

Researchers at the Swiss National Genebank carried out analysis of 502 bread wheat (*Triticum aestivum)* and 293 spelt (*Triticum aestivum spelta*) accessions using a 15K SNP array. It highlighted the importance of old landraces as promising sources of novel genes (Müller et al. 2017). Genomic data in combination with expert knowledge considerations facilitated the development of core collections for apple (Broggini et al. 2022), pear (Urrestarazu et al. 2019) and chestnut (Pereira-Lorenzo et al. 2020). In France, the COREPOM project (2011-2014), funded by the Foundation for Research in Biodiversity, enabled the genetic characterization of various collections of genetic resources of horticultural crops including apple for developing a core-collection that represents the cultivated biodiversity at the national scale (Roux-Cuvelier et al. 2021).

In summary, notable advances in plant genotyping and phenotyping capacities over the last ten years have opened up possibilities of having detailed genotype-phenotype maps of thousands of accessions to enable the selection of suitable accessions even in the absence

of phenotypic information. Acquisition of large-scale genotypic and phenotypic information on germplasm sets may also help develop genetic sequencing and prediction models to assess the value of a germplasm accession that lacks morphological characterization data.

4.3 Applying focused identification of germplasm strategy with genotyping

In the context of climate change adaptation, the focused identification of germplasm strategy (FIGS) enhances the possibilities of finding adaptive traits through customization of large germplasm collections using an ecogeographic approach (Bohra et al. 2022a, b). The use of genotyping on subsets identified using FIGS have allowed for the rapid identification of accessions with traits of interest. For example, Bhullar et al. (2009) performed allele mining for powdery mildew resistance gene *Pm3* in a subset of 1 320 landraces of wheat that was customized by applying FIGS approach on a large genebank collection of 16 089 accessions. Analysis of wheat germplasm collection using FIGS has identified subsets for specific traits such as resistances to Russian wheat aphid (El Bouhssini et al. 2010), stem rust (Endresen et al. 2012), yellow or stripe rust (Bari et al. 2014). Similarly, FIGS facilitated identification of sources for resistance to net blotch in barley (Endresen et al. 2011) and drought adaptation in faba bean (Khazaei et al. 2013).

As discussed by Bohra et al. (2022a), the recent research has underscored the great scope for the implementation of emerging machine learning algorithms for creation of FIGS sets to facilitate selection and breeding decisions. A combination of FIGS and core collection approach has been found to be more effective in creation of objective-driven core subsets to accelerate utilization of germplasm collection in crop improvement. For instance, Haupt and Schmid (2020) applied FIGS-informed core collection approach in > 17 000 soybean accessions from the USDA Soybean Germplasm Collection and identified two diversity panel of 183 and 366 accessions for future utilization in abiotic stress adaptation in soybean.

4.4 Direct application of genotyping to crop improvement

The use of PGRFA, particularly the crop wild relatives (CWRs) and landraces, for broadening the genetic base of breeding programs faces a variety of challenges. Pre-breeding and germplasm enhancement programs provide an avenue to devise strategies to accelerate PGRFA use for varietal improvement. The identification and introgression of quantitative trait loci (QTL) in different genetic backgrounds, particularly from wild

germplasm often results in unpredictable breeding outcomes (Bohra et al. 2022a). There remains tremendous scope for combining this approach with other breeding methods such as those involving multiple diverse parents e.g. wild nested association mapping (NAM) populations in barley (Nice et al. 2017).

The development of multi-parent populations (MPPs) takes advantage of the availability of cost-effective genotyping in combination with the need for high-resolution trait mapping has stimulated the efforts to generate such valuable mapping resources across different crops (Bohra et al. 2020). These MPP s harbouring genes for biotic and abiotic stress tolerance are important to incorporate yield stability and climate change adaptation of future cultivars, such as in the case of (*Hordeum vulgare* L.) with cases, such as scald, powdery mildew, leaf rust and Fusarium head blight (Novakazi et al. 2020).

5. Database development to accelerate PGRFA use

Recent advances in the field of omics, which is a term used to encompass different fields of biology, such as genomics, transcriptomics, proteomics, metabolomics and phenomics, have created novel avenues to accelerate PGRFA use by integrating data from multiple omics platforms. For example, the DivSeek International Network (https://divseekintl.org/) intends to harness omics and big data technologies with the help of crop researchers, genebank managers, database and computational experts. There has been a greater emphasis on providing all germplasm and associated information, such as seed or propagation materials, passport data, characterization data and evaluation data held by the genebanks, as open-access to the public for research, breeding and education. Various types of descriptors of internationally-agreed standards, such as the multi-crop passport descriptor (MCPD V.2.1; Alercia et al., 2015), have been developed and adopted to facilitate this exchange of the information related to plant genetic resources.

5.1 Current examples of genetic databases and genebanks providing genetic passport data for PGRFA

The information on several of these subsets can be obtained from online platforms that store details about PGRFAs from genebanks worldwide. For example, Genesys (https://www.genesys-pgr.org/) is a global platform providing passport data of about 4 185 326 accessions with 256 subsets, including core and mini-core collections as well as 444 characterization and evaluation sets of 40 crops. This also includes accessions with specific traits, such as tolerance to biotic and abiotic stresses, photoperiod insensitivity and quality. Core collections are available for sorghum (2 246 accessions), pearl millet (2 094

accessions), soybean (small seeded: 1 466 accessions; large seeded: 111 accessions), subterranean clover (97 accessions), cassava (629 accessions), whereas mini core collections have been reported in rice (600 accessions), *Oryza glaberrima* (350 accessions), *Triticum timopheevii* (92 accessions), and *A. tauschii* (40 accessions), cowpea (376 accessions). As demonstrated in rice (Kumar et al. 2020), wheat (Pascual et al. 2020) and common bean (Kuzay et al. 2020), generation of large-scale sequencing and genotyping data in recent years has facilitated more detailed investigations of the existing diversity panels, which has in turn allowed optimization of these subsets to have better representations of the genetic diversity of the crop species. Similarly, the germplasm resource information system (GRIN)-global (https://www.grin-global.org/) has been developed in a joint effort by the Global Crop Diversity Trust, Bioversity International, and the Agricultural Research Service of the United States Department of Agriculture (USDA) that is being used and evaluated by various national and international genebanks worldwide.

5.2 Characterization and evaluation data

Characterization and evaluation data are made available through the European Search Catalogue for Plant Genetic Resources (EURISCO), which also provides passport data related to nearly over 2 million accessions conserved *ex situ* by about 400 institutes from 43 member countries. Such data will also soon be available through GENBIS, the NordGen search catalogue (https://www.nordic-baltic-genebanks.org/gringlobal/about). A number of national genebanks also have searchable databases. For example, molecular and morphological characterization of diverse accessions of the national PGRFA genebank of Switzerland is freely available (https://www.pgrel.admin.ch/pgrel/), where specific traits can be queried and any registered user can create a virtual subset/list for the desired traits. Similarly, the Government of Poland reported that the characterization and evaluation data of 83 percent (16 506 accessions) of their national grasses collection has been made available at the NRI National Centre for Plant Genetic Resources (https://wyszukiwarka.ihar.edu.pl/pl).

The availability of global and independent information systems enhances the interoperability among genebank passport, phenotypic and genotypic data for improved germplasm management and use. In addition, the generation of massive sequencing information and other omics data on germplasm resources has led to the development of digital sequence information (DSI), which entails information, *inter alia*, from DNA and RNA sequences, protein sequences to metabolites and other macromolecules. It may also

include associated information and traditional knowledge (http://www.fao.org/3/cc6622en/cc6622en.pdf). Fair and equitable sharing of DSI arising from germplasm resources remains arguable in the context of biodiversity-related international instruments, such as the Convention on Biological Diversity (CBD) and the Nagoya Protocol and International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA). Therefore, evidence-informed discussions, training programs, coordinated efforts and favourable policy regimes are needed to address the challenges associated with access and benefit-sharing arising from germplasm in the current digitization era, particularly in relation to the developing countries that lag behind in the utilization of new technologies and the acquisition of biological data.

6. International and national initiatives in PGRFA phenotyping and genotyping

International collaborations have increasingly used whole genome sequencing approach and other high-density genotyping platforms for in-depth genetic characterization of the large crop germplasm collections. High-throughput DNA sequencing has allowed screening of genetic diversity in large sets of germplasm accessions, unleashing the potential of "Germplasm genomics" for plant improvement in the post-NGS sequence era. Sharing of DNAs and associated information is imperative to understand within-species diversity. Recent efforts on genome sequencing of hundreds of diverse accessions of a species, as witnessed in rice (Wang et al. 2018) and chickpea (Varshney et al. 2021b) foster strong collaborative environments to leverage phylogenetic research for guiding future improvement strategies.

The utility of the large germplasm collections was greatly constrained by the lack of knowledge about the traits for genetic and breeding potential, with <1 percent accessions used in crop improvement. Searching traits or alleles from large germplasm collections for breeding is complicated by resource- and labour-intensive nature of knowledge generation and germplasm management. In this context, acquisition and access to the quality data on germplasm characterization and evaluation paves the way for creation of trait-specific collection subsets. There is a need to improve the generation and quality of characterization and evaluation information, implement mechanisms for documentation and access to information, so that it can be used in plant breeding programs. Workable subsets of germplasms capturing sufficient genetic variation have been created for efficient use in breeding in rice, maize soybean, common bean, chickpea, groundnut, pigeon pea, cassava, cowpea, bean, sorghum and millets (Guo et al. 2014, Kuzay et al 2020).

7. Conclusions

Diversity in phenotypes and genetic architectures of germplasm archived in genebanks serves as a raw material for selection and breeding. Estimation of this diversity is thus crucial to accelerate their utilization in research and breeding, while guiding conservation and management efforts. The progress has been greatly constrained by the challenges encountered in germplasm characterization efforts through conventional phenotyping and genotyping methods. The increase in accuracy and scales of both plant phenotyping and genetic profiling over the last decade has been remarkable. Evolving landscape of sequencing and genotyping technologies have facilitated low-cost and time-efficient analysis of large plant DNA samples. The concurrent advances in plant phenotyping systems have complemented modern genetic profiling technologies for providing new insights into genomic architecture of germplasm resources. International collaborations involving multi-disciplinary teams are key to harness the enormous potential of genotyping and phenotyping advances.

Modern sequencing technologies have massively scaled-up the discovery of genome-wide variants, including SSR, but more importantly, with the use of SNP markers. With the cost-effectiveness of evolving sequencing and genotyping platforms, it is now possible to genotype large samples in a cost-effective, automated and accurate manner. Generation of the large-scale genotypic information in combination with phenotypic and other datasets provides novel opportunities for incorporating new breeding techniques like GS and prediction models to guide germplasm selection, conservation and management efforts (Yu et al. 2016; Schulthess et al. 2022). Equally importantly, detailed characterization of PGRFA creates novel avenues for the development of manageable subsets of germplasm accessions for research and breeding.

Table 1. List of selected high-throughput phenotyping platforms and example studies on different crop species.

HTPPs	Phenotype measured	Crops	References
LemnaTec Scanalyzer 3D	Biomass, leaf water content, plant height, width, compactness, drought stress, nitrogen content, chlorophyll status	Wheat, barley, sorghum, maize, chickpea, oak	Neumann et al. (2015), Chai et al. (2021), Mazis et al. (2020), Ge et al. (2016)
HRPF	Drought stress, tiller number	Rice	Yang et al. (2014), Duan et al. (2018)
PHENOARCH	Growth rate of ear and silk	Maize	Brichet et al. (2017)
GLYPH	Water use efficiency (WUE), drought stress	Soybean	Peirone et al. (2018)
GiARoots	Root system architecture (RSA)	Rice	Galkovskyi et al. (2012)
DART; Ez-Rhizo; DIRT; ROOTNAV	RSA	Legumes and cereals	Atkinson et al. (2019)
Crop Phenology Recording System (CPRS)	Nitrogen content, leaf area index	Rice	Fukatsu et al. (2012)
Field Scanalyzer	Canopy height, spike number, canopy closure, canopy temperature, Normalized Difference Vegetation Index (NDVI)	Wheat	Virlet et al. (2017)
LeasyScan	Canopy transpiration, plant height, leaf area, WUE	Peanut, cowpea, pearl millet, maize	Sunil et al. (2018), Vadez et al. (2015)
Phénofield	Water stress, nitrogen stress	Wheat	Beauchene et al. (2019)

Mini-Plot	Disease severity	Barley	Thomas et al. (2018)
NU-Spidercam	Plant height, ground cover, canopy temperature	Soybean, maize	Bai et al. (2019)
BreedVision	Moisture content, lodging, tiller density or biomass yield	Wheat	Busemeyer et al. (2013)
Phantom (UAV)	Carotenoid content	Lettuce	Mascarenhas Maciel et al. (2019)
3DR Solo Quadcopter (UAV)	NDVI, chlorophyll red-edge index (CHL), hemispherical-conical reflectance factors (HCRF)	Maize	Fawcett et al. (2020)
Octorotor (UAV)	Canopy height, canopy coverage, vegetation index	Rice	Wan et al. (2020)
Matrice 600 Pro (UAV)	Yield	Cotton	Feng et al. (2020)
Ebee, UAV	Yield	Wheat	Hu et al. (2020)
Robinson R44 Raven helicopter, (MAP)	Canopy temperature	Wheat	Deery et al. (2016)
Air Tractor AT-402B, (MAP)	Pest severity	-	Yang and Hoffmann (2015)
RapidEye, (Satellite)	Nitrogen stress	Wheat	Basso et al. (2016)
Fluorescence explorer (Satellite)	Terrestrial vegetation (Photosynthesis)	-	Drusch et al. (2017)
Phenomobile Lite	Plant height, biomass, ground cover	Wheat	Jimenez-Berni et al. (2018)
GPhenoVision	Canopy height, width, growth rate, projected leaf area, volume, yield	Cotton	Jiang et al. (2018)
phenoSeeder	Seed morphological and chemical constituent of seed	Rapeseed, barley	Jahnke et al. (2016)

Table 2. Characterization of large germplasm collections using sequencing and high-density genotyping.

Crop	Accessions	Genotyping method	Reference
Barley	22 626	Genotyping by sequencing (GBS)	Milner et al. (2019)
Chickpea	3 366	Whole genome resequencing (WGRS)	Varshney et al. (2021b)
Maize	2 815	GBS	Romay et al. (2013)
Pepper	10 038	GBS	Tripodi et al. (2021)
Rice	3 134	Diversity arrays technology sequencing (DArTseq)	Ndjiondjop et al. (2018)
Rice	3 010	WGRS	Wang et al. (2018)
Sorghum	1 628	GBS	Girma et al. (2020)
Soybean	14 000	50K SNP chip	Bandillo et al. (2015)
Wheat	80 000	DArT-seq	Sansaloni et al. (2020)
Wild wheat (*Aegilops tauschii*)	1 143	GBS	Singh et al. (2019)
Cassava	388	WGRS	Hu et al. (2021)
Common bean	683	WGRS	Wu et al. (2020)
Tomato	838	WGRS	Zhou et al. (2022)
Wheat	8 000	GBS	Schulthess et al. (2022)
Lettuce	445	WGRS	Wei et al. (2021)
Sorghum	400	WGRS	Boatwright et al. (2022)

Box 1. Developing next-generation genomic resources for key African crops

In the Africa region, the African Orphan Crops Consortium (AOCC) aims at understanding and exploiting the genetic diversity of key crops for providing adapted and high-yielding varieties for integration into African farming systems (http://africanorphancrops.org/about/). The consortium has started developing genome assemblies of 101 crops including okra, onion, amaranth, custard apple, papaya, watermelon. A complete list of crops selected for genome sequencing can be found at https://africanorphancrops.org/meet-the-crops/. Equally importantly, AOCC via African Plant Breeding Academy (AfPBA) initiative, led by UC Davis, focuses on building long-term capacity in Africa to efficiently use modern genomic and breeding tools for enhancing crop variety development programs.

References

Alercia, A., Diulgheroff, S. and Mackay, M. (2015). FAO/Bioversity Multi-Crop Passport Descriptors V.2.1 [MCPD V.2.1]. Rome, FAO and Bioversity International, 11 p.

Andrade-Sanchez, P., Gore, M. A., Heun, J. T. et al. (2014) Development and evaluation of a field-based high-throughput phenotyping platform. Func. Plant Biol. 41: 68–79.

Atieno, J., Li, Y., Langridge, P. et al. (2017) Exploring genetic variation for salinity tolerance in chickpea using image-based phenotyping. Sci. Rep. 7:1-11.

Atkinson, J. A., Pound, M. P., Bennett, M. J. and Wells, D. M. (2019) Uncovering the hidden half of plants using new advances in root phenotyping. Curr. Opin. Biotechnol. 55: 1-8.

Bai, G., Ge, Y., Hussain, W. et al. (2016) A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. Comput. Electron. Agric. 128: 181-192

Bai, G., Ge, Y., Scoby, D. et al. (2019) NU-Spidercam: a large-scale, cable-driven, integrated sensing and robotic system for advanced phenotyping, remote sensing, and agronomic research. Comp. Electron. Agric. 160: 71–81.

Bandillo, N., Jarquin, D., Song, Q. et al. (2015) A population structure and genome-wide association analysis on the USDA soybean germplasm collection. Plant Genome 8:eplantgenome2015.04.0024.

Bari, A., Street, K., Mackay, M. et al. (2012) Focused identification of germplasm strategy (FIGS) detects wheat stem rust resistance linked to environmental variables. Genet. Resour. Crop Evol. 59: 1465–1481.

Basso, B., Fiorentino, C., Cammarano, D. and Schulthess, U. (2016) Variable rate nitrogen fertilizer response in wheat using remote sensing. Precis. Agric. 17: 168–182.

Bayer, M.M., Rapazote-Flores, P., Ganal, M. et al. (2017) Development and Evaluation of a Barley 50k iSelect SNP Array. Front. Plant Sci. 8:1792.

Beauchene, K., Leroy, F., Fournier, A. et al. (2019) Management and characterization of abiotic stress via phenofield (R), a high-throughput field phenotyping platform. Front. Plant Sci. 10:904.

Bertioli, D.J., Clevenger, J., Godoy, I.J. et al. (2021) Legacy genetics of *Arachis cardenasii* in the peanut crop shows the profound benefits of international seed exchange. Proc. Natl. Acad. Sci. U. S. A. 118:e2104899118.

Bhullar, N.K., Street, K., Mackay, M., et al. (2009) Unlocking wheat genetic resources for the molecular identification of previously undescribed functional alleles at the Pm3 resistance locus. Proc. Natl. Acad. Sci. U. S. A. 106:9519-24.

Boatwright, J.L., Sapkota, S., Jin, H. et al. (2022) Sorghum Association Panel whole-genome sequencing establishes cornerstone resource for dissecting genomic diversity. Plant J. 111:888-904.

Bohra, A., Jha, U.C., Godwin, I., and Varshney, R.K. (2020) Genomic interventions for sustainable agriculture. Plant Biotechnol. J. 18: 2388-2405

Bohra, A., Kilian, B., Sivasankar, S. et al. (2022a) Reap the crop wild relatives for breeding future crops. Trends Biotechnol. 40:412-431.

Bohra, A., Tiwari, A., Kaur, P., et al. (2022b) The key to the future lies in the past: Insights from grain legume domestication and improvement should inform future breeding strategies. Plant Cell Physiol. 63:1554-1572.

Bohra, A., Satheesh Naik, S.J., Kumari, A. et al. (2021) Integrating phenomics with breeding for climate-smart agriculture. In: Omics Technologies for Sustainable Agriculture and Global Food Security (Vol II) (Kumar A, Kumar R, Shukla P, Patel HK Eds). Springer, Singapore. Pages 1-24.

Bredeson, J.V., Lyons, J.B., Prochnik, S.E. et al. (2016) Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. Nat. Biotechnol. 34:562-70.

Brichet, N., Fournier, C., Turc, O. et al. (2017) A robot-assisted imaging pipeline for tracking the growths of maize ear and silks in a high-throughput phenotyping platform. Plant Methods 13:96.

Brien, C., Jewell, N., Watts-Williams, S. J., Garnett, T. and Berger, B. (2020) Smoothing and extraction of traits in the growth analysis of noninvasive phenotypic data. Plant Methods 16: 1-21.

Broggini, G.A.L. Kölliker, R., Kellerhals, M. and Studer, B. (2022) Exploiting the Swiss Apple Core Collection for breeding tomorrow's apple cultivars. XXXI International Horticultural Congress (IHC2022): International Symposium on Conservation and Sustainable Use of Horticultural Genetic Resources. Angers, France, August 14-20, 2022

Bruning, B., Liu, H., Brien, C. et al. (2019) The development of hyperspectral distribution maps to predict the content and distribution of nitrogen and water in wheat (Triticum aestivum). Front Plant Sci. 10: 1380.

Burton, A. L., Brown, K. M. and Lynch, J. P. (2013) Phenotypic diversity of root anatomical and architectural traits in Zea species. Crop Sci. 53:1042-1055.

Busemeyer, L., Mentrup, D., Moller, K. et al. (2013) BreedVision–a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors 13:2830–2847.

Cai, C., Zhu, G., Zhang, T. and Guo, W (2017) High-density 80 K SNP array is a powerful tool for genotyping G. hirsutum accessions and genome analysis. BMC Genomics 18:654.

Chai, Y. N., Ge, Y., Stoerger, V. and Schachtman, D. P. (2021) High-resolution phenotyping of sorghum genotypic and phenotypic responses to low nitrogen and synthetic microbial communities. Plant Cell Env. 44:1611-1626.

Chen, H., Xie, W., He, H. et al. (2014) A high-density SNP genotyping array for rice biology and molecular breeding. Mol. Plant. 7:541-53.

Chopin, J., Laga, H., Huang, C. Y., Heuer, S. & Miklavcic, S. J. (2015) RootAnalyzer: a cross-section image analysis tool for automated characterization of root cells and tissues. PloS One 10: e0137655.

Clarke, W.E., Higgins, E.E., Plieske, J. et al. (2016) 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. 129: 1887–1899.

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A. et al. (2016) Methodology for high-throughput field phenotyping of canopy temperature using airborne thermography. Front. Plant Sci. 7:1808.

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A. et al. (2019) Evaluation of the phenotypic repeatability of canopy temperature in wheat using continuous-terrestrial and airborne measurements. Front. Plant Sci. 10:875.

Drusch, M., Moreno, J., Del Bello, U. et al. (2017) The fluorescence EXplorer mission concept—ESA's earth explorer 8. IEEE Trans. Geosci. Remote Sens. 55: 1273–1284.
Duan, L., Han, J., Guo, Z. et al. (2018) Novel digital features discriminate between drought resistant and drought sensitive rice under controlled and field conditions. Front. Plant Sci. 9:492.

El Bouhssini, M., Street, K., Amri, A. et al. (2011) Sources of resistance in bread wheat to Russian wheat aphid (*Diuraphis noxia*) in Syria using the Focused Identification of Germplasm Strategy (FIGS). Plant Breed. 130:96–97

Endresen, D.T.F., Street, K., Mackay, M. et al. (2012) Sources of resistance to stem rust (Ug99) in bread wheat and durum wheat identified using focused identification of germplasm strategy. Crop Sci. 52:764.

Endresen, D.T.F., Street, K., Mackay, M., Bari, A. and De Pauw, E. (2011) Predictive association between biotic stress traits and ecogeographic data for wheat and barley landraces. Crop Sci. 51: 2036–2055.

Fawcett, D., Panigada, C., Tagliabue, G. et al. (2020) Multi-scale evaluation of drone-based multispectral surface reflectance and vegetation indices in operational conditions. Remote Sens. 12:514.

Feng, A., Zhou, J., Vories, E. D., Sudduth, K. A. and Zhang, M. (2020) Yield estimation in cotton using UAV-based multi-sensor imagery. Biosyst. Eng. 193: 101–114.

Fukatsu, T., Watanabe, T., Hu, H., Yoichi, H. and Hirafuji, M. (2012) Field monitoring support system for the occurrence of Leptocorisa chinensis Dallas (Hemiptera: Alydidae) using synthetic attractants, field servers, and image analysis. Comput. Electron. Agric. 80:8–16.

Galkovskyi, T., Mileyko, Y., Bucksch, A. et al. (2012) GiA Roots: software for the high throughput analysis of plant root system architecture. BMC Plant Biol. 12:116.

Ge, Y., Bai, G., Stoerger, V. and Schnable, J. C. (2016) Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. Comput. Electron. Agric. 127: 625-632.

Geibel, J., Reimer, C., Weigend, S., Weigend, A., Pook, T., and Simianer, H. (2021) How array design creates SNP ascertainment bias. PLoS One 16:e0245178.

Girma, G., Nida, H., Tirfessa, A., et al. (2020) A comprehensive phenotypic and genomic characterization of Ethiopian sorghum germplasm defines core collection and reveals rich genetic potential in adaptive traits. Plant Genome 13:e20055.

Großkinsky DK, Svensgaard J, Christensen S, Roitsch T (2015) Plant phenomics and the need for physiological phenotyping across scales to narrow the genotype-to-phenotype knowledge gap. J. Exp. Bot. 66:5429-40.

Guo, Y., Li, Y., Hong, H. and Qiu, L.J. (2014) Establishment of the integrated applied core collection and its comparison with mini core collection in soybean (*Glycine max*). Crop J. 2:38-45.

Haupt, M. and Schmid, K. (2020) Combining focused identification of germplasm and core collection strategies to identify genebank accessions for central European soybean breeding. Plant Cell Environ. 43:1421-1436.

Hu, W., Ji, C., Liang, Z., et al. (2021) Resequencing of 388 cassava accessions identifies valuable loci and selection for variation in heterozygosity. Genome Biol. 22:316.

Hu, Y., Knapp, S., and Schmidhalter, U. (2020) Advancing high-throughput phenotyping of wheat in early selection cycles. Remote Sens. 12:574.

Huang, X., Feng, Q., Qian, Q et al. (2009) High-throughput genotyping by whole-genome resequencing. Genome Res. 19:1068-76.

Hübner, S., Bercovich, N., Todesco, M. et al. (2019) Sunflower pan-genome analysis shows that hybridization altered gene content and disease resistance. Nat. Plants 5:54-62.

Hulse-Kemp, A.M., Lemm, J., Plieske, J. et al. (2015) Development of a 63K SNP Array for Cotton and High-Density Mapping of Intraspecific and Interspecific Populations of Gossypium spp. G3 5:1187-209.

Jahnke, S., Roussel, J., Hombach, T. et al. (2016) PhenoSeeder - a robot system for automated handling and phenotyping of individual seeds. Plant Physiol. 172:1358–1370.

Jiang, Y., Li, C., Robertson, J. S., Sun, S., Xu, R., and Paterson, A. H. (2018) GPhenovision: a ground mobile system with multi-modal imaging for field-based high throughput phenotyping of cotton. Sci. Rep. 8:1213.

Jimenez-Berni, J. A., Deery, D. M., Rozas-Larraondo, P. et al. (2018) High throughput determination of plant height, ground cover, and above-ground biomass in wheat with LiDAR. Front. Plant Sci. 9:237.

Kumar, A., Kumar, S., Singh, K.B.M., Prasad, M. and Thakur, J.K. (2020) Designing a mini-core collection effectively representing 3004 diverse rice accessions. Plant Commun. 1:100049.

Kuzay S, Hamilton-Conaty P, Palkovic A, Gepts P (2020) Is the USDA core collection of common bean representative of genetic diversity of the species, as assessed by SNP diversity? Crop Sci. 60:1398–1414.

Kuzay, S., Hamilton-Conaty, P., Palkovic, A. and Gepts, P. (2020) Is the USDA core collection of common bean representative of genetic diversity of the species, as assessed by SNP diversity? Crop Sci. 60:1398–1414.

Lazzaro M, Bàrberi P, Dell'Acqua M, et al. (2019) Unraveling diversity in wheat competitive ability traits can improve integrated weed management. Agron. Sustain. Develop. 39:6.

Li, D., Quan, C., Song, Z. et al. (2021) High-throughput plant phenotyping platform (HT3P) as a novel tool for estimating agronomic traits from the lab to the field. Front. Bioeng. Biotechnol. 8: 623705.

Liu, W., Liu, C., Jin, J. et al. (2020) High-throughput phenotyping of morphological seed and fruit characteristics using X-ray computed tomography. Front. Plant Sci. 11: 601475.

Madec, S., Baret, F., de Solan, B. et al. (2017). High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground LiDAR estimates. Front. Plant Sci. 8:2002.

Mammadov, J., Aggarwal, R., Buyyarapu, R., & Kumpatla, S. (2012) SNP markers and their impact on plant breeding. Int. J. Plant Genomics 2012: 728398.

Mascarenhas Maciel, G., Gallis, R.B.D.A., Barbosa, R. L. et al. (2019) Image phenotyping of inbred red lettuce lines with genetic diversity regarding carotenoid levels. Int. J. Appl. Earth Observ. Geoinform 81: 154–160.

Mazis, A., Choudhury, S. D., Morgan, P. B. et al. (2020) Application of high-throughput plant phenotyping for assessing biophysical traits and drought response in two oak species under controlled environment. For. Ecol. Manag. 465: 118101.

Micheletti, D., Dettori, M.T., Micali, S. et al. (2015) Whole-genome analysis of diversity and SNP-major gene association in peach germplasm. PLoS One 10: e0136803.

Milner, S.G., Jost, M., Taketa, S., et al. (2019) Genebank genomics highlights the diversity of a global barley collection. Nat. Genet. 51:319-326.

Mir RR, Hiremath PJ, Riera-Lizarazu O, Varshney RK (2013) Evolving Molecular Marker Technologies in Plants: From RFLPs to GBS. In: Lübberstedt T, Varshney RK Eds., Diagnostics in Plant Breeding, Springer, Berlin, pp 229-247.

Muñoz-Amatriaín, M., Mirebrahim, H., Xu P. et al. (2017) Genome resources for climate-resilient cowpea, an essential crop for food security. Plant J. 89:1042-1054.

Ndjiondjop, M.N., Semagn, K., Zhang, J. et al. (2018) Development of species diagnostic SNP markers for quality control genotyping in four rice (Oryza L.) species. Mol. Breed. 38:131.

Neumann, K., Klukas, C., Friedel, S. et al. (2015) Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis. Plant Cell Env. 38:1980-1996.

Nguyen, G. N., Maharjan, P., Maphosa, L., et al. (2019). A robust automated image-based phenotyping method for rapid vegetative screening of wheat germplasm for nitrogen use efficiency. Front. Plant Sci. 10: 1372.

Nice, L.M., Steffenson, B.J., Brown-Guedira, G.L. et al. (2016) Development and genetic characterization of an advanced backcross-nested association mapping (AB-NAM) population of wild × cultivated barley. Genetics 203:1453-67.

Novakazi F, Krusell L, Jensen JD, Orabi J, Jahoor A, Bengtsson T, On Behalf Of The Ppp Barley Consortium (2020) You Had Me at "MAGIC"! Four Barley MAGIC Populations Reveal Novel Resistance QTL for Powdery Mildew. Genes (Basel) 11:1512.

Ormoli L., Costa C., Negri S., Perenzin M. and Vaccino P. (2015) Diversity trends in bread wheat in Italy during the 20th century assessed by traditional and multivariate approaches. Sci. Rep. 5: 8574

Pandey, M.K., Agarwal, G., Kale, S.M. et al. (2017) Development and Evaluation of a High Density Genotyping 'Axiom_Arachis' Array with 58 K SNPs for Accelerating Genetics and Breeding in Groundnut. Sci. Rep. 7:40577.

Pascual, L., Fernandez, M., Aparicio, N. et al. (2020) Development of a multipurpose core collection of bread wheat based on high-throughput genotyping data. Agronomy 10:534

Peirone, L. S., Pereyra Irujo, G. A., Bolton, A., Erreguerena, I., and Aguirrezábal, L. A. (2018) Assessing the efficiency of phenotyping early traits in a greenhouse automated platform for predicting drought tolerance of soybean in the field. Front. Plant Sci. 9:587.

Pereira-Lorenzo, S., Bischofberger, Y., Conedera, M. et al. (2020) Reservoir of the European chestnut diversity in Switzerland. Biodivers. Conserv. 29: 2217-2234.

Ramu, P., Esuma, W., Kawuki, R. et al. (2017) Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation. Nat. Genet. 49:959-963.

Rasheed, A., Hao, Y., Xia X. et al. (2017) Crop Breeding Chips and Genotyping Platforms: Progress, Challenges, and Perspectives. Mol. Plant 10:1047-1064.

Reuzeau, C., Pen, J., Frankard, V. et al. (2010) TraitMill: a discovery engine for identifying yield-enhancement genes in cereals. Mol. Plant Breed. 5:753–759.

Romay, M.C., Millard, M.J., Glaubitz, J.C. et al. (2013) Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 14:R55.

Roorkiwal, M., Jain, A., Kale, S.M. et al. (2018) Development and evaluation of high-density Axiom® CicerSNP Array for high-resolution genetic mapping and breeding applications in chickpea. Plant Biotechnol J. 16:890-901.

Roux-Cuvelier, M., Grisoni, M., Bellec, A. et al. (2021) Conservation of horticultural genetic resources in France. Chron. Hortic. 61: ffhal03372865ff

Sagan, V., Maimaitijiang, M., Sidike, P. et al. (2019) UAV/Satellite multiscale data fusion for crop monitoring and early stress detection. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-2/W13: 715–722.

Sansaloni, C., Franco, J., Santos, B. et al. (2020) Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. Nat. Commun. 11:4572.

Saxena, R.K., Hake, A., Bohra, A., et al. (2021) A diagnostic marker kit for fusarium wilt and sterility mosaic diseases resistance in pigeonpea. Theor. Appl. Genet. 134: 367–379

Schulthess, A.W., Kale, S.M. et al. (2022) Genomics-informed prebreeding unlocks the diversity in genebanks for wheat improvement. Nat. Genet. 54:1544-1552.

Shafiekhani, A., Kadam, S., Fritschi, F. B., and Desouza, G. N. (2017) Vinobot and vinoculer: two robotic platforms for high-throughput field phenotyping. Sensors 17:214.

Singh, N., Jayaswal, P.K., Panda, K. et al. (2015) Single-copy gene based 50 K SNP chip for genetic studies and molecular breeding in rice. Sci. Rep. 5:11600.

Singh, N., Wu, S., Raupp, W.J., et al. (2019) Efficient curation of genebanks using next generation sequencing reveals substantial duplication of germplasm accessions. Sci. Rep. 9:650.

Song, P., Wang, J., Guo, X., Yang, W., & Zhao, C. (2021). High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. Crop J. 9 : 633-645.

Sun, C., Dong, Z., Zhao, L. et al. (2020) The Wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat. Plant Biotechnol. J. 18:1354-1360.

Sunil, N., Rekha, B., Yathish, K., Sehkar, J., and Vadez, V. (2018) LeasyScan-an efficient phenotyping platform for identification of pre-breeding genetic stocks in maize. Maize J. 7: 16–22.

Talini, R.F., Brandolini, A., Miculan, M. et al. (2020) Genome-wide association study of agronomic and quality traits in a world collection of the wild wheat relative Triticum urartu. Plant J. 102: 555-568.

Tang, W., Lin, J., Wang, Y., et al. (2022) Selection and Validation of 48 KASP Markers for Variety Identification and Breeding Guidance in Conventional and Hybrid Rice (Oryza sativa L.). Rice 15:48.

Taranto, F., Di Serio, E., Miazzi, M.M. et al. (2022) Intra- and inter-population genetic diversity of "Russello" and "Timilia" landraces from Sicily: A proxy towards the identification of favorable alleles in durum wheat. Agronomy 12:1326.

Thomas, S., Behmann, J., Steier, A. et al. (2018) Quantitative assessment of disease severity and rating of barley cultivars based on hyperspectral imaging in a non-invasive, automated phenotyping platform. Plant Methods 14:45.

Thomson, M.J., Singh, N., Dwiyanti, M.S. et al. (2017) Large-scale deployment of a rice 6 K SNP array for genetics and breeding applications. Rice, 10: 40. https://doi.org/10.1186/s12284-017-0181-2

Tripodi, P., Rabanus-Wallace, M.T. Barchi, L. et al. (2021) Global range expansion history of pepper (Capsicum spp.) revealed by over 10,000 genebank accessions. Proc. Natl. Acad. Sci. U S A 118 :e2104315118.

Unterseer, S., Bauer, E., Haberer, G. (2014) A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. BMC Genomics 15:823.

Vadez, V., Kholova, J., Hummel, G. et al. (2015) LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget. J. Exp. Bot. 66: 5581–5593.

Varshney, R.K., Bohra, A., Yu, J. et al. (2021a) Designing future crops: genomics-assisted breeding comes of age. Trends Plant Sci. 26: 631-649

Varshney, R.K., Roorkiwal, M., Sun, S. et al. (2021b) A chickpea genetic variation map based on the sequencing of 3,366 genomes. Nature 599(7886):622-627.

Varshney, R.K., Saxena, R.K., Upadhyaya, H.D. et al. (2017a) Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. Nat. Genet. 49:1082-1088.

Varshney, R.K., Shi, C., Thudi, M., et al (2017b) Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. Nat. Biotechnol. 35:969-976.

Varshney, R.K., Thudi, M., Roorkiwal, M. et al. (2019) Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. Nat. Genet. 51:857-864.

Verde, I., Bassil, N., Scalabrin, S. et al. (2012) Development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. PLoS One 7: e35668.

Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P., and Hawkesford, M.J. (2017) Field Scanalyzer: an automated robotic field phenotyping platform for detailed crop monitoring. Funct. Plant Biol. 44:143–153.

Wan, L., Cen, H., Zhu, J. et al. (2020) Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer – a case study of small farmlands in the South of China. Agric. Forest Meteorol. 291:108096.

Wang, W., Mauleon, R., Hu, Z. et al. (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature 557(7703):43-49.

Wei, T., van Treuren, R., Liu, X. et al. (2021) Whole-genome resequencing of 445 Lactuca accessions reveals the domestication history of cultivated lettuce. Nat. Genet. 53:752-760.

Winfield, M.O., Allen, A.M., Burridge, A.J. et al. (2016) High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. Plant Biotechnol. J. 14:1195-206.

Wu, J., Wang, L., Fu, J., et al. (2020) Resequencing of 683 common bean genotypes identifies yield component trait associations across a north–south cline. Nat. Genet. 52:118-125.

Xie, J., Zhou, Z., Zhang, H. et al. (2022) Combining canopy coverage and plant height from UAV-based RGB images to estimate spraying volume on potato. Sustainability 14:6473.

Yang C., Everitt, J.H., Bradford, J.M. and Escobar, D.E. (2000) Mapping grain sorghum growth and yield variations using airborne multispectral digital imagery. Trans. ASAE 43: 1927–1938

Yang, C. and Hoffmann, W. C. (2015) Low-cost single-camera imaging system for aerial applicators. J. Appl. Remote Sens. 9:096064.

Yang, W., Feng, H., Zhang, X., et al. (2020) Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. Mol. Plant 13:187-214.

Yang, W., Guo, Z., Huang, C. et al. (2014) Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. Nat. Commun. 5:1-9.

Yu, X., Li, X., Guo, T. et al. (2016) Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nat. Plants 2:1615

Zhou, Y., Zhang, Z., Bao, Z. et al. (2022) Graph pangenome captures missing heritability and empowers tomato breeding. Nature 606:527-534.

Zhu, Y., Sun, G., Ding, G. et al. (2021) Large-scale field phenotyping using backpack LiDAR and CropQuant-3D to measure structural variation in wheat. Plant Physiol. 187:716-738.

Annex: Glossary of Terms

Crop Wild Relatives (CWRs): Close relatives of cultivated species, representing a rich source of genetic diversity for crop improvement.

Diversity Panel: A collection comprising accessions representing a diverse population of a species.

Digital Sequence Information (DSI): A placeholder term that covers nucleotide sequence data (DNA/RNA) to other omics information (transcriptomes, proteins and metabolites) associated with genetic resources.

Focused Identification of Germplasm Strategy (FIGS): An approach that efficiently identifies suitable germplasm with desirable traits e.g., disease and pest resistance and adaptation traits by using geographic and agro-climatic information.

Genome-wide association studies (GWAS): An approach that tests association of a genetic marker with traits of interest from a diverse set of genetic material.

High-throughput phenotyping: Automated and non-invasive methods that record precise and accurate measurements on hundreds of plants at different levels of biological organization.

KASP™: A high precision technology used for genotyping individuals based on polymerase chain reaction specific to an SNP or indel. The assay uses two allele-specific forward primers and one common reverse primer for genotyping.

Landrace: A heterogenous form of domesticated species that is well adapted to local environments.

Nested association mapping (NAM): A multi-parent strategy for dissection of complex traits that uses series of interconnected RILs derived from crossing a common parent with diverse founders.

Next generation sequencing (NGS): A group of high throughput technologies that enables sequencing of millions of small fragments of DNA in a highly paralleled fashion.