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

EVOLUTION AND APPLICATIONS OF NOVEL BIOTECHNOLOGIES AND THEIR IMPACTS ON THE CONSERVATION AND SUSTAINABLE USE OF PLANT GENETIC RESOURCES FOR FOOD AND AGRICULTURE (PGRFA)

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

Vanika Garg, [1]State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Food Futures Institute, Murdoch University, Murdoch, WA 6150, Australia

Rishi K. Tyagi, Asia-Pacific Association of Agricultural Research Institutions (APAARI), Bangkok 10100, Thailand

Anuradha Agrawal, ICAR-National Bureau of Plant Genetic Resources (ICAR-NBPGR), New Delhi 110012, India

Sunil Archak, ICAR-National Bureau of Plant Genetic Resources (ICAR-NBPGR), New Delhi 110012, India

Leena Tripathi, International Institute of Tropical Agriculture (IITA), Nairobi 00100, Kenya

Mario Caccamo, National Institute of Agricultural Botany (NIAB), Cambridge CB3 0LE, UK

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

Abstract

Achieving the food security targets for the global population requires innovations related to tools and strategies used for the conservation and use of plant genetic resources for food and agriculture (PGRFA). The last decade has witnessed the evolution and adoption of novel biotechnologies that have led to significant improvements in the conservation and management of PGRFA, including their accelerated use in crop breeding programs. The advances in targeted genome modification have been remarkable. Concurrent advances in DNA banking and cryopreservation offer tremendous opportunities to transform genebanks to 'biodigital resource centers' for effective utilization of PGRFA. Here, we review the latest trends in the application of emerging biotechnologies for crop improvement and PGRFA germplasm management.

1. Introduction

Global food demand is projected to increase by 35 to 56 percent between 2010 and 2050 (van Dijk et al. 2021). The current rate of crop improvement in not sufficient to meet the food demand projected for 10 billion people in 2050. Furthermore, about 663 million people representing 8.9 percent of the total population are undernourished, which means that their diets fail to meet the minimum energy requirements. Nearly 29 percent of the population in low-income countries is undernourished (https://ourworldindata.org/hunger-and-undernourishment).

Safeguarding food and nutritional security of the global population requires global crop yields to be increased significantly with minimal external inputs, amid frequent extreme weather events and evolving pest and pathogens. The narrow genetic base and low yield potential of elite stocks poses a big challenge to achieve the projected rate of 2 percent gain in crop productivity (Li et al. 2018). Conventional plant breeding methods have played a significant role in developing new crop varieties harboring new traits that impart higher yield, disease and pest resistance, tolerance to abiotic stresses and nutritional quality and make them amenable to machine harvesting. However, innovative methods based on novel biotechnologies are required to broaden the genetic base and for enhancing yield potential of new cultivars. By accelerating genetic changes, genomics-assisted breeding and gene editing approaches have dramatically reduced the time required to deliver new crop cultivars from the laboratory to market (Menz et al. 2020, Varshney et al. 2021a). These novel biotechnologies, including the latest advances in genomics, DNA banking, and cryopreservation, have emerged as great tools to support methods and strategies to efficiently conserve, manage and use plant genetic resources for food and agriculture (PGRFA). This thematic background study reviews the evolution and application of emerging biotechnologies, including marker aided

selection, genomics, gene editing, DNA banking and cryopreservation to support the conservation and use of PGRFA.

2. Novel biotechnologies and their impacts on conservation of PGRFA

2.1. Application of DNA sequencing technologies and bioinformatics for species identification

Accuracy of species–level identification is vital to the conservation and utilization of plant germplasm resources. Plant systematics and taxonomy have traditionally played a key role in the identification of species in plant collections. Phenotypic dissimilarities enable the construction of phylogenetic trees in plants (Patwardhan et al. 2014). However, finding a suitable morphological or phenotypic trait for species identification is tedious and often the subtlety of these morphological dissimilarities requires taxonomic expertise, which is rapidly diminishing (Ricciardi et al. 2021). The challenge of species identification is often aggravated by the presence of overlapping morphotypes.

DNA marker systems are efficient in detecting the similarities and differences among individuals of same or different species. Methods of DNA marker discovery have been greatly benefitted by the technological advances that enable sequencing of millions of DNA fragments in a highly paralleled fashion and the concurrent increase in computational power, data handling capabilities and efficient bioinformatics and analytical tools (Bohra et al. 2020, Kanzi et al. 2020, Varshney et al. 2021a). The development of molecular markers has contributed to improving the resolution of phylogenetic understanding (Zhang et al. 2021). Nucleotide diversity assessed by comparing DNA sequence data reflects genetic divergence resulting from molecular evolution and can provide additional information on species relationships (Dong et al. 2019). As evident from the growing body of literature (Mishra et al. 2015, Jin et al. 2023), the advantage of molecular markers and

sequence data is that they can detect the variation between species that is difficult to distinguish using morphological-based procedures.

The use of DNA sequences for species identification and discovery, referred to as DNA barcoding, has gained popularity in recent years. A DNA barcode is a tool for rapid species identification based short standardized sequences of DNA (400–800 bp) (Kress 2017). For instance, Jarret (2008) already highlighted the importance of DNA barcoding-based markers to facilitate the identification of *Capsicum annuum*, *C. chinense* and *C. frutescens* that show morphological similarities. Campanaro et al. (2019) discussed the growing importance of DNA barcoding in the identification and conservation of neglected and underutilized species (NUS), such as *Amaranthus* whose discrimination is difficult because of similarities in the morphological characteristics.

Research has documented the success of a variety of plant plastid markers such as *atpF-atpH, matK, rbcL, rpoB, rpoC1, rps16, trnC-rpoB, trnH-psbA, ycf1*, *ycf5* for DNA barcoding (Guo et al. 2022). The preference of plant researchers for regions of plastids for barcoding plants stems from the extensive internal rearrangements in plant mitochondria that render the development of DNA barcodes from mitochondrial regions quite challenging as compared to animals. Plastid genomes remain highly suitable for plant barcoding owing to low evolving rate, uniparental inheritance, lack of recombination and structure stability (Mishra et al. 2017, Rogalski et al. 2015). The coding regions from plastids *rbcL* and *matK* regions have been the most preferred system for barcoding plants and are supported by The Consortium for the Barcode of Life (CBOL) (https://www.ibol.org/phase1/cbol/; Ho et al. 2021). Recently, Liu et al. (2021) highlighted the need to "augment" the standard DNA barcodes with additional sequence data for improved species identification in plants. The growing popularity and affirmed efficacy of DNA barcoding has resulted in establishing a global public resource 'Barcode of Life Data (BOLD)', an online

78 workbench and database that currently hosts about 14 million barcodes, of which 72 000 belong
79 to plants (https://www.boldsystems.org/).

80 *2.2. Application of novel biotechnologies for conservation of PGRFA in situ*

81 Growing access and affordability of high-throughput sequencing of plant genomes have allowed
82 detailed inquiries on demographic history, population structure and adaptive variations. The
83 increased scale of genomic datasets coupled with methods having greater statistical power has
84 improved the resolution of studies on genetic elements crucial to population dynamics. New
85 insights into population structure facilitated by large genomic datasets have greatly strengthened
86 conservation efforts and germplasm management strategies (Hohenlohe et al. 2021). The loss of
87 genetic diversity or inbreeding has been associated with the decline in population's ability to meet
88 the adaptive requirements sought by the changing environment. The identification of genomic loci
89 associated with climate change adaptation can greatly help prioritizing the conservation of natural
90 populations that hold immense relevance to agriculture in future climate changes scenarios.

91 Population genomic approaches can also help identify the population structure, demographic
92 events, including bottlenecks and expansions, and can guide the evaluation of conservation actions
93 and informed decisions on biodiversity management. For example, population genomic analysis
94 conducted on whole genome re-sequencing (WGRS) data obtained from 105 individuals of *Acer*
95 *yangbiense* (maple species endemic to China) provided insights into conservation genetics and
96 strategies for the further conservation of this endangered species. The comprehensive analysis in
97 the study highlighted the factors, including bottlenecks that might have contributed to the small
98 population sizes of this species (Ma et al. 2022).

Studies have highlighted the ability of genomic tools to identify adaptive traits and genomic regions controlling them, thus playing an important role in conservation and use of PGRFA. Associations between the genomic data (SNPs) and environment data in wild populations can pinpoint and identify adaptive loci (functional variations that determine the fitness of a population in a given environment) and predict phenotypic variations (Turner et al. 2010). These functional variations have a crucial role in local adaptation and should be prioritized for conservation using *in situ* or *ex situ* strategies. By studying the associations between environmental, genomic, and phenotypic data collected from 1943 geo-referenced sorghum landraces, the genomic basis of local adaptation and genotype-by-environment interactions in response to abiotic stress were characterized (Lasky et al. 2015).

2.3. Application of novel biotechnologies for enhanced management of germplasm collections

In recent years, detailed genotypic characterization of germplasm accessions by cutting-edge biotechnologies has emerged as a promising approach in germplasm management as these can allow the identification of identical accessions (duplicates) when phenotypic and passport information is not reliable (Singh et al. 2019). In this context, pairwise identity-by-state comparisons following high-density genetic profiling of germplasm collections can help the identification of near-identical samples from large collections. For example, Milner et al. (2019) generated large-scale single nucleotide polymorphism (SNP) data of collections at Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) genebank in Gatersleben, Germany (https://www.ipk-gatersleben.de/en/), encompassing cultivars, landraces and crop wild relatives by using genotyping by sequencing (GBS) technique. The high-density genetic profiling and IBS analysis led to the identification of a substantial proportion of IPK's barley collection (33 percent) representing "potential duplicates". Similarly, IBS analysis performed on a large number

of *Aegilops tauschii* accessions using genome-wide SNPs derived from diversity array technology (DArT)-Seq and GBS platforms identified nearly 50 percent duplicates in collections from the Wheat Genetics Resource Center (WGRC)-USA, the International Maize and Wheat Improvement Center (CIMMYT)-Mexico and Punjab Agricultural University (PAU), Ludhiana (Singh et al. 2019). On the basis of analysis, a total of 564 unique accessions were identified from the total 1,143 accessions used in the study.

Recent studies based on genome-wide analyses of germplasm collections have shown that genomics tools can assist in identifying and correcting the biological status of accessions in the passport records of the accessions stored in genebanks. For example, principal component analysis (PCA) of wild and domesticated barleys based on more than 100 000 SNP markers helped in rectification of the biological status of the accessions (Mascher et al. 2019). The grouping of Ethiopian accession with the domesticated barley agreed with the archaeological evidence. The analysis thus allowed identification and correction of the passport records of Ethiopian accessions from 'wild' to 'domesticated' in the information system at IPK. More recently, Varshney et al. (2021b) performed genome sequencing of 3,171 cultivated and 195 wild accessions of chickpea accessions conserved at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India. The genetic clustering based on these data clearly revealed that the accession (ICC 16369) belongs to 'wild' chickpeas, which had been labelled as 'cultivated' in the passport data record. Additionally, the presence of wild specific 'T' allele for *SHATTERPROOF2* gene in the accession ICC 16369 also corroborated it's mislabeling in biological status in genebank records.

2.4. Advances in application of in vitro conservation and cryopreservation

Plant tissue culture including embryo-rescue and other *in vitro* techniques and cryopreservation, have been applied for *ex situ* conservation of PGRFA that are not amenable or feasible for safe storage in conventional seed genebanks (FAO 2014; 2022a, b). Conservation techniques applied for orthodox seeds are not suitable for several other species, including (a) those producing either recalcitrant, intermediate or no seed; (b) vegetatively propagated crops; (c) crops producing seed but vegetatively propagated, (d) species which take many years to produce seeds, and (e) exceptional plant species (threatened and wild species). Conventionally these species are conserved in field genebanks, botanical gardens, arboreta or herbal gardens. Field collections often require large area, are difficult to maintain and protect from natural disasters, exposed to pests and pathogen, labour-intensive and, therefore, imply high maintenance cost (Panis et al. 2020).

Since the application of *in vitro* techniques for PGRFA conservation in the early 1980s, these techniques have become an important part of PGRFA management including germplasm collecting, exchange and utilization. In the last four decades, improvements have occurred in *in vitro* techniques for PGRFA conservation, *i.e.,* from short-term approaches using simple *in vitro* propagation/micropropagation and storage at standard culture room conditions to the present day cost-effective slow growth conservation approaches for short- to medium-term conservation (Agrawal et al. 2019, Panis et al. 2020). *In vitro* conservation of PGRFA under normal growth is achieved by conserving cultures under standard culture room conditions on *in vitro* multiplication/propagation medium. Typically subculture frequencies range from one to three months (Agrawal et al. 2019). This approach of normal growth is useful in species that are naturally slow growing *in vitro*. This technique has many advantages, as it does not require low temperature facility, germplasm is readily available for exchange and distribution and can elude stress-induced

variants. However, there are several drawbacks as it requires frequent subculturing, has greater risk of microbial contamination, handling error (accession mixing/mislabeling), occurrence of somaclonal variation, loss of regenerative capacity, as well as being cost and labor intensive (Agrawal et al. 2022). Major impediments include prolonged duration under an artificial tissue culture environment are (i) the occurrence of somaclonal variations, (ii) cellular ageing and senescence, (iii) appearance of slow growing/cryptic endophytic microbes which are detrimental for the plant cultures. Further, contamination of tissue culture labs by microbes or insects, handling errors of cultures, and electrical failure can be very challenging in the long-term maintenance of *in vitro* genebanks (Panis et al. 2020).

To overcome these bottlenecks, protocols have been developed for maintaining *in vitro* cultures by various 'slow-growth conservation strategies', wherein subculture period can be extended up to one to two years. Three pathways are usually adopted for achieving slow growth, either singly or in combination: (i) physical growth limitation (lowering temperature and/or light); (ii) chemical growth limitation (use of osmoticum or growth retardants); and (iii) nutrient limitation (reduced supply of carbon and inorganic nutrients), all aimed to reduce the metabolic activity of the *in vitro* tissues (Panis et al. 2020). Conservation under slow growth has several advantages over normal growth, *viz.* reduced frequency of subculture, reduction in maintenance, and labour cost.

In recent years, many cryotechniques have been refined with the aim of minimum cell damage and maximum of post-thaw viability with ease of application and without altering the genetic constitution of biological material conserved. Cryopreservation of PGRFA has been achieved using diverse group of explants or plant material, depending on the species and its post-thaw regeneration tissue (Pence et al. 2020, Wang et al. 2021). Orthodox seeds of many species are easily cryopreserved by using whole seeds and following a simple air desiccation method.

Recalcitrant seeds, intermediate seeds and vegetatively propagated crops are conserved by using zygotic embryos, embryonic axes, dormant buds, pollen, and several *in vitro* derived explants (shoot/root tips, meristems, axillary buds, nodal segments, bulbils, somatic embryos, callus, hairy roots, and embryogenic cell suspensions) (Agrawal et al. 2019).

In the last six decades of research on cryopreservation, conservation protocols have been documented for more than 200 plant species (Panis 2019). However, the range of crops and collections represented in cryogenebanks is rather limited, as cryopreservation for PGRFA conservation was only initiated towards the end of twentieth century. A study by Acker et al. (2017) estimated that the diversity of nearly 100 000 unique accessions of vegetatively propagated and recalcitrant seed crops need to be secured through long-term conservation, with only 10 percent currently cryopreserved in 15 genebanks. Only a handful of vegetatively propagated crop species represent >100 accessions conserved, including potato (5 021), cassava (2 101), bananas and plantains (1 100), garlic (925), mint (207), apple (183), strawberry (199) pear (120) and almond (134) (Ruta et al. 2020, FAO, 2023). The major factors that hamper the larger scale application of cryopreservation for long-term conservation of vegetatively propagated crops include the lack of efficient cryopreservation protocols for several species, requirement to tailor the protocols to specific materials or individual genotypes, funding constraints limiting cryobanking capacities along with lack of skilled personnel and the required infrastructure (Acker et al. 2017, Panis 2019). In spite of significant progress made in protocol development with respect to *in vitro* conservation and/or cryopreservation of 'difficult-to-conserve' species, many important crops continue to be conserved in field genebanks and face imminent threats from natural vagaries.

2.4. Advances in DNA banking and implication of Digital Sequence Information

To enhance conservation efforts, plant DNA banks have gained increasing attention as a conservation resource. DNA banks facilitate long-term storage of genetic information contained in genomic DNA. During initial genome sequencing projects, a variety of genomic resources including expressed sequence tags (ESTs), full-length cDNAs, bacterial artificial chromosomes (BACs), etc. are often stored in DNA banks and shared among users to accelerate collaborative uses (https://cropgenebank.sgrp.cgiar.org/index.php/procedures-mainmenu-243/conservation-mainmenu-198/dna-bank-mainmenu-202). The storage of DNA in plant conservation systems has been greatly benefitted by advances in extraction, purification and amplification of DNA that have dramatically reduced the quantity of the plant material required for molecular analysis. The management practices for sample preservation in DNA banks are guided by the recommendations provided by the International Society for Biological and Environmental Repositories (ISBER; https://www.isber.org/page/BPR). In 2019, the addendum, "Liquid Nitrogen-Based Cryogenic Storage of Specimens" has been added to the Fourth Edition of ISBER *Best Practices* that deals with various aspects of cryogenic storage of DNA.

DNA banks serve as the reservoirs of the genetic information which facilitates detailed inquiries into evolutionary ecological, physiological, and behavioral biology of species, thus, helping conservationists to preserve biodiversity for the present and future. For example, the DNA and Tissue bank of the Royal Botanic Garden Kew stores genetic information in the form of 60 000 samples (48 000 DNA samples and 12 000 tissue samples) from 35 000 plant species covering more than 7 000 genera represents one of the most biodiverse plant conservation systems on earth (https://www.kew.org/science/collections-and-resources/collections/dna-and-tissue-bank).

Similarly, the DNA Bank at the Missouri Botanical Garden, St. Louis, Missouri, USA aimed at

supporting population genetics and genetic relationship studies, holds a collection of more than 25,000 leaf tissue samples from diverse geographic regions, with special emphasis on endemic plants (https://www.missouribotanicalgarden.org).

The availability of DNA from DNA banks, coupled with the advanced technologies, have revolutionized the scale of data generated on plant genetic resources. These large datasets are deposited in international databases, which are further linked to other databases and publications. For example, the 228 million annotated sequences hosted at the International Nucleotide Sequence Database Collaboration (INSDC; https://www.ncbi.nlm.nih.gov/genbank/statistics/) are downloaded either partially or completely approximately 34 million times a year and used by more than 10-15 million users per year (Scholz et al., 2022). The free access and sharing of the data have been vital to research and innovation. The continuous generation of vast sequencing and other associated information on genetic resources has given rise to the placeholder term 'digital sequence information' (DSI), which is broadly defined as the genetic information accruing from genetic resources (CBD, 2023;). However, an internationally agreed definition of the DSI is currently lacking. DSI covers nucleotide sequence data (DNA/RNA) and a future-proof definition of DSI will cover other omics information (transcriptomes, proteins and metabolites) associated with genetic resources (https://www.cbd.int/doc/c/fef9/2f90/70f037ccc5da885dfb293e88/dsi-ahteg-2020-01-03-en.pdf).

3. Novel biotechnologies for accelerating the use of PGRFA

3.1. Genomic technologies for efficient use of germplasm diversity

Evolving genomic technologies coupled with computational capabilities have provided unprecedented opportunities to harness breeding potential of large germplasm collections. The application of DNA marker technologies to germplasm characterization has contributed to

enhancing the use of germplasm collections for gene discovery and trait transfer by creating workable germplasm subsets having minimum repetition and optimum genetic diversity. For example, the genetic structure of a total of 3 367 sorghum accessions from the global composite germplasm collection of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India, comprised mostly of landraces (89.5 percent), was analyzed with 41 simple sequence repeat simple sequence repeat markers and resulted in the development of a reference set of 383 genetically diverse accessions for better use of germplasm diversity for cultivar development (Billot et al. 2013). Later, a genome-wide association study (GWAS) performed on 971 diverse sorghum accessions including the reference set (Billot et al. 2013), mini core collection (Upadhyaya et al. 2009) and the association panel (Casa et al. 2008), uncovered genomic loci associated with plant height and inflorescence architecture (Morris et al. 2013). Core and mini core collections have been developed in several crop species including grain crops like rice, wheat, maize, sorghum, pearl millet, soybean, common bean, chickpea, pigeonpea and faba bean (Schafleitner et al. 2015, Fatokun et al. 2018, Egan et al. 2022, Raturi et al. 2022, Gu et al. 2023).

Due to their cost-efficiency and suitability to assay hundreds of samples, next-generation sequencing (NGS) methods based on reduced representation such as genotyping by sequencing (GBS) andspecific locus amplified fragment (SLAF) sequencing have emerged as the methods of choice for genetic characterization of large number of germplasm collections (Bai et al. 2019, Milner et al. 2019, Morris et al. 2013, Schulthess et al. 2022, Yu et al. 2016). The genetic marker data, capturing the diversity stored in genebanks, could help creation of precision collections as a community resource to support rapid breeding of target traits (Mascher et al. 2019).

Modern genomics-assisted breeding approaches including the marker-aided selection (MAS) have opened exciting avenues to rapidly incorporate the novel genetic diversity into crop breeding

programs. MAS has been particularly successful in different crops in transferring simply inherited traits controlled by strong-effect QTL, such as disease resistance (Varshney et al. 2021a). For example, several markers have been developed in wheat for improving resistance against a variety of diseases including fusarium head blight, powdery mildew, leaf rust, stripe rust and stem rust (Song et al. 2023). Some of the latest examples demonstrating the utility of MAS in crop improvement include improved varieties of chickpea released for cultivation in India (Bharadwaj et al. 2021, 2022) and Africa (https://www.cgiar.org/news-events/news/first-ever-high-yielding-chickpea-variety-developed-using-marker-assisted-backcrossing-mabc-released-in-ethiopia/) that have high level of tolerance against biotic (Fusarium wilt) and abiotic (drought) stresses.

3.2. Advances in gene editing technologies for crop improvement

The unprecedented precision of site-directed nucleases to introduce breaks and subsequent repairs at a specific site on targeted DNA has led to targeted genome engineering. Zinc finger nucleases (ZFNs, Urnov et al. 2010), transcription activator-like effector nucleases (TALENs; Joung and Sander 2013) and more recently, clustered regularly interspaced palindromic repeat (CRISPR)-Cas9 (Wang and Doudna 2023) are different techniques used for genome editing. In the last decade, the CRISPR-Cas9 system has dominated the evolving landscape of genome editing technologies. A newer class of genome editing tools referred to as base editor and primer editor has emerged in recent years that precisely edit the genetic code (Anzalone et al. 2020).

An interactive database of articles showing the introduction of any crop trait through genome editing has been developed by the European Union Sustainable Agriculture through Genome Editing (https://www.eu-sage.eu/genome-search). Among the total of 660 articles in the database, the use of CRISPR-Cas9 technology was the most abundant in the reported studies (590) followed

by TALENs (30), base editing (23), ZFN (7) and others (10). China has published the maximum number of articles on genome editing (362) followed by the United States (145), Japan (35), France (29), South Korea (28), United Kingdom (27), Germany (24), Australia (15), India (14) and remaining by other countries. The majority of traits targeted in these articles are related to plant yield and growth, improved food/feed quality, biotic stress tolerance and industry utilization. Genome editing is being applied to more than 40 crops across 25 countries, mostly addressing agronomy, food and feed quality, or abiotic stress tolerance (Menz et al. 2020). So far, only six crops developed using genome editing, namely, rice, tomato, maize, canola, soybean and camelina, have been approved for commercialization (Pixley et al. 2022). However, several other crops such as banana, cassava, potato, teff and wheat are under development.

3.3. Advances in genetic modification in crop improvement

Genetic modification (GM) or genetic engineering (GE) refers to a process of "introducing, eliminating or rearranging specific genes" in the genome of an organism by using novel biotechnologies (https://www.usda.gov/topics/biotechnology/biotechnology-glossary). The resulting crops are known as genetically modified crops. By facilitating gene transfer between sexually compatible species, GM or GE overcomes various shortcomings of the traditional breeding for creation of plants with desired traits, in terms of crossability, time and specificity.

The International Service for the Acquisition of Agri-biotech Applications (ISAAA) reported that GM crops are grown in approximately 190.4 million hectare area in 29 countries, including 24 developing and 5 industrial countries (ISAAA, 2019). Of these, the USA**,** Brazil, Argentina, Canada, India, Paraguay, China, South Africa, Pakistan and Bolivia cultivate 98 percent of the global area planted with GM crops. Globally, soybean remains the most cultivated GM crop (91.9 M hectares or 48.2 percent of the total area under GM crops), followed by maize, cotton and

canola. The other GM crops planted in 2019 included alfalfa, apple, eggplant, papaya, pineapple, potato, safflower, squash, sugarbeet and sugarcane.

The adoption of GM technology over the last 25 years has contributed to the significant increase in global production levels of crops. For instance, four major GM crops, soybean, maize, cotton and canola have witnessed a rise in their global production by approximately 330, 595, 37 and 16 million tonnes, respectively, between 1996 and 2020 (Brookes 2022). Insect resistance (bollworm/budworm pests in cotton, stalk-boring pests in maize), herbicide tolerance (tolerance to glyphosate in soybean, maize, cotton, canola) and drought tolerance (maize) have been the main traits introduced through GM.

4. Conclusions

Increased throughput and affordability of genotyping and sequencing systems have led to the generation of high-density genetic data on large collections conserved in germplasm repositories across the globe, thus facilitating their efficient management through identification of potential duplicates and correcting the biological status of accessions in the passport records. The functional diversity (e.g., adaptation to weather extremes) uncovered by the analysis of the phenotype and genotype data should inform future conservation strategies. Thus, the conservation and sustainable use of PGRFA would require better linkages between genebank passport data, phenotypic and genotypic data. At the same time, availability of high-density genotyping data on germplasm collections necessitates increased storage space and multiplication efforts to conserve these precision collections and provide material to users. Novel molecular technologies that accelerate variation analysis among different species have also contributed to improve the accuracy of species identification, which is essential for understanding the biodiversity and has tremendous implications for conservation. Molecular markers and DNA barcoding represent rapid, automated

and accurate species identification tools which can support morphological classification. Growing availability of the sequence data when combined with standard DNA barcodes will further improve species identification in plants.

In vitro techniques have been consistently used for PGRFA management and conservation, and recent technical advances have led a shift towards greater use of cost-effective and slow growth conservation strategies. In this context, cryogenebanks have emerged as one of the most suitable solutions for long-term storage of PGRFA owing to their efficacy for various tissues from a wide range of species. However, large-scale implementation of cryostorage calls for capacity building in terms of human skill and expertise, and infrastructure to facilitate research in various areas, including seed biology, *in vitro* biology and cryopreservation techniques. Storing DNA has also been a promising method for long-term conservation and of genetic material. DNA banking has been greatly benefitted by the recent advancements including techniques that enable quick, cost-effective and high throughput extraction of high-quality nucleic acid.

Genetic modification has made significant contribution to the global agriculture, and GM crops are currently grown by more than 17 million farmers in 29 countries. With the discovery of customizable nuclease systems, GE has offered an easy and cost-effective alternative to incorporate desired genetic manipulations in the target sequence. Importantly, GE-products are subject to less scrutiny from time-consuming and costly regulatory processes. Concerning laboratory to market transition of the GE products, discussions among policy makers, experts, the public, and NGOs are critical to develop sustainable and evidence-informed policy. Greater engagement of stakeholders is required to provide assessment of the risks and challenges associated with novel biotechnologies for the benefits of consumers, farmers and the environment.

Developing the required skills or capabilities among researchers and genebank managers from developing nations through capacity building is imperative to harness the full potential of novel biotechnologies and new breeding techniques for PGRFA conservation and use. Both technical and functional capacitie of researchers will be necessary to take advantage of these fast-evolving fields of plant phenotyping and biotechnologies including genome editing and next-generation genomics.

References

Acker JP, Adkins S, Alves A, Horna D, Toll J (2017) Feasibility Study for a Safety Backup Cryopreservation Facility; Bioversity International: Maccarese-Stazione, Italy.

Agrawal A, Gowthami R, Chander S, Srivastava V (2022) Sustainability of In Vitro Genebanks and Cryogenebanks. Indian J Plant Genet Resour 35: 180-184.

Agrawal A, Singh S, Malhotra EV et al. (2019) In vitro conservation and cryopreservation of clonally propagated horticultural species. In: Rajasekharan PE, Rao VR (eds) Conservation and Utilization of Horticultural Genetic Resources. Springer Singapore, Singapore, p 529-578.

Anzalone AV, Koblan LW, Liu DR (2020) Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. Nat Biotechnol 38: 824–844

Bai Q, Cai Y, He B, Liu W, Pan Q, Zhang Q (2019) Core set construction and association analysis of Pinus massoniana from Guangdong province in southern China using SLAF-seq. Sci Rep 9:13157.

Bharadwaj C. Tripathi S. Soren KR et al. (2021). Introgression of "QTL-hotspot" region enhances drought tolerance and grain yield in three elite chickpea cultivars. Plant Genome 14: e20076.

Bharadwaj C, Jorben J, Rao A et al. (2022) Development of high yielding Fusarium wilt resistant cultivar by pyramiding of "Genes" through marker-assisted backcrossing in chickpea (Cicer arietinum L.). Front Genet 13:924287.

Billot C, Ramu P, Bouchet S et al. (2013) Massive sorghum collection genotyped with SSR markers to enhance use of global genetic resources. PLoS One 8:e59714.

Bohra A, Jha UC, Godwin I, Varshney RK (2020) Genomic interventions for sustainable agriculture. Plant Biotechnol J 18: 2388-2405

Brookes G (2022) Farm income and production impacts from the use of genetically modified (GM) crop technology 1996-2020. GM Crops Food 13:171-195.

Campanaro A, Tommasi N, Guzzetti L, Galimberti A, Bruni I, Labra M (2019) DNA barcoding to promote social awareness and identity of neglected, underutilized plant species having valuable nutritional properties. Food Res Int 115:1-9.

Casa AM, Pressoir G, Brown PJ et al. (2008) Community resources and strategies for association mapping in sorghum. Crop Sci 48:30–40

Dong Y, Chen S, Cheng S, et al. (2019) Natural selection and repeated patterns of molecular evolution following allopatric divergence. Elife 8:e45199.

Egan LM, Conaty WC, Stiller WN (2022) Core Collections: Is There Any Value for Cotton Breeding? Front Plant Sci 13:895155.

FAO (2014) Genebank Standards for Plant Genetic Resources for Food and Agriculture; Food and Agriculture Organization of the United Nations: Rome, Italy.

FAO (2022a) Recent developments in biotechnologies relevant to the characterization, sustainable use and conservation of genetic resources for food and agriculture. FAO Commission on Genetic Resources for Food and Agriculture. Rome. https://doi.org/10.4060/cb8956en

FAO (2022b) Practical guide for the application of the Genebank Standards for Plant Genetic Resources for Food and Agriculture: Conservation via in vitro culture. Commission on Genetic Resources for Food and Agriculture. Rome. https://doi.org/10.4060/cc0025en

FAO (2023). WIEWS—World Information and Early Warning System on Plant Genetic Resources for Food and Agriculture; Food and Agriculture Organization of the United Nations: Rome, Italy. https://www.fao.org/wiews/en/

Fatokun C, Girma G, Abberton M et al. (2018) Genetic diversity and population structure of a mini-core subset from the world cowpea (Vigna unguiculata (L.) Walp.) germplasm collection. Sci Rep 8: 16035.

Guo M, Yuan C, Tao L, Cai Y, Zhang W (2022) Life barcoded by DNA barcodes. Conserv Genet Resour 15:1-15.

Gu R, Fan S, Wei S, Li J, Zheng S, Liu G (2023) Developments on core collections of plant genetic resources: Do we know enough? Forests 14:926.

Hohenlohe PA, Funk WC, Rajora OP (2021) Population genomics for wildlife conservation and management. Mol Ecol 30:62-82.

Ho VT, Tran TKP, Vu TTT, Widiarsih S (2021) Comparison of matK and rbcL DNA barcodes for genetic classification of jewel orchid accessions in Vietnam. J Genet Eng Biotechnol 19:93.

ISAAA (2019) Biotech crop highlights. Available at: https://www.isaaa.org/resources/publications/pocketk/16/#:~:text=The%20most%20planted%20biotech%20crops,crops%20or%2091.9%20million%20hectares (Accessed 28 September 2023).

Jin DP, Sim S, Park JW et al. (2023) Identification of the plant family caryophyllaceae in Korea using DNA barcoding. Plants 12:2060.

Jarret RL (2008) DNA barcoding in a crop genebank: Resolving the Capsicum annuum species complex. Open Biol 1:35-42

Joung JK, Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol 14:49-55.

Kanzi AM, San JE, Chimukangara B et al. (2020) Next generation sequencing and bioinformatics analysis of family genetic inheritance. Front Genet. 11:544162.

Kress WJ (2017) Plant DNA barcodes: Applications today and in the future. J Syst Evol 55: 291–307

Lasky JR, Upadhyaya HD, Ramu P et al. (2015) Genome-environment associations in sorghum landraces predict adaptive traits. Science Adv 1:e1400218.

Li H, Rasheed A, Hickey LT, He Z (2018) Fast-forwarding genetic gain. Trends Plant Sci 23:184-186.

Liu ZF, Ma H, Ci XQ et al. (2021) Can plastid genome sequencing be used for species identification in Lauraceae? Bot J Linn Soc 197: 1– 14.

Ma Y, Liu D, Wariss HM et al. (2022) Demographic history and identification of threats revealed by population genomic analysis provide insights into conservation for an endangered maple. Mol Ecol 31:767-79.

Mascher M, Schreiber M, Scholz U, Graner A, Reif JC, Stein N (2019) Genebank genomics bridges the gap between the conservation of crop diversity and plant breeding. Nat Genet 51:1076-1081.

Menz J, Modrzejewski D, Hartung F, Wilhelm R, Sprink T (2020) Genome edited crops touch the market: A view on the global development and regulatory environment. Front Plant Sci 11:586027.

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

Mishra P, Kumar A, Nagireddy A, Shukla AK, Sundaresan V (2017) Evaluation of single and multilocus DNA barcodes towards species delineation in complex tree genus Terminalia. PLoS One 12: e0182836.

Mishra P, Kumar A, Nagireddy A et al. (2015) DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. Plant Biotechnology J 14:8–12.

Morris GP, Ramu P, Deshpande SP et al. (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc Natl Acad Sci U S A 110:453-8.

Panis B (2019) Sixty years of plant cryopreservation: From freezing hardy mulberry twigs to establishing reference crop collections for future generations. Acta Hortic 1234: 1–7.

Panis B, Nagel M, den Houwe IV (2020) Challenges and prospects for the conservation of crop genetic resources in field genebanks, in in vitro collections and/or in liquid nitrogen. Plants 9: 1634.

Patwardhan A, Ray S, RoyA (2014) Molecular Markers in Phylogenetic Studies – A Review. J Phylogen Evolution Biol 2: 131.

Pence VC, Ballesteros D, Walters C et al. (2020) Cryobiotechnologies: Tools for expanding long-term ex situ conservation to all plant species. Biol Conserv 250:108736.

Pixley KV, Falck-Zepeda JB, Paarlberg RL et al. (2022) Genome-edited crops for improved food security of smallholder farmers. Nat Genet 54:364-367.

Raturi D, Chaudhary M, Bhat V et al. (2022) Overview of developed core and mini core collections and their effective utilization in cultivated rice and its related species (Oryza sp.)-a review. Plant Breed 141: 501–512.

Ricciardi A, Iacarella JC, Aldridge DC et al. (2021) Four priority areas to advance invasion science in the face of rapid environmental change. Environ Rev 29: 119-141.

Rogalski M, do Nascimento Vieira L, Fraga HP, Guerra MP (2015) Plastid genomics in horticultural species: importance and applications for plant population genetics, evolution, and biotechnology. Front Plant Sci 6:586.

Ruta C, Lambardi M, Ozudogru EA (2020) Biobanking of vegetable genetic resources by in vitro conservation and cryopreservation. Biodivers Conserv 29:3495-3532.

Schafleitner R, Nair RM, Rathore A et al. (2015) The AVRDC – The World Vegetable Center mungbean (Vigna radiata) core and mini core collections. BMC Genomics 16:344.

Scholz AH, Freitag J, Lyal CHC et al. (2022) Multilateral benefit-sharing from digital sequence information will support both science and biodiversity conservation. Nat Commun 13:1086.

Schulthess AW, Kale SM, Liu F et al. (2022) Genomics-informed prebreeding unlocks the diversity in genebanks for wheat improvement. Nat Genet 54:1544-1552.

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

Song L, Wang R, Yang X, Zhang A, Liu D (2023) Molecular markers and their applications in marker-assisted selection (MAS) in bread wheat (Triticum aestivum L.). Agriculture 13:642.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of Arabidopsis lyrata to serpentine soils. Nat Genet 42: 260–263

Upadhyaya HD, Pundir RPS, Dwivedi SL et al (2009) Developing a mini core collection of sorghum for diversified utilization of germplasm. Crop Sci 49:1769–1780.

Urnov F, Rebar E, Holmes M et al. (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11: 636–646.

van Dijk M, Morley T, Rau ML et al. (2021) A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. Nat Food 2: 494–501.

Varshney RK, Bohra A, Yu J, Graner A, Zhang Q, Sorrells ME (2021a) Designing future crops: genomics-assisted breeding comes of age. Trends Plant Sci 26: 631-649

Varshney RK, Roorkiwal M, Sun S et al. (2021b) A chickpea genetic variation map based on the sequencing of 3,366 genomes. Nature 599 :622-627Wang JY, Doudna JA (2023) CRISPR technology: A decade of genome editing is only the beginning. Science 379:6629.

Wang MR, Lambardi M, Engelmann F, Pathirana R, Panis B, Volk GM et al. (2021) Advances in cryopreservation of in vitro-derived propagules: technologies and explant sources. Plant Cell, Tissue and Organ Cult 144:7-20.

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

Zhang T, Huang S, Song S et al. (2021) Identification of evolutionary relationships and DNA markers in the medicinally important genus Fritillaria based on chloroplast genomics. PeerJ 9:e12612.

Annex. Glossary of Terms

CRISPR-Cas9: A gene editing system involving an endonuclease that cleaves a double-stranded DNA at a specific site as directed by the guide RNA.

Cryopreservation: Long-term maintenance of cells, tissues, and organs at sub-freezing temperatures.

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.

DNA banking: Storage of an individual's genetic material for future analysis.

Genome editing: Alteration in the native genetic material (DNA) through deletions, insertions, substitutions or point mutations.

Genotyping by sequencing (GBS): A high-throughput method for identification of single nucleotide polymorphisms (SNPs) using restriction enzyme-based reduced representation DNA sequencing of multiple samples.

Identity-by-state: Presence of same allele at a specific locus even if the individuals are not related to each other.

Neglected and underutilized species (NUS): Wild, semi-domesticated and cultivated species, which were used as traditional/local crops but are now neglected due to limited commercial interest.

New breeding techniques: Precise and faster breeding methods to incorporate targeted changes in the genetic constitution to develop new improved varieties with desired traits.

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

Restriction-site associated DNA sequencing: Use of multiple restriction enzymes to cleave DNA, followed by DNA sequencing to obtain targeted representation of the genome.

Simple sequence repeat (SSR): Stretches of DNA made up of short tandem repeats of nucleotide motifs ranging in length from 1-6 base pairs.

Site-directed nucleases: Proteins or enzymes used to target the dsDNA repair for modification via three approaches:

Single nucleotide polymorphism (SNP): DNA sequence variation at a single position among individuals.

Specific locus amplified fragment sequencing: A method of optimized reduced representation sequencing with the use of training data for large-scale genotyping of multiple individuals.

Transcription activator-like effector nucleases (TALENs): Engineered restriction enzymes to cleave DNA at specific points.

Whole genome re-sequencing (WGRS): Comprehensive next-generation sequencing of entire genomes of individuals and comparing them with that of a known reference genome.

Zinc finger nucleases (ZFNs): A chain of zinc finger proteins linked to a bacterial nuclease for site-specific dsDNA breaks.