Report of Workshop held Jun 2023 at Cornell University

PLANT GENETIC RESOURCE DIGITAL COMMONS:

conservation to consumption

Mar 2024

A Tactical Roadmap for investment to increase knowledge exchange and harness Crop Plant Genetic Resources curated in genebanks for the benefit of humanity

DIVSEEK INTERNATIONAL NETWORK INC.

Plant Genetic Resource Digital Commons: *conservation to consumption*

A Tactical Roadmap for investment to increase knowledge exchange and harness Crop Plant Genetic Resources curated in genebanks for the benefit of humanity.

EXECUTIVE SUMMARY

What? PGR collections contain an intrinsic capacity to meet global and regional challenges of climate change, environmental degradation, food supply and human health. **This document presents an analysis of the current state and the future potential of information flow around crop Plant Genetic Resource (PGR) collections conserved in genebanks**.

We identify current barriers that impede greater access and utility of these genetic resources for breeders, farmers, researchers, and the wider bio-economy. Proposed actions are framed in the context of historical and future benefits for food, nutritional and national security, with significant health and environmental outcomes. This represents the crop PGR value chain.

We present the value proposition and a theory of change for democratizing access to information about PGR, facilitated by a multi-domain community of practice. Our analysis builds on discussions within the DivSeek International Network over the past few years, and a focused workshop in 2023 co-organized with the Global Crop Diversity Trust.

Current successes and ongoing challenges are identified, together with working practices required to achieve a more functional and interoperable information flow connecting PGRs with R&D, plant breeding, the wider agri-food and bioeconomy sectors. The analysis identifies barriers to communication and the need to democratize access to information derived from PGR in parallel to their access and use.

We address the key role for implementation of data standards, and present a vision for increased investment aligned to practical next steps. This includes actions focused on extending and establishing open interoperable information systems with a broader, easily accessible, and robust portfolio of data description standards, alongside increasing awareness of the impact of PGRs and interactions between key stakeholder groups.

Why now? Intervention in adding value to PGR information builds on rapid advances in genomics, genome editing, information technologies, and artificial intelligence (AI) that have had transformational effects on sectors such as clinical health, engineering, and advanced materials science. PGR collections contain an intrinsic capacity able to contribute to meeting global and regional challenges of climate change, environmental degradation, food supply and human health.

Who cares? Crop diversity, particularly that **contained within genebanks,** can function as a driver of plant breeding advancement and is a critical component of **global food security**, **climate adaptation and diversification of the bioeconomy**. **Information flow** between different domains is key to realizing the potential of PGR. We highlight the economic, social and environmental benefits expected arising from coordinated development of robust data standards and interoperable working practices. As outlined, investment in a crop PGR *Digital Commons* will leverage the contribution of PGR biodiversity to meet targeted regional and global challenges.

Vision:

We present a vision and tactical roadmap to address major challenges through collective effort in a series of well-defined coordinated activities. There is a generic need to link PGR with high quality characterization and evaluation[1] data. We see value in such data being generated by different stakeholders and annotated using generic vocabularies and ontologies[2] to describe crop phenotypes, traits and downstream attributes. We argue for timely and targeted investment for infrastructure development, and specifically where this supports information flow that adds value to *ex situ* PGR collections through open sharing of characterization information.

The proposed actions require concerted global engagement of public and private sectors, with the expansion of an active identifiable community of practice built on common vested interests in conserving and harnessing plant biodiversity.

This document has been prepared by a community of practice led by DivSeek International, as guardians and mediators of an open content process that includes workshops, publications, open-data resources and wider consultation. Readers may also wish to consult the DivSeek International Network Strategic Plan[3], including the Glossary of Terms (p14-15).

[1] PGR *characterization* involves i) a description of the distinctive nature or features of an organism; ii) detection of variation as a result of differences in either DNA sequences or specific genes (genotype), or any aspect of morphology, physiology, biochemistry, interaction with biotic or abiotic environment, appearance or quality traits (phenotype). *Evaluation* includes assessment of germplasm for any agronomic (production) or end-use quality traits.

[2] An ontology is a structured set of concepts and categories in a subject area or domain that shows their properties and the relations between them.

[3] https://divseekintl.org/strategic-plan/

CONTENTS

1. The Value Proposition

Plants underpin earth systems[4] and human civilization. **Biodiversity** within plant gene pools represents an irreplaceable source of variation to meet challenges of climate adaptation and mitigation, food security and human health.

The crop value chain globally spans the PGR, R&D, breeding, production, post-harvest, processing, supply-chain, consumption sectors with associated socio-economic and health outcomes. In this context the capacity for exchange and annotating complex datasets related to PGR characterization is limite

Ex situ genetic resource collections (genebanks) capture, characterize, conserve, document, and distribute representative variation within defined taxa or ecosystems and may conserve these in the form of seeds or live plant collections. Seeds are plant time capsules for some of the most complex self-replicating adaptive systems in the universe.

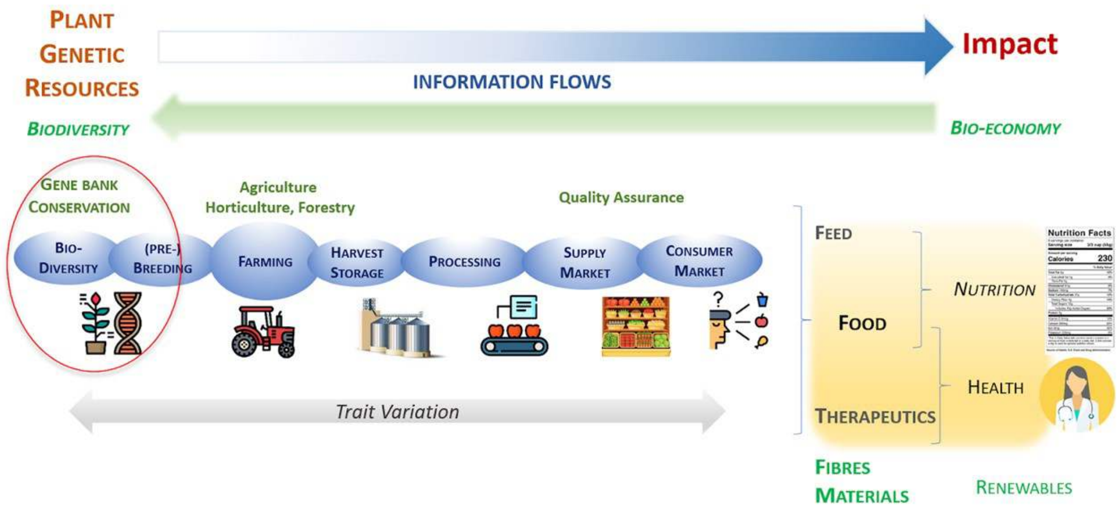
Securing and characterizing **germplasm** representing genetic variation provides a powerful platform for ensuring access to combinations of genes that have been subject to natural and human selection through millenia. Here our primary concern is with collections that aim to conserve within-species variation for **crops** and crop wild relatives. These are generally distinct from efforts aimed to conserve species' diversity such as the Millennium Seed Bank at Kew[5]**.**

Plant breeding involves the ongoing generation, selection and deployment of new crop cultivars adapted to Ag production and supply systems. Plant breeders depend on access to *ex situ* plant genetic resources to gain access to novel and unique diversity alongside reliable and high quality data. However, current infrastructure and information sharing mechanisms are limited in their ability to link collection inventories with R&D, breeding, and bio-economy assessment of useful variation (evaluation data).

[4] Coates, Juliet C., et al. "Plants and the Earth System - Past Events and Future Challenges." The New Phytologist, vol. 189, no. 2, 2011, pp. 370–73. JSTOR, http://www.jstor.org/stable/40983839. Accessed 7 Nov. 2023.
[5] http://tinyurl.com/jvwp76k

Increasing demands and challenges of changeable growing environments, market diversification and functional end-uses require **access to genetic variation** alongside characterization and **evaluation data**.

The under-valuation of PGR variation in modern socio-economic systems is an existential risk. Political, economic, social and technological factors highlight the need for public-good and private sector investment in enhanced **information** and access infrastructure to harness the potential of PGR.



Governance and management of Plant Genetic Resource (PGR) collections

Ex situ PGRs are primarily visible as represented by a system of national and international crop-specific genebanks. Public PGR collections are organized under national institutions, such as the USDA National Plant Germplasm System[6], CGIAR[7] and FAO Biodiversity International[8]. In addition, the Svalbard Global Seed Vault[9] is managed under the auspices of the Global Crop Diversity Trust.

Collectively, these are responsible for curation, regeneration and implementing conservation policies, typically managed within public sector repositories that interact with public and private R&D institutions, public sector breeders (some economies), private sector breeders and seed companies (from multinationals such as Bayer to sole traders). They thus support both large-scale and specialist Ag/horticulture.

- A far wider range of research organizations, breeders and farmers manage PGR collections around the world. These often support regionally relevant vertically integrated grower/producer collectives, added value enterprises and regional industries.

[6] http://www.ars-grin.gov/
[7] http://isa.ciat.cgiar.org/urg/main.do?language=en
[8] http://www.bioversityinternational.org/
[9] http://www.nordgen.org/sgsv/

- In parallel many commercial private breeders increasingly manage their own diversity collections independently from the global PGR system, due in part to uncertainties and inconsistencies around consequences of international treaties and protocols (see Section 2).

Access, exchange and downstream use of PGR are regulated by **international agreements, including** the Convention on Biological Diversity (CBD) and principles Access and Benefit Sharing (ABS), the International Treaty on Plant Genetic Resources for Food and Agriculture (FAO ITPGRFA), the FAO Global Plan of Action for Plant Genetic Resources, and the Nagoya Protocol[10].

- These have resulted in ongoing confusion, with the lack of a workable ABS framework in particular often limiting freedom to operate.

- Many commercial private breeders increasingly manage their own diversity collections independently from the global PGR system, due in part to uncertainties and inconsistencies around consequences of international treaties and protocols.

A recent paper outlines a context and guide to best practices for PGR[11].

ANALYSIS

- Plant Genetic Biodiversity is defined variously at different levels including genomic haplotypes, genes, genotypes, populations and species. Rarely are these measures of diversity integrated.

- Genetic Resource (germplasm) collections provide the platform for mining and harnessing novel genetic and chemical variation. Prior studies[12],[13],[14] have demonstrated remarkably high relative returns on investment derived from genetic improvement of plants and livestock when compared to other activities.
    - It is recognized that PGR collections only remain useful if resources are conserved long-term and benefit from ongoing review, collection and assessment of coverage.

- Whilst it is apparent that both *in situ/ex situ* are important for retaining material and preventing genetic erosion, the full socio-economic value or potential in PGR collections is unknown. This knowledge gap is due to a lack of quantitative tools, as well as rapid changes in environmental, end-user, and market factors, including global challenges such as working to net zero and addressing nutritional security.

- The impact of PGR conservation, along with characterization and utilization data has been assessed in a number of Crop Trust case studies, including:

[10] https://doi.org/10.1073/pnas.2205773119

[11] https://doi.org/10.1073/pnas.2205773119

[12] McCouch et al. (2013) Agriculture: Feeding the future. Nature.DOI: 10.1038/499023a

[13] https://doi.org/10.1080/21513732.2011.593557;

[14] Gollin (2020) Conserving genetic resources for agriculture: economic implications . Food Security 12, 919–927; Brush (1996) Valuing Crop Genetic Resources. https://doi.org/10.1177/107049659600500403

- the Global Public Goods 2 project[15], Genebank platform[16], Crop Wild Relatives initiative[17], and the ongoing BOLD[18] programme.

The role of information systems in the realm of PGR

Conservation:

The crop PGR conservation sector is defined primarily by an international community of public sector '*ex situ'* genebanks (seed vaults, plantations etc). Beyond those responsible for managing physical plant resources, a complex but often fragmentary ecosystem of information systems manage the digital footprint and characterization/evaluation of PGR.

Working towards greater integration and interoperability of information systems is expected a) to help the network of public genebanks to work as a single, virtual genebank following agreed standards of characterization and data sharing; b) help researchers, breeders and other end-use stakeholders to maximize the benefits of accessing and sharing information about PGR.

A loosely connected data ecosystem has been promoted through combined actions of the FAO/Treaty GLIS[19] and the public sector R&D community across the world [see diagram], with majority of investment from G7/G20 nations. This is primarily focused on germplasm inventory, and targeted characterization and evaluation programs. Although there has been good progress in international training and exchange of expertise, ongoing digital inequalities[20] represent a major barrier to the conservation and utilization of PGR in much of the world.

Formal inventories are managed under the auspices of the FAO/ITPGRFA and implemented by the Crop Trust (the Global Crop Diversity Trust), with > 4.3 million germplasm accession records listed in the Genesys-PGR database[21]. Beyond this, genetic resources are managed by public sector and regional farmer cooperatives, as well as in academic and NGO research collections.

---

[15] https://www.genebanks.org/resources/publications/collective-action-for-the-rehabilitation-of-global-public-goods-in-the-cgiar-genetic-resources-system-phase-2-gpg2/
[16] https://www.cgiar.org/the-genebank-platform/
[17] https://cwr.croptrust.org/
[18] https://bold.croptrust.org/
[19] https://glis.fao.org/glis/
[20] doi: 10.1080/01436597.2022.2079489
[21] https://www.genesys-pgr.org/

Additional tools to facilitate access and analysis of PGR information include GRIN-Global[22], Eurisco[23], and other national systems alongside specific tools for analysis of germplasm variation, including exploration of genomic data and associated derived genotype and phenotype data.

No coordinated or comprehensive system or portal is available for users to explore the landscape and impact of PGR data, or to link this with downstream data and information processing resources. The *DivSeek Commons* 'matrix'[24] is a first step in this direction.

Utilization:
Maximizing the value derived from available PGR is dependent on access to appropriate descriptive data. This may include formal characterization to distinguish and systematize the relationships between accessions (including DNA genotyping[25]), as well as pre-emptive evaluation of phenotypic variation (biological, chemical, physical or other attributes).

---

[22] https://www.grin-global.org/
[23] https://www.ecpgr.cgiar.org/resources/germplasm-databases/eurisco-catalogue/
[24] https://divseekintl.org/commons-landscape-matrix/; https://divseekintl.org/divseek-commons-the-matrix-is-born/
[25] https://excellenceinbreeding.org/module3

The DivSeek International Network

- DivSeek was established in 2012 as an informal global community of practice for sharing information about plant genetic resources.

- A White Paper '*Harnessing the power of crop diversity to feed the future*' published[26] in 2014 outlined how assessment and identification of new sources of genetic variation is a critical part of any long term strategy to enhance the productivity, sustainability and resilience of crop varieties and agricultural systems. This highlighted the opportunity to bring together genebanks, breeders, plant and crop scientists, database and computational experts to enhance the use of genebank materials, promote effective utilization of genetic variation in plant improvement.
  - Since 2014, there has been considerable change, including rapid advances in informatics, low cost whole genomes, global warming, and a 706m (9.6%) increase in population.

- DivSeek International Network Inc. was formally established as a not-for-profit with a Board elected by the ¬70 institutional members, with co-funding from Genome Prairie (Genome Canada), and the Global Institute for Food Security. This phase ran from 2018-2023 and was successful in:
  - articulating an overall vision and strategic requirements in the *DivSeek Strategic Plan*[27] (2021-2026);
  - establishing an initial set of globally regional and thematic Hubs[28];
  - agreeing a Memorandum of Understanding with the FAO International Treaty on Plant Genetic Resources for Food & Agriculture.
  - proposing the *DivSeek Commons*[29] concept as an ecosystem of Standards, Open Data Resources, Analysis Tools and Best Practices' for Plant Genetic Resources.

CONSIDERATIONS

Promoting sustainable use of PGR for key strategic end-uses and development of the bioeconomy can only be achieved where information is available to guide conservation and ensure that coverage of ongoing collections meets future needs.

- The tangible connections between health outcomes and PGR variation are not widely considered in policy decision-making.

- Improved interoperable PGR information systems are expected to facilitate de-risking of global and regional system failures to:
  - Contribute to AgroBiodiversity and land productivity
  - Accelerate transition of food systems to climate change mitigation (climate smart targets)
  - Improve food security through locally adapted crops, including
  - Health/nutrition enhanced targets

---

[26] https://divseekintl.org/wp-content/uploads/2019/12/WhitePaperDivSeek.pdf
[27] https://divseekintl.org/strategic-plan/
[28] https://divseekintl.org/regional-hubs/
[29] https://divseekintl.org/commons/

    - Find ideal genotype**s** with good taste and good resistance (ideotype)
    - Diversify future products (food, bio-renewable, therapeutic)
    - Address emergent unknown targets

- Market supply economics and biological diversity have resulted in a mismatch in investment between the major commodity starch, oil and protein crops *vs* the extensive and the diverse nutrient-dense vegetable, fruit and other culinary crops that contribute to human health and wellbeing worldwide.

- Alongside widely grown horticultural crops, investment is lacking in strategically important and regional traditional crops that are often regarded as minor or underutilized[30].

The following sections outline specific barriers to communication, challenges, gaps and proposed actions for democratizing access to information about PGR, and outline the role of standards in facilitating information flow throughout the crop PGR value chain. A vision is presented, along with a prioritized set of actions requiring coordinated effort and where return on investment will be tangible.

[30] https://doi.org/10.1007/s00425-019-03179-2

2. Democratizing Access to PGR Information through the Crop Value Chain

The crop value chain globally spans the PGR, R&D, breeding, production, post-harvest, processing, supply-chain, consumption sectors with associated socio-economic and health outcomes. In this context the capacity for exchange and annotating complex datasets related to PGR characterization is limited.

ANALYSIS

Political, Legal, Socio-economic

The AgFood and wider bioeconomy depend on reliable supplies of cultivated plants adapted to regional environments and markets.

- It is recognized that a tension exists between local and global food systems.
    - More inclusive access to PGR and related information is required to meet the needs of an increasing 21st century population, environmental challenges for Ag production, working to net zero, and achieving nutritional goals. Overall, PGR hold the key to developing locally adapted crops that also meet a more diversified bioeconomy.
    - Although there has been good progress in international training and exchange of expertise, ongoing digital inequalities[31] represent a major barrier to the conservation and utilization of PGR in much of the world.

- International and national regulatory frameworks, including CBD, International Treaty, and Nagoya[32] have resulted in ongoing confusion[33], often limiting freedom to operate.
    - The CBD and Nagoya concept of Access and Benefit Sharing (ABS)[34] has yet to be resolved or universally implemented for PGR, affecting data sharing, R&D and utilization.
    - The lack of an ABS framework limits commercial breeders FTO (freedom to operate).

- Historically, access and systematic characterization of PGR has led to transformational outcomes. eg - from 1960s onwards with 'Green Revolution' dwarfing wheat, rice etc; canola introduced as new oil crop in 1970s, increasing use of natural disease and pest resistances)[35]
    - Use cases demonstrate the strategic value of conservation and characterization[36].

- Major crops have been successful in utilizing the global depth of genetic resources - for annual arable crops the materials required are covered well.
    - However, minor and underutilized crops are mostly too small to generate cohesive communities of practice (CoP), with poor access to relevant education, information management tools, and limited human and physical resources.
    - A gap in R&D CoPs also exists for perennial crops, and for regional crop diseases.

- A tension between conservation and utilization has often inhibited development of sustainable management of PGR, with large imbalances in investment for conservation and downstream use.

---

[31] doi: 10.1080/01436597.2022.2079489
[32] https://doi.org/10.1073/pnas.2205773119
[33] https://doi.org/10.1073/pnas.2205773119
[34] https://www.cbd.int/abs/
[35] https://library.oapen.org/bitstream/handle/20.500.12657/61485/9781786768827_web.pdf?sequence=1&isAllowed=y
[36] https://www.genebanks.org/resources/impacts/working-papers/

- PGR management and some breeding may not be directly linked to or aware of the wider value chain/network. The full chain of interactions and interfaces (human, live plant, machine) may not be clear to individual players.

- Food supply beyond primary production spans multiple processing and supply chain sectors.
  - Opportunity to build on and existing limited consumers' awareness and concern about plant varieties, where attributes may be associated with quality, provenance and regional identity, as well as nutrition and health outcomes - eg, consumers are actively engaged in cultivar variation data for sectors such as fresh fruit and wine[37].
  - Demand may increase to diversify available crops and varieties of vegetables and grains, including chickpea, lentils, wheat.

- Specification frameworks within the fiber supply and bioproducts industry remain patchy and poorly connected with PGR systems.

Technical

- Documentation about PGR accessions is often patchy or lacks visibility about the history of conserved material, limiting traceability and attribution.
  - The extent of interoperability and adoption of controlled vocabularies in the PGR, breeding and R&D sector has been very limited[38].
  - This is particularly apparent when associated PGR-derived materials with DNA genotyping, genomic and high throughput phenotyping data.

- Requirement for Information systems that add value and are easily accessible (FAIR[39]) + traceable

- An ongoing requirement for translation of concepts to different languages. There is a requirement for consistent vocabularies and integration of other data-sharing standards (see Section 4).

OTHER CONSIDERATIONS AND COUNTER-ARGUMENTS

- Demonstrating the value added to PGR from data interoperability, alongside formal and informal good management and governance practices, is likely to stimulate wider adoption of best-practices.

- Many institutional breeding programs feel they have sufficient diversity for years to come, with breeding companies often controlling the message about varieties based on their available genetic diversity.

- Provision of data may lead to unexpected, innovative uses.

PROPOSED APPROACHES and ACTIONS

[37] https://sca.coffee/research/coffee-tasters-flavor-wheel
[38] https://doi.org/10.1093/database/baab028
[39] The FAIR data principles for scientific data management and stewardship aim to optimize the re-use of data by ensuring it is Findable, Accessible, Interoperable and Reusable

- Increase policy awareness of the role of PGR in health (nutritional security), defense/food security, climate change mitigation.
  - Stimulate economic analyses to quantify the research value chain arising from PGR.
  - Develop narratives to engage the wider bio-economy in understanding the value of cultivar and PGR germplasm variation held within PGR collections.
  - Use demonstration projects to encourage adoption of robust and consistent specifications through the crop value chain.
  - Recognize that access and sharing is dependent upon availability of traceable data and information about PGR variation.

- More generally, build on consumer-oriented narratives, making it easier to access information about plant varietal variation that is relevant in the market-place.

- Develop scalable sets of best practices and operating procedures, particularly at the interfaces between domains (PGR to R&D, R&D to breeder, breeder to farmer, farmer to market) that recognize the value of PGR variation.
  - Extend available examples of traceability of germplasm from source to consumer.

- Within the PGR conservation domain - define minimum data quality standards for decision-making (breeding) – this will resolve uncertainty about legacy and incomplete data. Consider dataset triage methods based on Uniqueness, Completeness and Information Content (UCIC).

- In the R&D and breeder sector foster and incentivize a culture of data deposition by demonstrating cumulative benefits; ensure data are credited and tools are available to track and deposit data.
  - Address capacity of data generators to submit
  - Provide quantifiable reward for effort (eg returnable)

- Extend strategies as adopted in the CGIAR Generation Challenge Program (GCP)[40] - approaches include socializing PGR and breeding protocols, greater emphasis on visualization of data, public involvement in curation.

- An opportunity exists to build on the *DivSeek Commons*[41] and platforms such as an information ecosystem matrix to help players in different domains navigate specific technical tools.
  - Recognize that some redefinition of relationships between the PGR conservation, R&D and Ag sectors is required to demonstrate the value to be derived from an inclusive PGR-centric *Data Commons*, especially to quantify collective benefits and mutual dependence.

[40] https://www.generationcp.org/; doi: 10.1007/s11032-010-9512-3; a global crop research consortium directed toward crop improvement, from comparative biology, PGR characterization to plant breeding.
[41] https://divseekintl.org/commons/

- In the context of Access & Benefit Sharing, CARE principles may become more prominent through adoption of TK/BC labels[42] for recognizing traditional knowledge and biocultural contributions to plant germplasm.

[42] https://divseekintl.org/tk_bc_labels/

3. Addressing Barriers to Communication

The complex challenges of integrating data from genebanks, breeding programs and the wider bioeconomy limit harnessing the full socio-economic potential of crop PGR. Barriers to communication and information flow between different domains include technological, social, economic and political factors.

On a technical level, there are disparate data formats and varying levels of data quality, placed alongside the divergent goals and methodologies of conservation (genebanks) and downstream user (R&D, breeders, bioeconomy) communities. Additional interconnected sociological, legal and regulatory factors contribute to under-investment in resourcing interoperable platforms to facilitate information flow and access to PGR.

ANALYSIS

The PGR and immediate stakeholder communities have identified a series of barriers that currently limit the potential of accessible and well-characterized plant genetic diversity to address environmental, health, societal and economic challenges. These barriers include:

Technical

Data Compatibility and Standardization: Genebanks and breeding programs operate at different levels of abstraction and complexity in terms of underlying genetics and use different data formats, platforms, and terminologies to store and manage their data. Integrating and tracing data relating to primary genetic resource collections requires establishing common data standards, ontologies, and formats to ensure seamless data sharing and exchange.

Data Quality and Consistency: Ensuring the traceability, accuracy, completeness, and reliability of data is crucial for meaningful integration. Datasets from different sources may have varying levels of contextual meta-data, quality and consistency, which can lead to errors, ambiguities and misinterpretations if not addressed.

Representative Sampling: Many genebank accessions are not genetically pure lines and may represent a mixture of 'genotypes', especially those sampled from landraces and wild crop relative populations. Adequately accounting for sampling bias can enable better estimates of identity, diversity and differentiation.

Correspondence between accession and derived genomic data Genomic data may only be tangentially traceable back to the original accession because heterogeneity (above) or because materials were backcrossed and recurrently selfed before genotyping.

Diverse Data Types: Genebanks and breeding programs deal with a variety of data types, including genetic sequences, phenotypic traits, environmental data, and pedigree information. Data systems allowing efficient integration of all types of data are required to extract information from complex data sets.

Data Systems: are often user- or crop-specific, and not suited for multiple purposes or user groups.

DNA genotype data: functional gaps exist in datasets, genome coverage and current pipelines; the role and value of legacy data is seldom reviewed. Concepts such as pan-genomes are not yet widely integrated into systematic analysis of germplasm collections. Moreover, plant materials used for genotyping are not necessarily representative of curated accessions.

Attribute (phenotypic trait) data: lack consistent systems of organization or description, with different solutions developed *ad hoc* for different crops, and often independent of considerations of production, processing, supply chain or consumer descriptors.

Sociological

Legal and Intellectual Property Issues: Genetic resources and breeding materials are mostly subject to intellectual property rights and other legal restrictions. In particular**,** there are legal uncertainties around Access and Benefit Sharing (ABS) that inhibit freedom to operate (FTO), innovation and investment. This is often a key barrier for engagement of the private sector with public PGR collections.

Privacy and Security Concerns: Genetic and breeding data can contain sensitive information, such as genetic markers and personal identifiers. Ensuring data privacy and security while enabling integration requires implementing robust data anonymization, access controls, and encryption mechanisms.

Awareness of the Impact of Plant Genetic Diversity: Many processors, suppliers, marketeers, and consumers of added-value products are unaware of the available scope for raw material product quality variation embedded within PGR collections.

Lack of Incentives and Resistance to Change: Integrating data requires changes in workflows, processes, and sometimes even cultural norms within organizations. Resistance to adopting new technologies and practices can slow down integration efforts. If there are no clear incentives or benefits for these entities to share their data, they might be hesitant to collaborate.

Language Barriers: Description of crops and their attributes often differ in different regions of the world. Even within a country, scientific jargon and consumer adjectives may differ. In addition, PGR materials are used and assessed by different actors in the bio-economy.

Cultural and Organizational Differences: Genebanks and breeding programs often operate with different goals, priorities, and organizational structures. Individual breeders may overestimate or have a short-term view of the extent of variation to which they have access. Additional challenges exist to informed interactions between these sectors, R&D, Ag production, supply chain and end-product consumers.

Infrastructure

Long-Term Sustainability: **In**tegrating data is not a one-time effort; it requires continuous maintenance, updates, and investments. Ensuring the long-term sustainability of integrated systems can be a challenge.

Resource Constraints: Both genebanks and breeding programs may have limited human and financial resources to allocate towards data integration efforts. Lack of **access to** funding, skilled personnel, and technical expertise can hinder progress.

Technological Infrastructure: The technical infrastructure and IT systems of genebanks and breeding programs might not be compatible for seamless data integration. This includes challenges related to data collection and curation, data storage, data transfer protocols, data synchronization, and computing resources.

CONSIDERATIONS

- Use and description of genetic diversity differ between and within conservation, R&D, breeding and user communities.
- Achieving successful integration and interoperability may require aligning these differences and fostering collaboration between different stakeholders (eg adopting downstream terminology used by the food and health sectors in evaluating nutritional variation within PGR).
- Successful integration of information sharing systems hinges on collaborative efforts that bridge the gap between these distinct disciplines, fostering mutual understanding and shared objectives to mobilize and safeguard plant genetic resources.

PROPOSED APPROACHES TO INCREASING COMMUNICATION

We anticipate a series of prioritized and concerted actions that work towards seamless integration of information from conservation to consumption. These approaches address social, legal and technical issues, with an emphasis on:

- developing compatible data exchange standards (Section 4),
- navigating organizational and privacy concerns surrounding the sharing of sensitive genetic information,
- balancing the urgency of breeding programs with the meticulous preservation goals of genebanks.

Social, legal

- Clarifying ownership, usage rights, and licensing agreements is essential for integrating data while respecting legal boundaries.
- Identify incentives for traceability compliance by down-stream users of PGR - through provision of online registries and services that embed traceability.
- Development of training programs, including peer to peer help-desks and international support
- Engage community to introduce and extend involvement in development of dataset descriptive ontologies.
- Extend CoP to increase the number and diversity of informed stakeholders. Draw on the 'diverse lenses' of CoP to develop system-wide prioritization of information-sharing requirements, standards, and systems to capture value-added data.
- Stimulate engagement throughout supply chains, using regional and global initiatives to raise public interest in PGR, increasing awareness of genetic variation and difference between food (bio) products and the economic impact of PGR variation. Use analogies drawn from eg apple, wine grape and elsewhere.
- Encourage participatory involvement efforts, including establishing distributed 'common gardens' (*cf* the successful EU INCREASE program[43]).

[43] https://www.pulsesincrease.eu/

Technical:

- Integrate diverse data types by developing scalable generic methodologies and specialized tools that can handle such complexity.
- Aim for increased horizontal integration, encouraging funding investment to be focused on establishing generic (i.e. crop agnostic) underlying data management, data description and analysis tools. This will increase the ability to access and compare heterogenous or specialized data for a wide range of stakeholders.
  - Collate and share examples of good data and information management practice, aiming for an underlying generic framework that can be tailored to specific user requirements.
  - Provision of cloud data brokers and templates to ease data deposition and encourage wider usage of aggregations systems such as the FAO/GLIS GeneSys db, and interoperability with GRIN and other platforms.

Overall, many of the gaps identified may be addressed by broadening the remit of a dedicated crop PGR *Digital Commons*. In addition to specific recommendations outlined in Section 4, development of a Commons community is more likely to identify generic features and solutions for data/information integration, and provide examples of good practice by:

- Deploying demonstration apps that reduce barriers to entry in developing regions, including approachable AI/chat bots to enhance usage and adoption.

- Deploying dedicated user interfaces (skins) for PGR and downstream use, with training and support to make use of interoperable platforms such as Galaxy, Tripal and others that have benefited from investment in human and clinical science.

- Extending the the *DivSeek Commons* 'matrix'[44] as a first step to address the physical digital barrier of a lack of coordinated approach to a wider level of interoperability for PGR related information.

- Providing feedback for PGR collection management - through establishment of information management systems that ensure characterization data are accessible to collections by addressing traceability and inheritance of unique identifiers, along with data-mining tools and infrastructure.

- Value in demonstrating the value of market-focused vs crop-specific hierarchical classification of characterization and evaluation terms – by ensuring ontologies are fit for purpose for many aspects in supply/consumption. Phenotypic characterization/evaluation may include any measurable variation, and required recognition of value of different types, precision and layers of data.
  - Lower precision but rapidly acquired data may be of high value - exploitable by relevant statistics and/or AI.
  - There is value in characterizing large collections (eg INCREASE[45] and other initiatives).

---

[44] https://divseekintl.org/commons-landscape-matrix/; https://divseekintl.org/divseek-commons-the-matrix-is-born/

[45] https://www.pulsesincrease.eu/

    - Value is magnified when Information systems are FAIR[46] - easily accessible, interoperable, with traceable attribution of data.

- There is scope for integration of information in proprietary platforms such as the Breeding Management System (BMS)[47] . However, global tiered licensing models could be addressed to increase accessibility and wider adoption.

[46] Findable, Accessible, Interoperable and Reusable
[47] https://bmspro.io/

4. Data Standards: role, awareness, and implementation for PGR conservation and utilization

The past two decades have seen a transformation of human and economic behaviors arising from highly interconnected web services and intuitive GUIs. Federated information ecosystems with increased functionality are essential in domains such as healthcare (health information exchanges), finance (cross-institutional financial data sharing), government (inter-agency data collaboration), and research (data sharing among scientific institutions).

Such activities are enabled by **data standards** - established specifications that define how data should be structured, formatted and represented. These standards provide a common framework to ensure consistency, interoperability, and quality in data management, sharing, and exchange. Given the reach of PGR resources (Sections 1, 3) a concerted effort is required to address and demonstrate the value of adopting relevant, generic and scalable standards.

- In many sectors, proprietary platforms benefit from incorporating open standards and encouraging growth of wider developer communities. (see Section 4). e.g. car manufacturers, food supply are able to track provenance in a competitive commercial environment.

ANALYSIS

Role of standards: Adherence to data standards helps organizations and systems ensure data consistency, accuracy, and compatibility, facilitating data sharing and integration across disparate systems and enabling more effective data analysis, re-analysis, reporting, and decision-making.

- Application of data standards provides a common mechanism to represent knowledge, facilitating a shared understanding between stakeholders and supporting interoperability.
  - Simpler standards tend to be more widely adopted than more prescriptive standards, as such generic standards are those required by repositories, but community agreed enhancements may be applied in addition, to ensure more widespread acceptance and reusability (e.g. MIAPPE checklist on BioSamples).
  - Controlled vocabularies and consensus ontologies[48] provide a formalism that can transform the ability of complex networks of players to communicate, through common sets of definitions and establishment of structured relationships between diverse entities.

- The FAIR data principles[49] for scientific data management and stewardship aim to optimize the re-use of data by ensuring it is Findable, Accessible, Interoperable and Reusable.
  - Adopting a PGR Open Science Framework with FAIR will contribute to a game changing culture.
  - Recognizing that FAIR data does not equate to freely accessible.

---

[48] Ontology: a structured set of concepts and categories in a subject area or domain that shows their properties and the relations between them.

[49] https://www.go-fair.org/fair-principles/

- There are cumulative benefits from describing data with rich and structured metadata (data which describes and gives information about your data). This helps support discovery of the data (Findability), and provides the context to enable Reuse.

- Other technical challenges include: Other technical requirements include: human language translation systems; data translation layers, plug and play apps, pedigree visualization, the need to separate the data *vs* visualization tools.

Awareness within a PGR information ecosystem:

Data standards are often developed and maintained by relevant industry bodies, standards organizations or regulatory agencies to promote data consistency and accuracy. This reduces overheads of curation, categorization, data-related errors and ambiguities.

- Other Ag R&D-focused communities may come together to develop new standards or adopt and enhance existing standards (e.g. AgBioData[50]; Model Organism Databases, Gene Ontology Consortium[51], Reference Plant and Trait Ontologies via Planteome[52] and Crop Ontology, MIAPPE, etc.).
- Standards are usually required for submission to specific data repositories (e.g. genomic sequence repositories INSDC, OBO Foundry[53], etc.).
- Awareness of required standards prior to data generation is key, to ensure that all essential metadata are captured and prepared ready for submission and sharing.
- The past decade has seen some progress and success in data integration and increased awareness to ensure PGR data adhere to F.A.I.R. (and C.A.R.E) principles.

Status of standards useful for PGR and downstream use:

- **Unique Identifiers** – data should be assigned a globally unique and persistent identifier to support Findability. Digital Object Identifiers (DOIs[54]) are assigned to Plant Genetic Resources upon request by the GLobal Information System[55] (GLIS) of the UN FAO.
    - Current protocols specify that only the genetic material held by the PGR can be granted a DOI, which creates a handicap for individual labs and independent communities that maintain their own collection and distribution.
    - Also, many PGR collections enforce their own persistent uniform resource locator (PURL). Therefore a global approach and consensus is needed to consistently use one single approach to maximize the data collection, referencing, adoption and analysis.
- **Multi-Crop Passport Descriptors (MCPD)[56]** – these include information about the origin and distribution of a plant, scientific name and common names, and descriptors for morphological

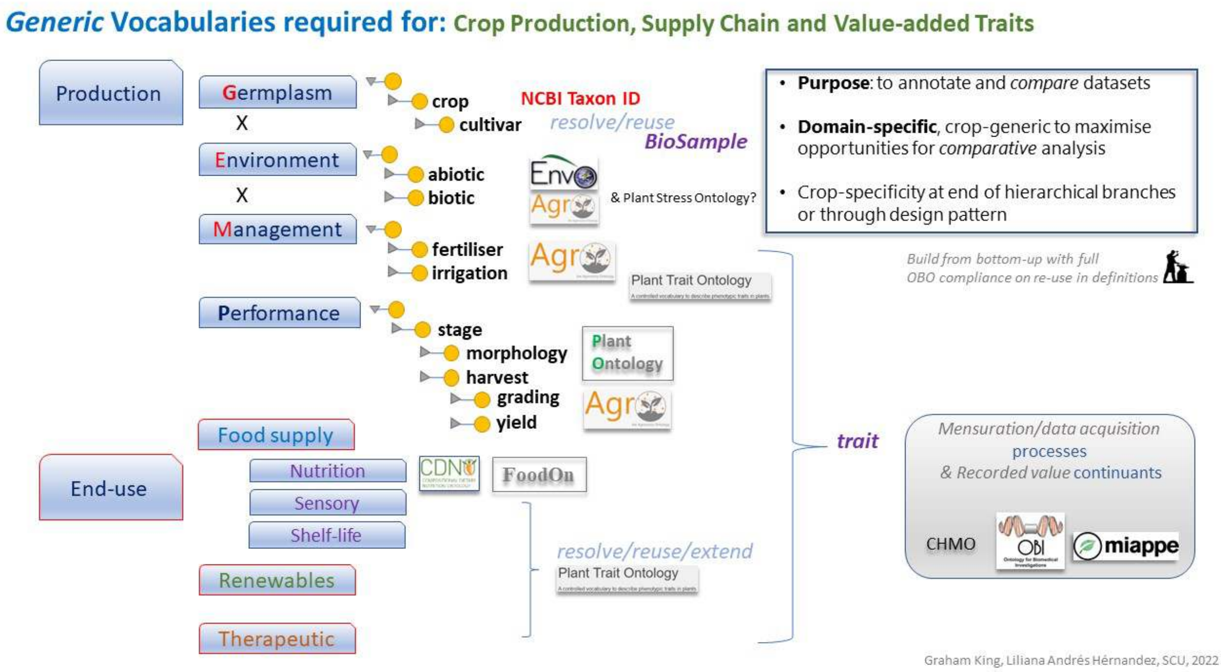
[50] https://www.agbiodata.org/
[51] https://geneontology.org/
[52] https://browser.planteome.org/amigo
[53] https://obofoundry.org/
[54] https://www.fao.org/3/I8840en/i8840en.pdf
[55] https://glis.fao.org/glis/
[56] https://alliancebioversityciat.org/publications-data/faobioversity-multi-crop-passport-descriptors-v21-mcpd-v21-december-2015

and agronomic traits. These may include factors that describe shape, size, yield, chemical composition, nutrient content, flavor, shelf-life of growing or harvested crops, pre and post-harvest crop management, as well as downstream use of crop raw materials in the food, fiber, medicinal or other sectors of the bio-economy.

- **Crop-specific descriptors**[57] have been developed independently for individual crops through international collaboration and translated into several languages.
- **Genomic Data Formats** – for deposition in sequence repositories (INSDC[58]), standard file formats such as FASTA, FASTQ, GFF3 and VCF are required to contain sequence reads & assemblies, genome annotations and genetic variations respectively. Metadata linking the samples to the genebanks should be captured also e.g. germplasm accession DOIs in the BioSamples records. When published these datasets should also be linked back to the genebank record, in accordance with the SMTA[59].
- **Minimum Information About Plant Phenotype Experiments** (MIAPPE)[60] – is an open, community driven, data standard designed to harmonize data from plant phenotyping experiments, providing a specification including a checklist and a data model of metadata required to adequately describe the experiments.
  - It is unclear how widely, if at all, the Darwin Core[61] (ex Dublin Core) meta-data frameworks are used in relation to PGR and downstream characterization data.
- **Breeding API** (BrAPI)[62] - is a RESTful web service API specification for communicating plant breeding data to enable interoperability among plant breeding databases.
- The CGIAR Generation Challenge Programme included development of meta-data standards and controlled vocabularies (ontologies).
- **Ontologies** - a few key generic ontologies such as TO (Trait Ontology) and PO (Plant Ontology) have been developed. Trait Ontology – has been a success and has some mapping to crop-specific ontologies[63] (and the FoodOn ontology covering food supply).
  - However, there are gaps between major vs minor crops (see Section 3), and is a gap for additional ontologies to cover relevant space through the PGR value chain.
  - Allow 'local' ontologies in different languages compatible with OBO[64] ecosystem, with apps to help AI-Plain language to OBO

---

[57] https://www.ecpgr.cgiar.org/resources/descriptors-technical-bulletins
[58] https://www.insdc.org/
[59] https://mls.planttreaty.org/itt/
[60] https://www.miappe.org/; https://nph.onlinelibrary.wiley.com/doi/full/10.1111/nph.16544
[61] https://dwc.tdwg.org/terms/; Darwin Core (often abbreviated to DwC) is an extension of Dublin Core for biodiversity informatics. It is meant to provide a stable standard reference for sharing information on biological diversity (biodiversity).
[62] https://brapi.org/
[63] https://cropontology.org/about
[64] Open Biological and Biomedical Ontologies; https://obofoundry.org/



CONSIDERATIONS AND COUNTER-ARGUMENTS

- Adhering to standards requires additional effort on the part of the data producer, especially when capturing rich metadata. Community culture around data submission and application of standards is key. The community should agree on where and how to share data in order to see collective benefits.
- Interoperability and adoption of FAIR and Open Science data standards are inhibited due to limited resources to maintain digital ecosystems serving PGR and downstream communities.
- Developments in AI and other IT are expected to help with interpretation of human language, although pragmatically information systems require compatibility at level of eg APIs[65] to facilitate interoperability, supplemented by flexible parsers.

PROPOSED ACTIONS

Ensuring progress on the areas outlined above requires concerted effort and visible benefits for the various stakeholders involved.

Social/organizational

- Stimulate innovation for adoption, completing and sustaining standards /ontologies by engaging and encouraging a broader community to introduce and extend involvement in the development of PGR descriptive ontologies. This will require careful stewardship to manage disruption and foster inter-community connections.
  - Establish a network for developers working on digital systems supporting PGR and

---

[65] Application Programming Interfaces

downstream communities.
    - Work with national/international informatics infrastructure providers for provision of lookup services to encourage consistent and persistent use of object identifiers, names, synonyms, etc.

- Coordinated /collective development of standards
    - AI ready - bring in expertise from other sectors (eg ICICLE initiative[66])

- Establishment of a Standing Group for standards authorisation and deployment
    - Ensure engagement with relevant communities and experts.
    - Potential to establish a Crop Use Ontology "authority" with wider remit and reach through the crop value chain - facilitate harmonization for federated local implementation.
    - Journals may also provide guidance on where data types must be submitted, requiring adherence to common standards.
        - it is important to ensure reviewers are aware of these requirements and helping them to validate associated files is key to supporting data reuse

- Tailor ontology use for specific production, supply chain and end-use applications.
    - Generate use case demonstrations of value of ontology annotation, where outcomes and insights are greater than the sum of parts.

- Apply generic ontologies to describe traits of minor and under-utilized crops, demonstrating value by focusing initially on areas such as nutritional composition, adaptation to future environments, food system resilience.

- Encourage evolution of best practices to map crop-specific breeder trait methods back to generic ontologies, recognizing that description of traits for genebanks do not have to be as specific as for breeding or research.

- Develop training material and host workshops to train/educate on data management and standards early in scientific careers and establish, and rewards for the additional effort (i.e. through recognition for submitting datasets and increased data citations linked to re-use) and data submission should be as easy as possible.

Technical

- Guidelines describing the standard and associated file formats should be publicly and readily accessible to all data generators.

[66] https://icicle.osu.edu/

- Data validation apps are typically provided for common data formats prior to submission to data repositories.

- Establish consistent use of DOI (not PURLs or other forms of IDs) for global consistency and adoption for every type of germplasm/seed material (held by PGRs and elsewhere) reported in data repositories, genomics data, and publications. This will allow successful implementation and use of citation indexing services to pick up and cross-reference genetic material, thereby allowing the PGR managers, resources, and authors to collect information, significance, and impact of their materials.
- Development of community standards for PGR sample metadata linked to genomic data records (similar to MIAPPE).

- Deposition of PGR-derived genomic data into relevant public repositories with clear attribution and traceability to a genebank's original accession ID (the DOI to BioSampleID link), requiring adequate IT tools and standard practises accepted by the relevant repositories and communities.

5. **Vision for a Digital Commons to harness the value of PGR**

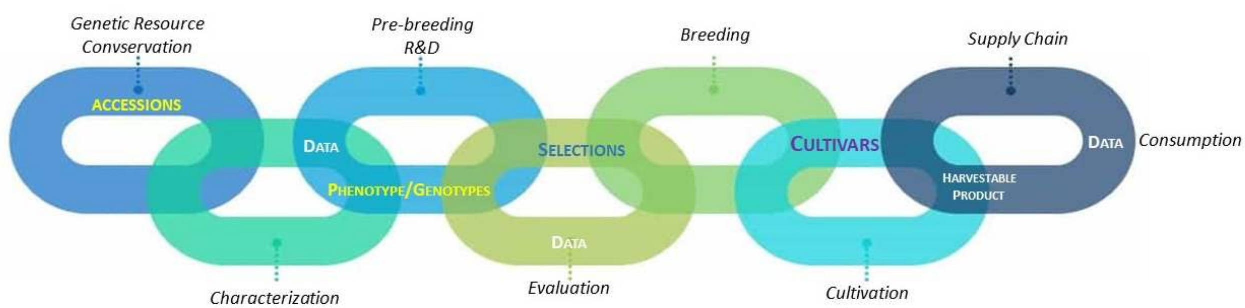
This document outlines key issues for enhancing information flows associated with conservation and use of crop plant genetic resources. Key strategic goals, objectives and activities were originally outlined in the Strategic Plan of DivSeek International Network published in 2021[67]. This Plan informed subsequent extensive discussion as presented in this document. Here we summarize a stepwise tractable roadmap for next steps, drawing on key points from Sections 1-4 above, followed by specific fundable actions requiring different levels of investment.

VISION

We propose **a global re-valorization of PGR-derived information flow and engagement to establish a PGR *Digital Commons*.** This will require more inclusive development, availability and adoption of best practices amongst different players in the crop research value chain.

Concerted efforts will be required to develop and adopt revised organizational and funding models to ensure sustainable cumulative benefits emerge from an evolving information *PGR Digital Commons* framework that may also support crowd-curation and open availability of characterization data.

Based on the economic, environmental and sociological imperatives outlined in sections 1-3, a key strategic requirement is to **raise awareness** of the value of cultivar and PGR derived information amongst Policy makers across the sectors of Health, Environment, Ag and the wider bioeconomy including biorenewables.



In particular, there is a timely opportunity to develop and demonstrate FAIR and CARE compliant information systems that highlight the role PGR in delivering tangible benefits for health and nutrition, climate mitigation and the wider bioeconomy.

SPECIFIC RECOMMENDATIONS

The recommendations here emphasize the potential for deriving and adding value to conserved crop Plant Genetic Resources. Implementation of the specific recommendations requires engagement of the

[67] https://divseekintl.org/strategic-plan/

wider PGR R&D, breeding support and end-use communities. This will be reliant on leadership, communication and active involvement spanning genomics and trait characterization data collection, informatics and standards development, drawing on expertise through the PGR value chain.

In addition to the Proposed Approaches & Actions outlined in Sections 1-4, we present a series of defined actions to develop a *PGR Digital Commons*.

FUNDABLE ACTIONS FOR ESTABLISHING A FUNCTIONAL DIGITAL KNOWLEDGE COMMONS

We have identified the need to develop a stratified approach to achieving the goals outlined above. These include a combination of long-term funding opportunities, along with specialized projects to focus on PGR characterization for distinct production environments, specific adaptive traits/phenotype, and quantifying the added value properties of harvested materials.

A trans-disciplinary cohort of R&D disciplines is required to address the current barriers and challenges to progress. This requires involvement and cooperation with economists and data scientists alongside commodity specialists. A series of resourcing levels can be used in achieving the vision outlined.

a. Small-scale < $200k, including postgrad projects
b. Medium-scale projects < $2m,
c. Larger-scale programs and capacity < $20m

Social, organizational

- Achieving the stated targets would benefit from extending the existing community of practice (CoP) established by DivSeek International Network[68], to include a broader set of vested stakeholders including those of the Crop Trust[69], One-CGIAR[70], the FAO ITPGRFA[71].

- Beyond PGR conservation activities, R&D, pre-breeding and public-sector breeders, the CoP would benefit from players in Ag production, supply, processing, added-value and marketing sectors, and consumers.

- Regular global and regional assessments of the wider value and socio-economic impact of *ex situ* PGR conservation and use is required in the context of traceable provenance, evaluation and characterization data through the research and wider crop value chain.
  - Such assessments may be defined in terms of quantified success, challenges and opportunity costs that addresses the changing value and sustainable use of PGR biodiversity for diverse end-uses in different regions of the world.
  - Include the specific impact of continued under-investment in development of generic data interrogation systems, standards and PGR networking.

[68] https://divseekintl.org/
[69] https://www.croptrust.org/
[70] https://www.cgiar.org/food-security-impact/one-cgiar/
[71] https://www.fao.org/plant-treaty/en/

- Identify and articulate incentives for a wider community through the Ag value chain to engage in a *PGR Data Commons.*
  - Identify opportunities to establish collectives or [virtual] communities of practice that interact with PGR repositories to enhance access to relevant germplasm.
  - Encourage innovation in the pre-competitive space through Agri-open targets and establishment of training networks.
  - Draw on advances in technological and biological knowledge to inform ongoing discussions around international treaties and Access & Benefit Sharing.

- DivSeek International Network represents an existing community of practice and are well placed to develop more cohesive and flexible governance frameworks to fill existing gaps as outlined in the Paper, by:
  - Capitalizing on the opportunity to extend the DivSeek Hub model[72] to establish a globally-federated set of independent but inter-communicative regional hubs. Vibrant hubs already operate in Latin/S. America, Oceania and parts of Africa. Current discussions indicate additional existing communities are well positioned to form active hubs with wider representation able to draw on complementary expertise in Africa, SE Asia, N. America and Europe.
  - Bringing together diverse communities to establish and disseminate best practice commons.
  - Providing incentives to establish a software and informatics developer network that is embedded in the PGR, research and breeder community.

- Opportunity to create networks that engage with specific target communities, focused around the wider bioeconomy beyond Ag production and food supply industries. Prioritized opportunities for demonstrating added-value in human nutrition and fiber and renewable raw materials sectors, alongside consumers and citizen science engagement programs.

- Stimulate creation of new centers that dock into genebanks to manage phenotyping of core-collections and connect with user communities via specialized projects.

Technical

- Encourage investment in generic tools that are applicable to a broad base of species/end-use attributes/regions.

- Development of information resources that integrate data and link back to PGR in genebanks.
  - Build on *DivSeek Common*s matrix to fill gaps for a federated and interoperable PGR-information ecosystem system.

- Generate and manage genomic databases, with a focus on genotyping complete sets of genebank collections (Appendix 2).

- Phenotypic characterization of complete or representative core PGR collections related to production, supply chain and consumption.

---

[72] https://divseekintl.org/regional-hubs/

- Recognize that low-cost, collaborative genotyping strategies could help to aggregate value to genebank collections. This could extend the approach adopted in the CGIAR "Excellence in Breeding" platform. To reach a far wider set of user communities, DivSeek and the Crop Trust are well placed to broker services for large-scale, cooperative genotyping for collectives of smaller crop genebanks.

- Establish integrated demonstration portals that incorporate
  - core reference sets of germplasm with reference genomes and training populations
  - adopt/enable information standards/vocabularies
  - collate phenotypic trait data relevant for a wide range of end users
  - provide seamless access to analysis and visualization tools.

Appendix 1: Indicative funding opportunities

(to be completed by relevant experts in each region)

In the USA

- RCN - Research Coordination Networks[73] ?
  - eg AgBioData
- NIFA[74] network to network
- National Science Foundation (NSF)[75]

Canada

- Genome Canada
- AAFC

S. America

Europe

- EU
- ELIXIR

Africa

- African Plant Breeding Association
- IAEA Collaborating Centre for Plant Breeding and Genetics

Asia

---

[73] https://new.nsf.gov/funding/opportunities/research-coordination-networks
[74] https://www.nifa.usda.gov/grants
[75] https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf23619

-

Oceania

- Bioplatforms Australia, Australian BioCommons/ARDC
- RDC network
- ACIAR

Appendix 2: Example funding initiative

Priority Research Initiative: Genomic Data Integration and Standardization for Enhanced Plant Genetic Resource (PGR) Utilization

A focused effort is required to develop and demonstrate how a PGR-derived information framework can address critical challenges in harnessing the potential of genomic data and improve the utilization of Plant Genetic Resources (PGR). This will enhance the value of PGR collections by standardizing how germplasm, genomic, evaluation and characterization data are integrated and promoting collaborative efforts. Key areas of focus include:

Data Quality Enhancement and Integration: Develop strategies to add value to PGR collections by improving data quality and combining diverse data sources. This includes addressing issues related to low-quality data layers and data fragmentation across various platforms.

Data Interoperability and Governance: Work on increasing data interoperability and unifying analytic tools used in the field. Explore different governance models to encourage greater participation and reduce data fragmentation. Consider the implementation of networks of networks and Federated database providers.

Standardization and Traceability: Establish data standards critical for managing and comparing diverse genomic data across gene banks. Develop methods for tracing and aggregating data from different gene banks and connecting it back to the initial data sources. Utilize hierarchical ontologies and reference trait ontologies to unify trait descriptions.

Value Proposition and Collaboration: Emphasize the broader value propositions of genetic resources, particularly untapped diversity, and their impact on various scientific communities and society at large. Promote the development of incentives for data sharing and demonstration projects to highlight the non-monetary benefits of data sharing.

Outcomes from this effort are expected to drive innovation, improve access to genomic data, and enhance the utility of PGR beyond breeding programs and immediate Ag production. It seeks to establish a collaborative ecosystem that connects the objectives of gene banks to broader scientific, economic, and societal contexts, ultimately reinforcing and characterizing gene bank collections through standardized data practices and shared knowledge.

Annex 3 Participants and contributors from workshop held at Cornell University June 2023, sponsored by DivSeek International Network. The Crop Trust supported selected attendees to travel to the session.

Achigan-Dako	Enoch G.	University of Abomey-Calavi, Republic of Benin	Benin
Bassil	Nahla	USDA-ARS National Clonal Germplasm Repository	USA
Breseghello	Flavio	Embrapa Rice and Beans, Brazil	Brazil
Coyne	Clare	Curator/Geneticist; United States Department of Agriculture	USA
Das	Aayudh	Pennsylvania State University	USA
Dyer	Sarah	European Molecular Biology Laboratory – European Bioinformatics Institute (EMBL-EBI)	UK
Fei	Zhangjun	Boyce Thompson Institute	USA
Jaiswal	Pankaj	Oregon State University	USA
Kehel	Zakaria	ICARDA	Morocco
Keilwagen	Jens	Julius Kuehn Institute	Germany
King	Graham	DivSeek International Network	Australia
Krishnan	Sarada	Crop Trust	Germany
Lange	Matthew	International Center for Food Ontology Operability Data and Semantics (IC-FOODS)	USA

Manzella	Daniele	FAO of the UN, Secretariat of the Plant Treaty	Italy
McClaughlin	Murray	Chair, Biomass Quality Network Canada	Canada
McCouch	Susan	Cornell University	USA
Muellar	Lukas	Boyce Thompson Instistute	USA
Norton	Sally	Australian Grains Genebank	Australia
Obreza	Matija	Crop Trust	Germany
Papa	Roberto	Università Politecnica delle Marche	Italy
Ribaut	Jean-Marcel	J.ribaut@cgiar.org	Mexico
Richards	Chris	USDA	USA
Sharma	Shivali	Crop Trust	India
Shaw	Paul	James Hutton Institute	Scotland
Sheenan	Moira	Cornell University	USA
Sibiya	Julia	University of KwaZulu-Natal; African Plant Breeders Association	South Africa