JUST AGRICULTURE
multidisciplinary e-Newsletter

Concept of Core collections of Plant Genetic Resources

***S.Sheelamary, K.Lakshmi, K.Dhanapal and R.Vivek Kumar**
Division of Crop Improvement, ICAR-Sugarcane Breeding Institute, Coimbatore

ARTICLE ID: 09

Abstract

Worldwide efforts to breed plants are dependent on the genetic diversity that exists in current cultivars and germplasm banks. In order to preserve valuable material before it vanished, breeding programs prioritized germplasm collection activities throughout the mid- to late 1900s. This led to the formation of the several vast germplasm resources that are currently available in various nations. Risks to the preservation of genetic resources have increased, along with a notable decline in the amount of expenditure spent on germplasm collection explorations, a rise in international biosecurity regulations and restrictions on breeders to freely exchange commercial germplasm. Plant breeders' capacity to use germplasm is limited by the size of most genebank collections, as well as issues related to germplasm passport data and the physical or genetic characterization of the accessions in collections. Core collection is a representative subset of the entire germplasm collection has grown in favor as a means of getting around these limitations. By including highly defined germplasm that tries to capture the majority of the variation in a collection as a whole, core collections seek to increase the usage of germplasm.

Key words: Core collection, quality evaluation, types and uses

Introduction

Under various environmental and climatic situations, plant genetic resources (PGR) are the cornerstone of sustainable agriculture and global food security and stability. One of the main issues causing the underutilization of material is the inadequate evaluation and characterization of germplasm conserved in genebanks, which also has an adverse effect on funding efforts and future opportunities. Plant breeding organizations use the germplasm stored in genebanks, frequently on a global scale, and breeding program depend on the genetic diversity that is available (Rajesekharan,2015). Elite crop cultivars are less able to adapt to environmental changes because they are typically bred from a limited genetic base and intended for high input, intensive agricultural production. In order to achieve future breeding targets that

JUST AGRICULTURE
multidisciplinary e-Newsletter

are climate resilient more work needs to be put into utilizing the germplasm found in genebanks.

At NBPGR national gene bank conserves as per the gene bank standards as base collections at -18° C. The current germplasm holdings in the National Gene Bank in the form of orthodox seeds are 3, 96,189 including released varieties and genetic stocks representing 1,584 species (http://www.nbpgr.ernet.in). However, the varieties released represent only a small proportion of the total variation. This is largely due to the negative effects of linkage drag, unfavourable alleles and the breakdown of the coadapted gene complexes that can be introduced from wild germplasm. Unless there is an extreme need, the breeders avoid adding highly unadopted germplasm into their breeding programs.

What is core collection?

The core collection is a small subset that minimizes genetic redundancy while preserving the maximum genetic diversity of the entire population. In order to represent the genetic diversity of the full germplasm resource with the fewest possible resources, core collection refers to the process of choosing a portion of the entire germplasm pool using certain techniques. The hierarchical structure model of genetic variety and the notion of neutral mutations provide the theoretical underpinnings for this idea. Germplasm resources are essential for genetic research because they provide a basic material and facilitate the identification and use of genes and traits with ecological and economic significance. Therefore, the preservation and utilization of germplasm resources are of significant importance for the development of new crop varieties, and a large number of germplasm banks have been established. But the diversity that has been gathered may not be completely and efficiently used because of their abundance, diversity, and lack of complete information about germplasm resources.

3. What makes the Core collection good?

- A good core collection should have no redundant entries. In the absence of significant differentiation among accessions, they can be considered as drawn from the same population, and hence be regarded redundant. Combining such material into a single accession is one of the possibilities to reduce the amount of redundancy in a collection, and to improve the composition thereof.
- Represent maximum the whole collection with regards to species, subspecies and geographical regions. Maximising the representativeness of genetic diversity implies

JUST AGRICULTURE
multidisciplinary e-Newsletter

also the inclusion of broadly adapted and heterotic materials containing 'generalist' alleles in a core collection. It may represent the full range of variation present in the whole collection or maximising the representativeness of the pattern of variation present in the whole collection.

- Should be small enough to be easily managed.
- Data completeness: The primary cause of the limited utilization of germplasm in crop improvement initiatives is the dearth of data on a vast number of accessions, especially for economically significant traits that exhibit high genotype × environment (G x E) interaction and necessitate multilocation assessment. Creating a "core collection," or roughly 10% of the total collection, has been suggested as a solution to the size-related collection problem. This collection would represent the genetic variability of the entire collection.
- Utilization: The fact that crop improvement scientists were not using genebanks to their full potential was evident from the rise in accession numbers within the collections and the lack of a corresponding increase in their use (Marshall, 1989). There is a huge discrepancy between the materials' actual use and their availability. This was valid for both national and international programs, including the CGIAR institutes. Similarly, in the national programs, the germplasm lines used in breeding programs are very limited (Upadhyaya et al.,2008).

Types of core collection

Based on the purposes for which they are formed core collections can generally be classified into three types or categories i.e. core collections representing (1) individual accessions; (2) extremes; and (3) distribution of accessions in the whole collection.

- **Type I:** A core collection that stand in for each of the collection's individual accessions (CC-I). In this instance, every item in the core collection stands in for one or more accessions that taken together comprise the entire collection. The entry in the core that most closely resembles each accession throughout the entire collection serves as its representation.
- **Type II:** The goal of a core collection of type CC-X is to accurately represent the genotypes, allele and phenotypic ranges of the entire collection. Entries in a good core collection of type CC-X are as dissimilar from one another as feasible.

JUST AGRICULTURE
multidisciplinary e-Newsletter

- **Type III:** A core collection that shows how the accessions for the entire collection are distributed (CC-D).In this instance, it is important to make sure that the proportion of accessions within a core collection accurately represents the numerical contributions made by the various regions or categories to the collection as a whole. For instance, the core collection ought to appropriately represent the significance of a particular geographic region if the bulk of the accessions originate from that area.

Common methods used for evaluating core collections

Several types of information can be used for selecting core collections. The most common type of data is (i) passport data (ii) agronomic data (iii) molecular marker data.

- **Passport data:** The identity, origin, and taxonomic classification of an accession, along with related information about domestication, distribution, breeding history, cropping patterns, and use, are all included in passport data. Examples of passport information include the origin nation, crop type (such as winter or summer wheat), and lineage.
- **Agronomic data:** Data related to agriculture can be categorical, discrete or continuous. Continuous variables include things like plant height, leaf area, and grain yield. Counts, such as the number of fruits or seeds in a pod, are dealt with by discrete variables. According to Crossa and Franco (2004), categorical variables can be classified as nominal (such as an organ's color or shape), ordinal (a visual scale set up to represent intensity, color, or size), or binary (presence or absence of a particular characteristic). Environmental factors and multiple genes are typically responsible for controlling agronomic traits.
- **Molecular data:** Data from molecular or biochemical marker systems can be treated as either continuous or categorical. Single nucleotide polymorphism (SNP), amplified fragment polymorphism (AFLP), random amplified polymorphic DNA (RAPD), and simple sequence repeats (SSR) are a few common types of molecular data (Odong et al.,2012).

Criteria used for evaluation of quality of core collections

Criteria	General comments
Summary statistics	Agronomic: Mean, range, variance, phenotypic correlation coefficients)

JUST AGRICULTURE
multidisciplinary e-Newsletter

	Molecular: gene diversity, heterozygosity, PIC, major allele frequency
PCA/ PCoA	Plot of the coordinates of the entries on main PCs to show spatial distribution of entries
Diversity index eg. SDI, Nei	Highest when uniform distribution and to be compared with the max possible value (logn)
Class coverage	The highest value (100%) is obtained when all the categories are represented in the core
Goodness of Fit	To test frequency distributions of important categorical traits.

Core collection developed at NBPGR, New Delhi

Crop	Entire collection	Core collection (Mini core)
Moong bean	1532	152
Sesame	3129	362
Brinjal	1798	181
Wheat	22469	2208 (224)
Rice (NEH)	6984	701
Lentil (wild)	405	96

(Source : NBPGR, New Delhi)

Use of Core Collection

Genomic Study and Marker Development

Molecular markers have broad prospects for application in the core collection, which can promote the protection and utilization of germplasm resources through evaluation of genetic diversity, help identify the species and varieties of core collections that are difficult to distinguish morphologically, an construct genetic maps of germplasm resources, which help understand their genetic structure and evolutionary history, providing a basis for comparative genomics research and the novel molecular markers derived from functional genes play a crucial role in improving agronomic characteristics such as yield, quality, disease resistance, and abiotic stress tolerance. Collectively, these novel molecular markers provide an effective approach for classifying germplasm resources, conducting genetic research, identifying the

JUST AGRICULTURE
multidisciplinary e-Newsletter

specific genes affecting important traits, and facilitating marker-assisted breeding (Gu et al., 2023).

Identification of Disease or Pest Resistance

Core collections, which represent the entire genetic diversity of germplasm resources, have been shown to improve the efficiency of identifying disease or pest-resistant accessions or genes which can greatly aid in disease-resistance breeding efforts and may even lead to gene editing and the introduction of novel alleles for single or multiple disease resistance in various crops.

Gene Discovery and Allele Mining

Core collections are valuable genetic resources for the identification of elite genes and mining alleles. Functional genes or markers can improve the efficiency and precision of selecting desirable traits and aid in accumulating favorable alleles for high-yield crop breeding. It is further necessary to work on mining of more underlying genes, manipulating the desired traits.

Conclusion

Integration of phenotypic and molecular data needs to be addressed more effectively and scientifically. It is necessary to improve phenotyping tools to enhance the efficiency of phenotyping. Optimizing the sampling strategy, The development of core collections for timber trees and bamboo and also the endangered forest species, is a top priority for the near future. We should make full use of genome sequencing information to develop molecular markers and gene discovery based on the core collection.

Reference

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.* https://doi.org/10.3390/f14050926

http://www.nbpgr.ernet.in

Marshall, D.R. (1989). Limitations to the use of germplasm collections. p. 105-120. *In*: The use

Odong, T.L, Jansen, J, van Eeuwijk, F.A, van Hintum, T.J. (2013). Quality of core collections for effective utilisation of genetic resources review, discussion and interpretation. Theor Appl Genet. 126(2):289-305. doi: 10.1007/s00122-012-1971-y.

of plant genetic resources [Brown AHD, Frankel OH, Marshall DR and Williams JT (eds)]. New York:Cambridge University Press.

Rajasekharan, P. (2015). "Gene banking for ex situ conservation of plant genetic resources," in *Plant Biology and Biotechnology.* eds. Bahadur B., Venkat R. M., Sahijram L. (New Delhi: Springer;), 445–459

Upadhyaya,H.D, Gowda, C.L.L and Sastry, DVSSR (2008). Management of Germplasm Collections and Enhancing Their Use by Mini Core and Molecular Approaches. APEC-ATCWG Workshop Capacity Building for Risk Management Systems on Genetic Resources 35~70