

## Core Collection for Efficient Management and Enhanced Utilization of PGR

**K.K. Gangopadhyay<sup>1</sup>, I.S. Bisht<sup>2</sup> and Rakesh Singh<sup>1</sup>**

<sup>1</sup> ICAR-NBPGR, Pusa Campus, New Delhi

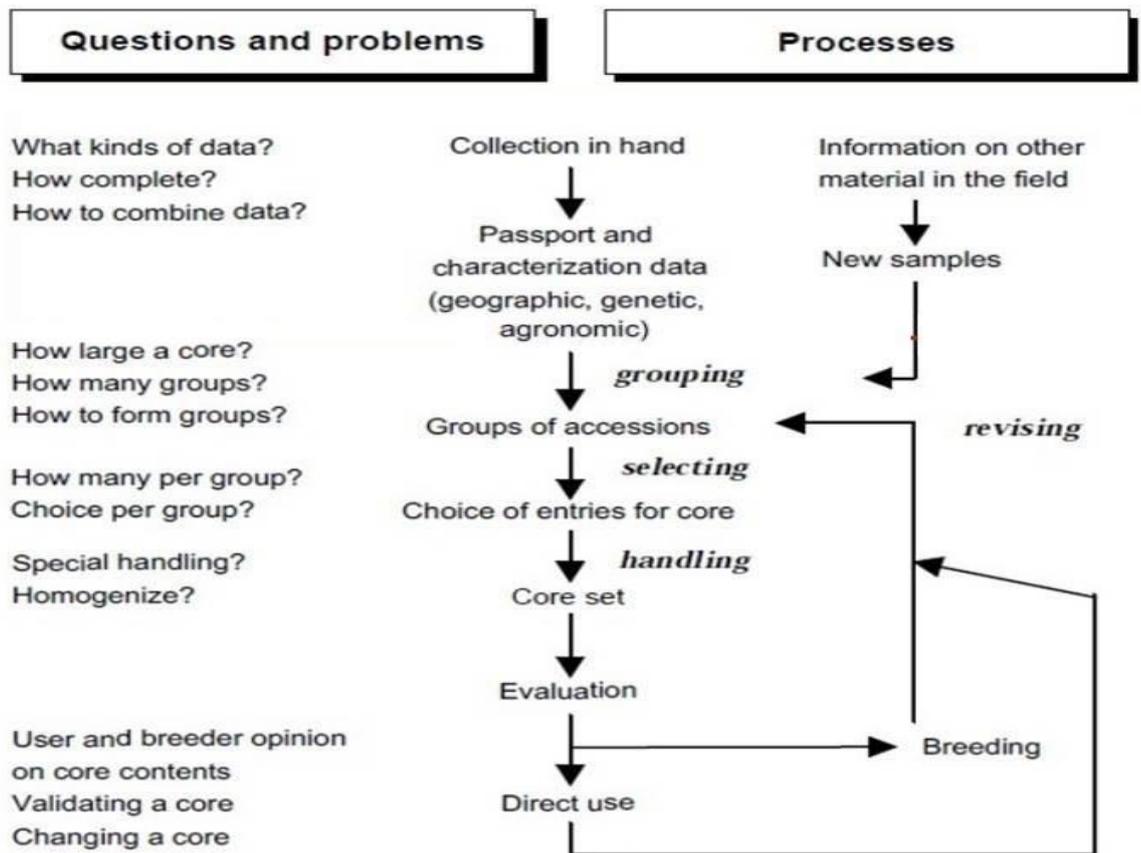
<sup>2</sup> ICAR-NBPGR Regional Station Bhowali

---

Genebanks (GBs) in the world are conserving large collections of genetic resources for easy access by the users/breeders/researchers to promote utilization. Ex-situ germplasm collections have increased enormously in number and size over the last three to four decades as a result of global efforts to conserve plant genetic resources for food and agriculture. This has complicated the characterisation, evaluation, utilisation and maintenance of the conserved germplasm. In some crops the collections are so large that may hinder the basic purpose of collections i.e. conservation through use of the genetic diversity. Perceiving the large collection may restrict/deter the use of genetic resources, Frankel (1984) defined core collection as 'a limited set of accessions representing, with a minimum repetitiveness, the genetic diversity of a crop species and its wild relatives'. Mini cores are developed from core collections having large number of accessions in order to have better management and usage. It is ensured that there is minimum similarity between the core entries and thus a reduced/limited size represents the genetic diversity of the entire collection of a crop, wild species, and group of species. The existing collections from which the core collection is obtained, is not replaced and the existing collection without the core collection is called as 'reserve collection'. With the Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture (FAO 1996) recommending core collection as one of the activities needed to improve the use of genetic resources, it has become accepted as efficient tools to improve conservation and use of genetic resources.

Van Hintum (1999) suggested a general procedure for developing a core collection. It comprises six steps viz. definition of the domain, decide on the size of the core collection, division into distinct groups, allocation of entries over the groups, choice of entries from each group that will be included in the core, representativeness of the core collection in terms of diversity and inclusiveness. The domain of the collection is dependent on the available material and on the objectives behind its establishment. A core collection can be global or regional or national core depending upon the domain of collections considered for its development. For example, International Barley Core Collection (Knüpffer and van Hintum 1995), Asian groundnut core collection from 25 countries in the Asian continent, core of sesame landraces from India (Bisht et al. 1998), etc. The domain may include the wild species alongwith the cultivated species for developing a core for both cultivated as well as wild species. The core collection when restricted to material with specific traits is termed as 'reference collection', such as local maize populations with a good combining ability (Radovic and Jelovac 1994) or *Pisum sativum* germplasm with disease resistance (Matthews and Ambrose 1994). The division of the domain into distinct group is done by divided into as genetically distinct groups as possible based on passport or taxonomic information or genotypic data. A stepwise, hierarchical procedure is usually followed to define the groups: first making the major divisions and subsequently splitting these subgroups into smaller ones depending upon the objectives. Core collections can be established based on taxonomy, phenotypes (Huaman et al., 1999), pedigrees (Martynov et al., 2003), geographic origins (Gangopadhyay et al., 2010), isozymes (Chandra et al., 2002) and DNA markers (Roy Choudhury et al.; 2014) and in combinations with two or more criteria. The number of entries in the core is arbitrary and based on the purpose of core collection. Usually, the number of entries chosen is substantially lower than the number of accessions in the domain, usually in the 5 to 20% range. Brown (1989a) using a sampling theory of selectively neutral alleles reported that 70% alleles can be retained in a core

collection containing about 10% entries of the entire collection. After the core collection size is decided, the number of accessions from each group to the core is estimated based on importance of or diversity in the subgroups. Generally, four strategies namely, constant (C), proportional (P), logarithmic (L) and diversity dependent (G) strategies is chosen for deciding the number of accessions from each group. Strategy G is superior to P strategy (Yonezawa et al., 1995). If quantitative information about the diversity in the subgroups is available the gene diversity (H) or maximization (M) strategy can be used (Schoen and Brown 1995). The final choice of entries from each group is generally made by scientists and curators with specific knowledge and experience with the groups. Crossa et al. (1995) suggested a multivariate analysis to select entries from each group wherever extensive marker, characterization or evaluation data are available. However, the relative contribution of each accession computed from Principal Component Score Strategy (PCSS) as suggested by Hamon et al. (1995) is a formal procedure to select entries from each group. Many approaches for selecting core collections have been proposed and used for example, M-Strat (Gouesnard et al., 2001), Genetic distance sampling (Jansen and van Hintum, 2007), PowerCore (Kim et al., 2007) and CoreHunter (Thachuk et al., 2009).



Flow chart to illustrate steps in developing a core collection.

Once, the core is developed, it is mandatory to evaluate the core to assess its homogeneity by comparing with the entire collection from which it has been obtained. It is better to evaluate the core from a different data set than the data set used for core development. Different numerical and graphical methods are used for evaluation of core collection. Among the numerical methods, summary statistics (mean, range, variance, co-efficient of variation, sign tests), frequencies, phenotypic correlation coefficients, principal component analysis, Shannon diversity index, class coverage, chi-square tests, per cent of alleles retained, etc. are predominantly followed. In addition, it is advisable that whenever possible or appropriate, the evaluation of core collections should be based on data that have not been

used for the selection of the accessions for the core collection. The core collection should be fairly stable after it is established, otherwise its management would be more difficult if its constitution is continually changing. Keeping in view the fresh requirement, the core set can be modified preferably by replacing rather than adding a new entry. The core collections can be used in developing collection strategy and management, information management, study and use of collections, and distribution of collections, etc. In India, core collections have been developed in mungbean, sesame, okra, brinjal, wheat and rice.

## References

- Bisht, I.S., R.K. Mahajan, T.R. Lokkathan and R.C. Agrawal. 1998 Diversity in Indian sesame collection and stratification of germplasm accessions in different diversity groups. *Genet. Resour. and Crop Evol.* 45:325-335.
- Brown, A.H.D. 1989a. Core collections: a practical approach to genetic resources management. *Genome* 31:818-824.
- Chandra S, Huaman Z, Hari Krishna S, Ortiz R. 2002. Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data: A simulation study. *Theor Appl Genet* 104: 1325–1334.
- Crossa, J., I.H. DeLacy and S. Taba. 1995. The use of multivariate methods in developing a core collection. Pp. 77-92 in *Core Collections of Plant Genetic Resources* (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.FAO (1996)
- FAO. 1996. *Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture*. FAO, Rome, Italy.
- Frankel, O.H. 1984. Genetic perspectives of germplasm conservation. Pp. 161-170 in *Genetic Manipulation: Impact on Man and Society* (W. Arber, K. Llimensee, W.J. Peacock and P. Starlinger, eds.). Cambridge University Press, Cambridge.
- Gangopadhyay, KK, R. K. Mahajan, Gunjeet Kumar, S. K. Yadav, B. L. Meena, Chitra Pandey, I. S. Bisht, S. K. Mishra, N. Sivaraj, Rajeev Gambhir, S. K. Sharma, and B. S. Dhillon. 2010. Development of a Core Set in Brinjal (*Solanum melongena* L.). *Crop Sci.* 50:755–76.
- Gouesnard B, Bataillon TM et al 2001. MSTRAT: an algorithm for building germ plasm core collections by maximizing allelic or phenotypic richness. *J Hered* 92(1):93–94
- Grenier C, Hamon P, Bramel-Cox PJ. 2000. Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum collection using random vs non-random sampling procedures A. Using morpho-agronomical and passport data. *Theor Appl Genet.* 101(1-2):190–196
- Hamon, S., M. Noirot, and F Anthony. 1995 Developing a coffee core collection using the principal component score strategy with quantitative data. In: T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales (ed) *Core collection of plant genetic resources*. Intl. Plant Genet. Resour. Inst. (IPGRI). John Wiley & Sons, NY.
- Huaman, Z., Anguilar, C. and Ortiz, R. 1999. Selecting a Peruvian sweet potato core collection on the basis of morphological, eco-geographical and disease and pest reaction data. *Theoretical and Applied Genetics*, 98: 840-844.
- Jansen J, van Hintum TJL. 2007. Genetic distance sampling: a novel sampling method for obtaining core collections using genetic distances with an application to cultivated lettuce. *Theor Appl Genet* 114(3):421–428
- Kim KW, Chung HK et al. 2007. PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. *Bioinformatics* 23(16):2155–2162
- Knüpfper, H. and Th.J.L. van Hintum. 1995. The Barley Core Collection - an international effort. Pp. 171-178 in *Core Collections of Plant Genetic Resources* (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK. Martynov et al. (2003)

- Martynov, S.P., Dobrotvorskaia, T.V., Dotlacil, L., Stehno, Z., Faberova, I. and Bares, I. 2003. Genealogical approach to the formation of the winter wheat core collection. *Russian Journal of Genetics*, 39: 917–923.
- Matthews, P. and M.J. Ambrose. 1994. Development and use of a ‘core’ collection for the John Innes *Pisum* collection. Pp. 99-107 *in* Evaluation and Exploitation of Genetic Resources, Pre-Breeding. Evaluation and exploitation of genetic resources: pre-breeding. Proceedings of the Genetic Resources Section Meeting of Eucarpia, 15-18 March 1994.
- Radovic, G. and D. Jelovac. 1994, The possible approach in maize core collection development. Pp.109-115 *in* Evaluation and Exploitation of Genetic Resources, Pre-Breeding. Evaluation and exploitation of genetic resources: pre-breeding. Proc. of the Genetic Resources Section Meeting of Eucarpia, 15-18 March 1994, Clermont-Ferrand, France (F. Balfourier and M.R. Perretant, eds.). Institut National de la Recherche Agronomique, Versailles, France.
- Roy Choudhury D, Singh N, Singh AK, Kumar S, Srinivasan K, et al. 2014. Analysis of Genetic Diversity and Population Structure of Rice Germplasm from North-Eastern Region of India and Development of a Core Germplasm Set. *PLoS ONE* 9(11): e113094. doi:10.1371/journal.pone.0113094
- Schoen, D.J. and A.H.D. Brown. 1995. Maximising genetic diversity in core collections of wild relatives of crop species. Pp. 55-76 *in* Core Collections of Plant Genetic Resources (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Thachuk C, Crossa J, Franco J, Dreisigacker S, Warburton M, Davenport GF (2009) Core Hunter: an algorithm for sampling genetic resources based on multiple genetic measures. *BMC Bioinformatics* 10:243
- van Hintum, Th.J.L. 1994. Hierarchical approaches to the analysis of genetic diversity in crop plants. Pp. 23-34 *in* Core Collections of Plant Genetic Resources (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- van Hintum, Th.J.L. 1999. The Core Selector, a system to generate representative selections of germplasm accessions. *Plant Genet. Resour. Newsl.* 118:64-67.
- Yonezawa, K., T. Nomura and H. Morishima. 1995. Sampling strategies for use in stratified germplasm collections. Pp. 35-54 *in* Core Collections of Plant Genetic Resources (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.