Indian J. Genet., 69(4) (Spl. issue): 294-299 (2009)

Role of core collection and pre-breeding in management and use of genetic resources for designing crops under changing climate

Anurudh K. Singh*

H.No. 2924, Sector 23, Gurgaon 122 017

Abstract

The success to designing new cultivars, adapted to the changed climate primarily depends on the information regarding the genetic variability available within the taxonomic gene pool of cultivated species. This needs quantification of genetic variability within manageable set of collections and information about the phylogenetic relationships between distant sources of genetic variability and the cultivated species to enable introgression of desirable gene(s) into cultivated gene pool in a usable from. The core collection approach and gene pool grouping followed by pre-breeding can play an effective role in providing access to wide range of genetic resources, bringing their desirable gene(s) into cultivated species to meet the challenges of climatic changed. The present article discusses the possible application of these approaches along with concerns and future perspective.

Key words : Core collection, genetic resources, gene pool

Introduction

Designing of crops adapted to foreseeable climatic changes, such as, unpredictable weather, frequent droughts, increasing global temperatures, decreasing soil fertility, salinity etc. needs improved management of available genetic resources to facilitate effective use and pre-breeding efforts for introgression of desired genes from crop wild relatives, which have evolved resilience against adverse climatic condition over centuries. Improved management of genetic resources will help easy access to desired variability, both genetic and allelic for their use in further genetic improvement/ designing of cultivars, suited to changing climatic conditions. The success of any crop improvement programs depends largely on the extent of quantification of available variability and access to desirable genetic variability available in germplasm collections to the breeders, in conventionally usable form.

Presently, the problem is that in most important crops, the collections have grown very large and it has become difficult for the breeders to identify appropriate genetic diversity for his use. In fact, it has been realised that due to this situation a very little (1.5% from the total collections) percent of collections are used by breeders [1]. In fact the large collections, instead of prompting greater use, are creating the “problem of plenty”, raising questions, like what germplasm to begin with for desired genetic enhancement programme. One of the reasons for this situation is that quantification of genetic diversity in large collections has become difficult to facilitate identification of desirable accessions. Therefore, the curators and breeders have to be empowered with tools that should fulfill basic needs of identification of desirable germplasm from large collections. The **core collection** approach**,** based on genetic, ecological and statistical parameters of quantification of total genetic variability of collections can help in capturing the total spectrum of genetic variability, distributed over agro-ecological/ geographical regions into a smaller set of collections representing total spectrum, which would be easy to manage and use.

Often it happens that the desired genetic variability available in primary gene pool is either in genetic background not adapted to the breeding or target climate or in poor agronomic background for its direct use in conventional breeding programme and needs **pre-breeding** for its conversion to facilitate its use in specific breeding programme. At other times the desirable genetic variability is not available within cultivated gene pool, and is available in genetically distant wild relatives of crops. These wild relatives, because of their genetic divergence, either refuse to cross by conventional means due to various levels of crossing barriers, restricting hybridisation with cultivated species or produce sterile or partially fertile hybrids or if crossable

*Corresponding author’s e-mail: anurudhksingh@gmail.com
Published by Indian Society of Genetics & Plant Breeding, F2, First Floor, NASC Complex, PB#11312, IARI, New Delhi 110 012
Online management by indianjournals.com

do not release regular segregates making the genetic introgression of desirable genes difficult or even impossible through conventional breeding means. Therefore, for making the genes/alleles available with these wild relatives accessible to conventional breeders, there is a need to understand the phylogenetic relationships of such species with cultivated species to identify appropriate breeding strategies for incorporating desired genes, for designing new cultivars. The application of **gene pool** concept, based on cross-compatibility relationships can play a basic role facilitating classification or grouping of genetic diversity within related taxa and help identification of appropriate breeding strategy for incorporation of desirable gene from distant sources into conventionally usable form of cultivated species. Based on this information **pre-breeding** efforts can facilitate hybridisation, establishing fertile hybrids and conventional cytogenetic manipulations to improve genetic recombination for incorporation of desirable genes and bringing them into a usable form. Alternatively, special biotechnological approaches may be needed to provide access to these genes through sexual or parasexual means of genetic transformation.

Role for core collection

Otto Frankel [2] suggested development of core collections that could represent total spectrum of variability of total collections. The basic elements of core are- 1) the original collection has taxonomic integrity; 2) the core subset has a small size; 3) the core subset is a representative sample; and 4) it is diverse. It expects unequal numbers from different classes like cultivated versus wild, or per subspecies, or per botanical varieties or geographical areas. Developing core collection involves-

1. Assembling all the relevant data, which would help grouping, sub-grouping and clustering of accessions based on discriminatory information about some strongly inherited characters.

2. Defining the collection to be represented in the core.

3. Deciding the size of core collection to ensure that it comprise a representative selection of accessions from each of the designated groups. On the basis of sampling theory of selectively neutral alleles, Brown [3] argued that the number of accessions in the core collection should be about 10% of the total collection, with a maximum of 3000. This level is effective in retaining about 70% of alleles of the entire collection.

4. Grouping of accessions into categories of genetic similarity. It may begin with the groupings by taxonomy (species, subspecies, and races) followed by major geographic groups (country, state), climate or agro-ecological regions. Further sub-division within each group should be based on strongly inherited characters such as genetic polymorphism, pathogen, pest and abiotic stress resistance, habit, etc. The abundant discriminating data will require multivariate analysis to define the groups. The clustering within the broad geographical group could be done to sort accessions into clusters using standard hierarchical clustering methods. The problem of how to use different traits (continuous, discrete, ordinal, multi-state, binomial) and eliminate scale differences by standardising each variable by means of either standard deviation or its range. The Ward [4] method was found the best clustering strategy when the sizes of the groups are similar and UPGMA (Unweighted pair group with arithmetic mean method) is appropriate when the groups are of different sizes [5]. Hence groups should reflect the major genetic and ecological categories within the entire collection.

5. Selecting the core entries - how many from each group and which ones is the next question. Three strategies of Constant number, Proportional number and Logarithmic number has been used of which, theoretically Logarithmic provides optimal number for fully differentiated loci, whereas Proportional is optimal for undifferentiated loci [6].

However, the review of literature on the core developed reflects that most are more based on multivariate analysis of selected quantitative traits without the concern for representativeness of variability in statistical terms. Therefore there is a need that representativeness of core to original collection must be ensured using both parametric and nonparametric statistical methods. This will ensure representativeness of genetic diversity for quantitative traits of entire collection, that *Variance of Homogeneity* and *Homogeneity of Distribution* of different classes is same between entire and core collections, conservation of the associations under genetic control, and that the percentage of significant differences for mean and variance, coefficient of variation and diversity index for

each trait in both core and entire collections is similar.

The other factor, because of which the present cores have not been very effective, is lack of dynamism of core collection. For this the core need to be multiplied, made homogeneous, conserved separately, and evaluated further for important traits. They may need to be made flexible to meet diverse requirement. For example, a smaller, but a representative set (mini core), or to meet a specific interest, more accessions with specific trait. van Hintum [7] has developed the core selector to meet such requirements. It is relatively simple based on a formalization of the normal procedures to create a core depending upon the purpose. The size of the core subset can be set between 1 to 10 percent for total or the domain trait. This can be divided into distinct groups depending upon the nature of traits for which a domain is defined. The stepwise division of groups of qualitative nature can easily be accomplished as they are distinct by nature. Quantitative traits can be divided into groups based on the distribution and range value for the trait among accessions. Nevertheless, there is a need for further evaluation of core for characters of breeding value involving breeders. Breeders must have information about core and should participate in the evaluation for- diversity within core collection, evaluating breeding potential of exotic collections, method for identification of promising parents and use of information on core for exploitation of reserves collections. Mini core may also be developed with another step of evaluation of the core subset for various morphological, agronomic, and quality traits, and selecting a further subset of about 10% accessions from the core, consisting of only about 1% of the entire collections.

Core collections have been established for most important field crops, such as wheat (ICARDA- 2002 wild *Triticum* core collection; 2003 durum wheat core collection developed), rice (The USDA rice core collection, including 1,794 accessions from 114 countries), sorghum (ICRISAT-GREP Patancheru), pearlmillet (ICRISAT-GREP Patancheru), chickpea (ICRISAT-GREP Patancheru; 2004 Kabuli chickpea core collection developed), pigeonpea (ICRISAT-GREP Patancheru), groundnut (ICRISAT-GREP Patancheru), rapeseed mustard (*B. napus* core collection from Europe), soybean (Chinese core; also developing mini core for US). Similarly, in vegetables, core is available for common bean (*Phaseolus vulgaris* L.), cassava (*Manihot esculenta* Crantz), okra (*Abelmoschus esculentus* L.), potato (*Solanum tuberosum* L.), quinoa (*Chenopodium quinoa* Willd.), mung bean [*Vigna radiata* (L.) Wilazek] and sweetpotato (*Ipomoea batatas* L.), as cited by Singh and Dhobal [8].

Therefore, potentially, **core collection** can play a major role in designing new cultivars by providing a working collection representing total spectrum of variability that can be extensively examined for economic traits, researched and used. However, it would be possible, provided the core has no redundancy, based on authentic information about accessions, sufficiently large with representatives from major sub-specific taxa and geographic regions and broadly adapted rather than having ecologically specialised alleles or genotypes, to function as more logical and effective tool to identify sources of desirable traits for their utilization in crop improvement, within the constraints of utility and genetic diversity as measured by the number of alleles per locus. Further evaluation of core may help in identification of gaps (of variability and regions) for prioritisation, may indicate sources of acceptable expression and on rare variant the accessions from a hot spot be evaluated or the rare character be searched in a core subset of related wild species.

Some concerns

However, there still remains some concern about use of core, as core is chosen to represent diversity and not to represent usefulness. Lack of general validity in sampling of variation is another concern that originates from reliability of information on genetic diversity, which is the basis of core and key for usefulness, and from the fact that the character of breeder's interest may be rare. If a core is formed on incomplete or misleading information or patterns of diversity of whole collection, it is possible that it could leave out important types. Some character of interest may be rare, and rare variant may be absent from core, as it may be only one in an entire collection. Therefore, identification of such variant from entire collection depends on luck or the capacity to cope with sampling of entire collection. Further evaluation of a core collection provides information on whether the variant is rare? Some cores are still large for easy use (10% contains thousands of accessions) and need further decrease in size. Therefore, diverse core is more likely to contain adequate sources of many characters.

Future perspective

In light of these concern and experiences, it is required that the core should be made more dynamic with

periodic revision, the receipt of new accessions into the collection from distinctively new sources, replacement of accessions of questionable authenticity, revision of groupings in the light of new data, review of users' needs, ensuring inclusion of better accessions for specific traits and rationalisation of core set using molecular markers, ensuring capture of both genetic and allelic diversity covering the whole genome. The core should help to develop knowledge about relationships among variables, which is much needed.

An alternative proposal

The present core are more based on quantitative traits as the emphasis is on representing the total diversity than use and therefore the characters accounted may or may not have association with characters of breeding value to meet various challenges. Therefore, there is a need to develop core based on qualitative or quantitative variability of useful traits, particularly related with stress, yield and quality traits. This may involve-

- Preparation of gene pool or group of desired or breeding value traits, both qualitative and quantitative, from base collection (resistance to biotic and abiotic stresses, dwarfness, male sterility, early maturity etc.) representing both taxonomic and geographic groups.
- Application of multivariate analysis on quantitative traits associated with mechanism of functioning of genes/factors.
- Develop clusters using appropriate statistical method.
- Pick up representative accessions from each cluster using appropriate sampling method.
- Develop core set with representativeness of each group from taxonomic and geographical point of view.
- In Indian context, the effort may be integrated by developing a core on Indian collections, which can be enriched with variability available in global core.

Role for pre-breeding

Harlen and de Wet [9] in order to provide a genetic perspective and focus for cultivated plants in relation to other components of genetic diversity, classified the constituent taxa based on cross-compatibility, into- (1) primary gene pool (GP-1), (2) secondary gene pool (GP-2) and tertiary gene pool (GP-3).

Primary gene pool (GP-1) corresponds with the traditional concept of the biological species. All components of this gene pool cross easily; produce hybrids that are generally fertile with good chromosome pairing; normal gene segregation, therefore gene transfer is easy. It includes spontaneous races (wild and/or weedy) belonging to subspecies, which may be of use, needing little pre-breeding effort.

Secondary gene pool (GP-2) refers to all species that will cross with the crop species and are experimentally cenospecies. They have barriers to hybridisation, which may be poorly or not at all developed and produce hybrids that may be weak, partially sterile, difficult to bring to maturity with recovery of desired types in advanced generations etc. Gene transfer is possible, but with **pre-breeding** to overcome the cross-incompatibility barriers and various possible manipulations to establish a fertile hybrid and improve chromosome pairing and recombination rates. This gene pool offers greatest variability and is available for use with **pre-breeding** efforts.

Tertiary gene pool (GP-3) includes species, which are cross-incompatible and if crosses with crop species, produce hybrids that tend to be lethal or completely sterile. It may require **pre-breeding** efforts, such as embryo rescue, grafting, or bridge species to establish hybrid and doubling chromosome number to overcome ploidy barriers and/or obtain some fertility or other biotechnological approaches. As the GP-3 defines the outer limits of potential conventional genetic reach, gene transfer is either not possible with known techniques or difficult. The most powerful tool known for introducing genes from GP-3 is either use of complex hybrids that may work as bridge or the biotechnological approaches.

The ploidy barriers between species are not always strong and gene transfer across ploidy levels may be rather easy in some cases. The concept of gene pool has been further developed by Smartt [10] with minor modification, giving equal value to considerations of genetic resources. GP-3 may be big and can be further classified on the basis of degree of genetic isolation expressed in terms of cross-compatibility relationship [11]. Following **pre-breeding** approaches may be resorted to produce interspecific derivatives carrying the gene(s) conferring desirable traits that would meet the challenges of changing climate.

A. Conventional cytogenetic manipulations

Direct hybridisation

1. *Same chromosome number:* easiest route for

incorporation of wild species germplasm through direct hybridisation followed by manipulations to improve chromosome pairing and recombination rates producing fertile segregates for screening against target stress.

2. *Different chromosome number:* Four basic methods are employed.

   a. Direct hybridisation at different ploidy level

   b. Raising the ploidy level of wild species followed by hybridisation

   c. Doubling chromosome number of the wild species or their hybrids before hybridisation to cultivated species and re-synthesising chromosome number equal to cultivated species.

   d. Reducing the chromosome number of cultivated species before hybridisation and re-synthesising chromosome number equal to cultivated species.

Bridge Crosses

This method has been used through crosses with a species common in cross-compatibility relationship between donor and recipient cultivated species. Bridge crosses have been used in wheat, *Cucurbita* and *Solanum.*

Chromosome Manipulations

Several addition lines have been produced in wheat with improved quality of disease resistance, earliness, protein content etc., but none of these have been accepted as commercial variety. Substitution lines can be produced only in species where extensive cytogenetic and genetic knowledge has accumulated.

Utilization of chromosomal translocation has also been used for gene transfer in case of wheat, tobaccos, rice etc., through various manipulations like irradiation where the desired gene is situated on non-homologous chromosome.

B. Embryo-rescue and cytogenetic manipulations

Embryo culture has been advantageously used for establishment of hybrids when the endosperm degenerates after fertilization. This technique has been widely used for obtaining interspecific or intergeneric hybrids. Successful application of this technique has resulted in production of hybrids in case of *Phaseolus, Trifolium, Gossypium, Cucurbita, Lycopersicon, Hordium, Triticum* etc. This is followed by various cytogenetic manipulations to obtain hybrid fertility, increased chromosome pairing and genetic recombination.

C. Genetic transformation

The recent development in biotechnology had made it possible for transfer of genes from tertiary gene pool and beyond the taxonomic boundaries through the asexual method of genetic transformation. This approach may involve several steps starting from gene isolation using molecular marker technology, production of gene constructs and incorporation of the genes through agro-bacterium based genetic transformation system or alternative in-plant method of transformation, such as particle bombardment with isolated DNA, co-cultivation, microinjection etc.

Other methods

In addition to above methods a variety of other methods have been used either to facilitate production of hybrids or to facilitate the chromosome clearing and genetic recombination. A variety of growth regulators have been applied to overcome the pre-fertilization and post-fertilization barriers to initiate seed production. Similarly, additional methods have been applied to utilize wild species germplasm through manipulation of sporophytic or gametophytic incompatibilities through use of mentor pollen. Also, attempts have been made to establish hybrids through somatic hybridisation technique where genomes of different species are combined through the fusion of somatic cells. However, these methods have not been very successful and therefore rarely use in the crop improvement programme.

The pre-breeding, particularly wide hybridization exploiting wild relatives has established credentials with successful stories of overcoming challenges in many crops. There has been periodical reviews [12-14] and international conferences and books reviewing the progress made in pre-breeding/wide hybridization incorporating desirable gene from alien sources, overcoming the various stress challenges. For example, in potato by 1980 49% of varieties had *S. demissum* late blight genes [15]. Similarly, nine *Lycopersicon* species, provided almost all traits for improvement of tomato [16]. In rice, several wild species, *Oryza rufipogon, O. longistaminata* and *O. glaberrima*, have been used through chromosome segmental substitution incorporating resistance to biotic and abiotic stresses [17]. In many other crops, such as wheat, maize, cotton,

sugarcane, groundnut, potato, tomato etc. they have been used for incorporating diverse traits, such disease and insect resistance (wheat, rice, potato, tomato), yield (QTL) (oat, tomato), quality (fruit size; TSS) (grasses, pigeonpea, tomato), earliness and adaptation (rye, potato, tomato, grape, strawberry), modes of reproduction (cytoplasmic male-sterility in several crops), miscellaneous traits (hard-seediness, colour, leaf texture, delayed ripening).

Future perspective

A recent review reflects that though there has been a steady increase in release of cultivars containing wild species genes, there is continued emphasis on their use for a wider range of characteristics, also, they are gaining in importance and prevalence, but, their contributions in development of new cultivars remain less than expected, given available improved procedures, advances in molecular methods for managing backcrossing programmes, increased numbers of wild species accessions in gene banks, and the substantial literature on beneficial traits [14]. Therefore, greater **pre-breeding** efforts are needed in meeting the challenges of climatic change for which the present cultivars has become more vulnerable due to funnelling towards a selected set of genes and narrowing down of genetic base [18].

References

1. **Duvick D.** 1984. Genetic diversity in major farm crops on the farm and in reserve. Economic Bot., **38**: 161-178.
2. **Frankel O. H.** 1984. Genetic perspectives of germplasm conservation. In: W.K. Arber, K. Llimensee, W. J. Peacock and P. Starlinger (eds). Genetic Manipulation: Impact on Man and Society. Cambridge University Press, Cambridge. pp. 161-170.
3. **Brown A. H. D.** 1989a. Core collections: a practical approach to genetic resources management. Genome, **31**: 818-824.
4. **Ward J.** 1963. Hierarchical grouping to optimise an objective function. J. Am. Stat. Assoc., **38**: 236-244.
5. **Sokal R. and Michener C.** 1958. A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull., **38**: 1409-1438.
6. **Brown A. H. D.** 1989b. The case of core collection. In: The use of plant genetic resources. Brown, AHD *et al.* (eds.) Cambridge University Press, Cambridge. pp. 135-56.
7. **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.
8. **Singh A. K. and Dobhal V. K.** 2002. Gene pools and core collection concepts for effective conservation and use of germplasm. *In*: Proceeding of International Conference on Vegetables. 11-14 November, 2002. Bangalore, India, pp. 33-40.
9. **Harlan J. R. and De Wet J. M. J.** 1971. Towards a rational classification of cultivated plants. Taxon., **20**: 509-517.
10. **Smartt J.** 1990. Grain legumes – Evolution and genetic resources. Cambridge Univ. Press, Cambridge. 379p.
11. **Singh A. K. and Simpson C. E.** 1994. Biosystematic and Genetic Resources. *In*: The Groundnut Crop, A scientific basis for improvement, J. Smartt (ed.). pp. 96-138. Chapman and Hall London.
12. **Stalker H. T.** 1980. Utilization of wild species for crop improvement. Advances in Agronomy, **43**: 199-241.
13. **Singh A. K., Moss J. P. and Smartt J.** 1990. Ploidy manipulations in interspecific gene transfer. Advances in Agronomy, **43**: 199-241.
14. **Reem H. and Toby H.** 2007. The use of wild relatives in crop improvement: A survey of developments over the last 20 years. Euphytica, **156**: 1-13.
15. **Ross H.** 1979. Wild species and primitive cultivars as ancestors of potato varieties. In: Proc. Conf. Broadening Genet. Base of Crops. 1978. Wageningen Pudoc. Wageningen, The Netherlands.
16. **Rick C. M.** 1979. Potential Improvement of tomatos by controlled introgression of genes from wild species. *In*: Proc. Conf. Broadening Genet. Base of Crops. 1978. Wageningen Pudoc. Wageningen, The Netherlands.
17. **Brar D. S.** 2005. Broadening the gene pool of rice through introgression from wild species. In: Rice is Life: Scientific Perspective for 21 Century. K. Toriyama, K. L. Hong and B. Hardy (eds.), IRRI, Phillipines. pp. 157-160.
18. **Tanksley S. D. and McCouch S. R.** 1997. Seed banks and molecular map: unlocking genetic potential of the wild. Science, **277**: 1063-1066.