

Reference sets of food crop germplasm for international collaboration GCP Workshop held in Montpellier, November 13-17, 2008

Jean-Christophe Glaszmann¹, Carmen de Vicente²

1 – CIRAD, avenue Agropolis, 34398 Montpellier Cedex 5, FRANCE
2 – GCP, MEXICO

Access to genetic diversity available in large crop germplasm collections requires identification of representative samples with smaller size to make them suitable for different surveys: screening of traits, evaluation of phenotypic diversity, evaluation of combining ability, assessment of molecular diversity, etc. Moreover, integrating diverse types of characterization on the same materials makes it possible to assess correlations among traits and investigate gene effects such as epistasis and pleiotropy. Passport data enable selecting based on eco-geographic information; molecular markers offer means to further refine assessment of relatedness and to reduce sample size. Use of standardized methods yields data that can be compared across materials, laboratories and time, providing a durable momentum to enrich global understanding and representativeness.

The first phase of the GCP has yielded massive data sets featuring mostly SSR diversity (12 to 50 loci) among large germplasm samples (300 to 3000 accessions) of the major world food crops. This has served for identifying reference samples of 50 to 500 accessions. These are being further handled as genetic stocks and data are being ascertained for a subset of high quality SSR markers. This is meant, among others, to allow comparison of any new material to this international reference. Thus any novel diversity born by newly accessed germplasm can be easily identified, and used for enriching the global panel; conversely the global panel can be used to enrich the diversity accessible anywhere, e.g. in a germplasm collection of specific mandate (e.g. regional) or any breeding programme.

Altogether this leads to a major GCP product: diverse germplasm reference samples with validated data of reference markers, accessible as a **global public good** in a robust form. This workshop was an opportunity to describe and discuss all these steps and aspects of reference set identification, as well as the perspectives and the mode of organization that is necessary for taking full advantage of the initiative.

71 participants from 28 countries, specialized scientists involved (30%) or not (20%) in GCP projects, germplasm scientists and managers (25%) and breeders (25%)



Programme

Session 1 "Sampling within germplasm: Where we are",

- Establishment of core collections from passport/phenotypic data
- Sampling from molecular data
- Sampling diversity, methodological aspects

Session 2 "Use cases of reference sets"

- Molecular vs phenotypic diversity: between historical and functional correlations",
- Regional diversity

Brainstorming session (BS1), "Type and quality of service"

Session 3 "Biological stocks in a context of management, access and distribution"

- Germplasm policy/management
- Use of germplasm by breeders"

Brainstorming session (BS2) "Link between genebanks and breeders, breeders expectations"

Session 4 "Challenges and obligations in an international context"

- Quality of the data
- Way ahead: such as a re-dynamised breeding community, the integration of geographic data, a set-up for genotyping and phenotyping networks....

Main conclusions

The Generation Challenge Program approach was affirmed in exploring, evaluating and using germplasm diversity, in defining reference sets on the principles that all GCP products must be delivered to users.

The expectation extends also to related initiatives in the System-Wide Genetic Resources Program (SGRP) of the CGIAR, the Global Crop Diversity Trust and the Global Partnership Initiative for Plant Breeding Capacity Building (GIPB). GCP-derived reference sets can play a prominent role provided that there is clarity on political availability of the resource and the attached data by placing them under the Multi-Lateral System and that research continues on their use as an entry to the broader collections. Their description and access have a natural place in the web system being put in place by the SGRP and Bioversity, which can therefore be used for publicizing and ordering.

Information has been shared about how many reference sets are available, how they were compiled and what their status is. Together with this, we have been made aware of the broad diversity of situations among crops and some limitations that need to be overcome, on representation and comprehensiveness of original composite collections, choice and justification of methodology to select reference sets, data quality of both collections, composite and reference, traceability of germplasm, availability of the minimum required information to be distributed with the material and quality control (label, certification, ...) of the material to be distributed

Next steps – Issues to be addressed (currently addressed) and action to be taken (currently taken)

Governance: Who will decide which sets are ready to get the GCP seal of quality and approval? How will this decision be made?

Multiplication: Who will do it? Who will finance it? Relationships between CG Centers and NARS, Development of the sets and ensuring their quality and their readiness should go together with the development of the information systems to support them (to be developed in conjunction with Bioversity, Compliance with Multilateral System (MLS), Availability of methodologies is a must

Evaluation of the demand: survey past and current demand for minicore/reference sets, Circulate questionnaire among curators and networks of crop researchers (questionnaire to be accompanied by a clear description), write simple document on concept, utility, methods, ... pertaining to the reference sets

Upgrade situation on a crop by crop basis: Collect/compile passport data and currently available information on composite set

Have molecular data analysis validated by outsider (see project G4007.01)

- Quality assessment with simple criteria: Inform original lab(s) of abnormalities found, Produce new (if necessary) GCP template data matrix (TDM): derived from labX and used as template by GCP for structure analysis, Gross structure analysis with standard methods and compared with reports and publications of lab X
- Have data compared between the GCP TDM and the results of the validation exercise (with reference markers): Analyze discrepancies between data; types, sources, distributions, Infer impact on global quality of gross structure
- Compare structure between GCP TDM and GCP reference data (reference set of materials characterized with reference markers; GCP reference data should reflect the structure obtained with GCC TDM)

Identify expert committee per crop: Establish expert committee with breeders and partners, Request expert advice for representativeness and complementation, Collect/compile/update information on status of seed production/availability

Promote visibility through SINGER-ALIS, SGRP, CDT, GIPB, Coordinate with Bioversity officer, Communicate, distribute, publish