V. Construction of coffee core collections

François Anthony[1] *and Stéphane Dussert*[2]

[1] Plant Genetics & Genomics, Institut de recherche pour le développement (IRD), UMR RPB, 911 avenue Agropolis, BP 64501, F-34394 Montpellier cedex 5, France
[2] Seed Conservation Biology & Technology, Institut de recherche pour le développement (IRD), UMR DIA-PC, 911 avenue Agropolis, BP 64501, F-34394 Montpellier cedex 5, France

Introduction

The concept of a 'core collection' was proposed to enable efficient and cost-effective management and utilization of crop genetic resources (Frankel 1984; Frankel and Brown 1984) and has been interpreted in various ways (van Hintum et al. 2000). Frankel (1984) defined a core collection as a limited set of accessions representing, with a minimum of repetitiveness, the genetic diversity of crop species and its wild relatives. For practical uses, the core collections allow setting up a large representation of the genetic diversity within a reduced set of genotypes, which can be intensively evaluated and widely distributed. However an IPGRI survey of 1346 genebanks and institutions worldwide pointed out considerable confusion and lack of knowledge on what a core collection is (Brown and Spillane 1999). For all intent and purposes, core collections were never intended for conservation purposes, but rather to facilitate use of conserved material. If core collections are to have a meaningful impact on management of germplasm collections, there is a need for greater consensus and knowledge among the curators and users on what is and is not a core collection.

The experience reported here is based on the construction and management of *C. arabica* core collections in CATIE, firstly for evaluation (Anthony et al. 2001), then for long-term conservation (Vasquez et al. 2005).

Coffee core collections

Genetic diversity is not randomly distributed among species and populations, but it can generally be represented by a hierarchical model, a tree (Hamon et al. 1995; Noirot et al. 2003). This model was adopted for constructing coffee core collections, using passport data combined with knowledge of the structure of the gene pools (Noirot et al. 1993).

Evaluation purposes

A representative core collection of the Ethiopian accessions conserved in the CATIE genebank was defined prior to genotypic evaluation (see Chapter 4). Considering the molecular analysis capacity in the CATIE biotechnology laboratory, the core collection was finally composed of 88 Ethiopian accessions (109 genotypes). The sampling was based on geographical data, namely the collecting sites of the FAO (Fernie et al. 1968) and ORSTOM (Guillaumet and Hallé 1978) surveys in Ethiopia. All accessions from the provinces outside south-west Ethiopia were analysed because of their low representation in the genebank (Table 5.1). A selection was necessary within the 482 accessions from south-west Ethiopia. The selection was based on the collecting sites, assuming that the accessions derived from spontaneous trees growing in forest or from subspontaneous plants cultivated on small farms presented a higher diversity than those collected in semi-industrial and industrial farms, where growing coffee trees were not usually derived from spontaneous plants. In fact, visitors to Ethiopia have reported that

it is very difficult to establish from where *C. arabica* is truly native, as it spreads rapidly from cultivation and becomes naturalized in clearings and along trails in the forest (Sylvain 1955; von Strenge 1956; Meyer 1965; Friis 1979). Based on RAPD analysis carried out at CATIE, four groups of accessions were identified (Anthony et al. 2001): one large group in the south-west (Kefa and Ilubabor provinces) and three smaller groups in the south and south-east (Sidamo and Harerge provinces) (see Chapter 4). This result was in accordance with historical data on coffee domestication in south-west Ethiopia (Sylvain 1955; von Strenge 1956; Lejeune 1958). A similar structure of Ethiopian diversity was found using amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) markers (Anthony et al. 2002).

Table 5.1. Geographical origin within Ethiopia of the accessions included in the *C. arabica* core collection for neutral marker analysis of diversity (Anthony et al. 2001).

Region	Province	No. of collection sites	Living accessions in CATIE genebank	No. of accessions selected
South-west	Kefa	337	356	38
South-west	Ilubabor	116	126	31
South	Sidamo	8	8	8
South-east	Harerge	2	2	2
Centre	Gojjam	6	6	6
Centre	Shoa	2	2	2
North	Eritrea[†]	1	1	1

Notes: [†] At the time the accessions were collected, the area was a province of Ethiopia.

Table 5.2. Composition of the *C. arabica* core collection (74 accessions) representative of the genetic diversity present in wild and cultivated accessions from Ethiopia and Yemen, and conserved in the CATIE field genebank.

Origin	Selected accessions
WILD (FAO collection)	T.4472, T.4476, T.4495, T.4497, T.4501, T.4505, T.4579, T.4619, T.4621, T.4661, T.4662, T.4664, T.4665, T.4666, T.4758, T.4759, T.4819, T.4824, T.4837, T.4857, T.4863, T.4864, T.4865, T.4893, T.4900, T.4938, T.4942, T.4945, T.4952, T.4958, T.4960
WILD (ORSTOM collection)	T.16689, T.16690, T.16691, T.16692, T.16694, T.16695, T.16697, T.16700, T.16702, T.16704, T.16705, T.16706, T.16707, T.16709, T.16712, T.16713, T.16714, T.16723, T.16724, T.16726, T.16729, T.16733, T.16737, T.16739, T.17177, T.17205, T.17207, T.17223, T.17232
CULTIVATED (locally in Ethiopia)	T.2710, T.2711, T.2722, T.2724, T.2727, T.2742, T.2748, T.2754, T.2915, T.3097, T.4007
CULTIVATED (locally in Yemen)	T.21233, T.21239, T.21240

Cryopreservation purpose

The classification of Ethiopian accessions into genetic groups was used for constructing a representative core collection of the *C. arabica* gene pool for long-term conservation in liquid nitrogen (see Chapter 6). The sampling also included accessions representing varieties locally cultivated in Yemen, which is considered as the primary centre of *C. arabica* dispersion outside Ethiopia (Meyer 1965). The maximum number of accessions to be included in the core collection was

imposed by financial constraints on storage capacity. It was decided to include 74 accessions in the core (Table 5.2). Because of self-compatibility of *C. arabica*, one genotype per accession was harvested in the CATIE genebank. It was selected on the basis of phenotypic data (not published), principally resistance to Coffee leaf rust (*Hemileia vastatrix*) and root-knot nematodes (*Meloidogyne* spp.). The core collection could thus be considered as representative of the diversity detected by neutral markers, and in addition it was improved in resistance genes.

Conclusion

The construction of core collections provides genebank managers, breeders and research scientists with a manageable number of accessions for their work. The strategy described for coffee could be easily applied to other crops, especially non-orthodox-seed species. Prior to evaluation, a whole collection can be stratified using the data on accession origin, which are commonly recorded in genebanks. After genotypic evaluation, the accessions can be classified according to their genetic group and then sampled within the groups. This contributes to optimize the genetic diversity retained in a subset of the whole collection for long-term preservation.

A pragmatic attitude was adopted for sampling the coffee accessions, taking into account phenotypic traits of interest for breeders as well as technical constraints limiting the core collection size. As frequently mentioned in other crops, the major constraint in constructing core collections is the availability and reliability of data (Ng and Padulosi 1992). It seems important to remember that a core collection formed by simple random sampling of the accessions has surprisingly good retention statistics (Brown 1989) and might actually be better than one biased by poor data (Brown and Spillane 1999).

References

Anthony, F., Bertrand, B., Quiros, O., Wilches, A., Lashermes, P., Berthaud, J. & Charrier, A. 2001. Genetic diversity of wild coffee (*Coffea arabica* L.) using molecular markers. Euphytica 118:53–65.

Anthony, F., Combes, M.-C., Astorga, C., Bertrand, B., Graziosi, G. & Lashermes, P. 2002. The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. Theoretical and Applied Genetics 104:894–900.

Brown, A.H.D. 1989. The case for core collections. pp. 136–156, *in:* A.H.D. Brown, O.H. Frankel, D.R. Marshall and J.T. Williams (editors). The Use of Plant Genetic Resources. Cambridge University Press, Cambridge, UK.

Brown, A.H.D. & Spillane, C. 1999. Implementing core collections – principles, procedures, progress, problems and promise. pp. 1–9, *in:* R.C. Johnson and T. Hodgkin (editors). Core Collections for Today and Tomorrow. IPGRI, Rome, Italy.

Fernie, L.M., Greathead, D.J., Meyer, F.G., Monaco, L.C. & Narasimhaswamy, R.L. 1968. FAO coffee mission to Ethiopia, 1964-65. FAO, Rome, Italy. 204 p.

Frankel, O.H. 1984. Genetic perspectives of germplasm conservation. pp. 161–170, *in:* W.K. Arker, K. Llimensee, W.J. Peacock and P. Starlinger (editors). Genetic Manipulation: Impact on Man and Society. Cambridge University Press, Cambridge, UK.

Frankel, O.H. & Brown, A.H.D. 1984. Current plant genetic resources – a critical appraisal. pp. 1–11, *in:* J.H.W. Holden and J.T. Williams (editors). Genetics: New Frontiers (vol. IV). Oxford and IBH Publishing, New Delhi, India.

Friis, I. 1979. The wild populations of *Coffea arabica* L. and cultivated coffee. pp. 63–68 *in:* G. Kunkel (editor). Taxonomic Aspects of African Economic Botany. Ayuntamiento de las Palmas, Las Palmas de Gran Canaria, Spain.

Guillaumet, J.-L. & Hallé, F. 1978. Echantillonnage du matériel récolté en Ethiopie. Bulletin IFCC 14:13–18.

Hamon, S., Dussert, S., Noirot, M., Anthony, F. & Hodgkin, T. 1995. Core collections: accomplishments and challenges. Plant Breeding Abstracts 65:1125–1133.

Lejeune, J.B.H. 1958. Rapport au Gouvernement Impérial d'Ethiopie sur la production caféière. FAO, Rome, Italy.

Meyer, F.G. 1965. Notes on wild *Coffea arabica* from southwestern Ethiopia, with some historical considerations. Economic Botany 19:136–151.

Ng, N.Q. & Padulosi, S. 1992. Constraints in the accessibility and use of germplasm collection. pp. 45–50, *in:* G. Thottappilly, L.M. Monti, D.R. Mohan Raj and A.W. Moore (editors). Biotechnology: Enhancing Research on Tropical Crops in Africa. Technical Centre for Agricultural and Rural Cooperation, Wageningen, The Netherlands.

Noirot, M., Anthony, F., Dussert, S. & Hamon, S. 2003. A method for building core collections. pp. 65–75, *in:* P. Hamon, M. Seguin, X. Perrier and J.-C. Glaszmann (editors). Genetic diversity of cultivated tropical crops. Collection Repères. CIRAD, Montpellier, France.

Noirot, M., Hamon, S. & Anthony, F. 1993. L'obtention d'une "Core Collection" de caféiers. Définition des groupes d'échantillonnage et méthodologie. pp. 183–190, *in:* Proceedings of the 15th International Scientific Colloquium on Coffee. Montpellier, France, 6–11 June 1993. ASIC, Montpellier, France.

Sylvain, P.G. 1955. Some observations on *Coffea arabica* L. in Ethiopia. Turrialba 5:37–53.

van Hintum, Th.J.L., Brown, A.H.D., Spillane, C. & Hodgkin, T. 2000. Core collection of plant genetic resources. IPGRI Technical Bulletin No. 3. IPGRI, Rome, Italy.

Vasquez, N., Salazar, K., Anthony, F., Chabrillange, N., Engelmann, F. & Dussert, D. 2005. Variability in response of seeds to cryopreservation within a coffee (*Coffea arabica* L.) core collection. Seed Science and Technology 33:293–301.

von Strenge, H. 1956. Wild coffee in Kaffa Province of Ethiopia. Tropical Agriculture 38:297–301.

Anthony François, Dussert Stéphane (2007)

Construction of coffee core collections

In : Engelmann Florent (ed.), Dulloo M.E. (ed.), Astorga C. (ed.), Dussert Stéphane (ed.), Anthony François (ed.). Conserving coffee genetic resources: complementary strategies for ex situ conservation of coffee (*Coffea arabica* L.) genetic resources : a case study in CATIE, Costa Rica

Rome (ITA) ; Paris : Biodiversity International ; IRD, 45-48

ISBN 978-92-9043-752-9