RESEARCH ARTICLE

Genetic diversity of a *Coffea* Germplasm Collection assessed by RAPD markers

Milene Silvestrini · Mirian P. Maluf · Maria B. Silvarolla · Oliveiro Guerreiro-Filho · Herculano P. Medina-Filho · Marina M. T. Vanini · Adalgisa S. Oliveira · Cristiana de Gaspari-Pezzopane · Luiz C. Fazuoli

Abstract Genetic diversity and relationships within and among nine species of *Coffea*, one species of *Psilanthus* and the Piatã hybrid from the Coffee Germplasm Collection of Instituto Agronômico de Campinas (IAC), Brazil were assessed using RAPD markers. Genetic diversity and relationships were evaluated by proportion of polymorphic loci (P), Shannon's genetic index (H' and G'_{ST}) and clustering analysis. The overall RAPD variation among all accessions was mostly partitioned between rather than within species. However, *C. canephora* and *C. liberica* showed a high genetic diversity within the species ($\underline{H}'_{sp} = 0.414$ and $\underline{H}'_{sp} = 0.380$, respectively) and this was highly structured (high $\underline{G}'_{ST}$). Genetic diversity from *C. congensis* and *C. arabica* was also structured, but with lower levels of genetic diversity ($\underline{H}'_{sp} = 0.218$ and $\underline{H}'_{sp} = 0.126$, respectively). The results were consistent with agronomic and molecular studies and demonstrated that the IAC *Coffea* Collection is representative of the phylogenetic structure observed in the genera. This study devises sampling strategies for coffee germplasm collections and provides genetic diversity parameters for future comparisons among them.

Keywords *Coffea* · Genetic variability · Molecular markers · Plant collections

M. Silvestrini (✉) · M. B. Silvarolla · O. Guerreiro-Filho · H. P. Medina-Filho · M. M. T. Vanini · A. S. Oliveira · C. de Gaspari-Pezzopane · L. C. Fazuoli
Centro de Café "Alcides Carvalho", Instituto Agronômico de Campinas (IAC), Av. Barão de Itapura, 1481, P.O. Box 28, Campinas, SP 13012-970, Brazil
e-mail: milene.silvestrini@gmail.com

M. P. Maluf
Empresa Brasileira de Pesquisa Agropecuária (Embrapa), Campinas, Brazil

Introduction

The development of new coffee cultivars by breeding programs depends on the existence of germplasm collections with significant and accessible genetic variability. The field Coffee Germplasm Collection of Instituto Agronômico de Campinas (IAC) includes accessions of 14 species of genus *Coffea* L. and three species of genus *Psilanthus* Hook.f (Fazuoli et al. 2007; Medina-Filho et al. 2007). The species better represented in the collection are the commercial *Coffea arabica* L. and *Coffea canephora* Pierre ex A. Froehner. The others are wild species that, although without commercial value, represent an important source of genetic variability to characteristics like architecture, disease and pest resistance and various agronomic and industrial traits. Currently, the IAC Coffee Germplasm Collection is one of the most representative *Coffea* collections

conserved *ex situ*. Once direct access to genetic resources from the center of origin of several *Coffea* species may be restricted, the availability of germplasm collections such as this is crucial for coffee breeding programs.

The efficient use of genetic resources of a germplasm bank requires, besides the introduction, characterization and proper maintenance of its accessions. Coffee Germplasm Collection of IAC has been maintained and characterized by traditional methods such as analysis of morphological (plant height and fruits, leaves and seeds color and size) and agronomic characteristics (productivity, disease and pest resistance, cup quality) (Silvarolla et al. 1999; Gaspari-Pezzopane et al. 2004; Aguiar et al. 2005; Medina-Filho et al. 2007). However, this type of evaluation has been insufficient to characterize all available genetic resources, in part due to intrinsic characteristics of coffee like long life span, low genetic diversity of commercial *C. arabica* plants (Lashermes et al. 1996; Maluf et al. 2005) and difficulties of in vivo maintenance of some species under the biotic and abiotic hazards of the field collection. In addition, field collections are costly and demand intensive labor, so it is necessary to accurately assess the level of genetic diversity in order to minimize duplication and establish core collections.

Molecular markers are a powerful tool for rapid and efficient access of genetic variability and have been used in germplasm banks and breeding programs of various crop species (Rafalski and Tingey 1993). Previous studies demonstrated that RAPD analysis is a reliable and effective method to assess genetic variability within and among *Coffea* species (Orozco-Castillo et al. 1994; Ruas et al. 2000; Maluf et al. 2005) as well to provide markers linked to genes of interest for genetic improvement (Agwanda et al. 1997; Ram and Sreenath 2000).

In this study, we analyzed the genetic diversity and relationships of *Coffea* Germplasm Collection of IAC through RAPD markers in order to (1) assess the level and distribution of genetic diversity of the collection, (2) determine the genetic distance of the accessions within the collection, and (3) use the information for the coffee breeding program as well as to devise sampling strategies for collections of coffee germplasm.

Material and methods

Plant material

A total of 180 individuals of nine species of genus *Coffea* L., one species of genus *Psilanthus* Hook.f. and the Piatã hybrid (Centro de Café's code C387) (Medina-Filho et al. 2007) present in the Coffee Germplasm Collection of IAC (Table 1) were selected for analysis. Sampling of individuals was representative of total species/variety/cultivars/ populations accession number at the Germplasm Collection of IAC. Eight cultivars of *C. arabica* developed by the coffee breeding program at IAC and commercially cultivated in Brazil as well as some mutants of this species were included. The cultivars are Acaiá (AC 474-4), Mundo Novo (MN 388-17), Bourbon Amarelo (BA), Catuaí Vermelho (CV 81), Catuaí Amarelo (CA 100), Icatu Vermelho (IV 4045), Obatã (OB), Tupi (TP) and the mutants are Maragogipe (Mar.), Volutifolia (Vol.) and Polyorthotropica (Pol.). Young leaves were collected from plants of each variety, cultivars or population, frozen in liquid N_2, and kept at $-80°C$ until used.

Genomic DNA extraction

Total genomic DNA was extracted from frozen young leaves according to Paillard et al. (1996), using CTAB as detergent. All DNA samples were diluted to a final concentration of 20 ng/μl.

RAPD amplification

A total of 40 ng of each DNA sample was used in PCR reactions for RAPD markers amplification. Random commercial 13-mer oligonucleotides from Operon Technology, kits A, G, and X were used (OPG 05, OPG 06, OPG 10, OPG 17, OPG 18, OPA 05, OPA 07, OPA 11, OPA 15, OPX 06, OPX 09, OPX 11, OPX 15). PCR reactions were set as follows in a final volume of 25 μl: 0.1 mM dNTP, 2 mM $MgCl_2$, 0.5 μM primer, 1× reaction buffer and 0.25 U Taq Polymerase. Samples were submitted to 45 cycles of the following conditions: 1 min at 94°C, 45 s at 35°C and 1.5 min at 72°C. Amplified

Table 1 List of *Coffea* germplasm evaluated

Species	Variety, cultivar or population	Accession code	Number of individuals
Coffea eugenioides S. Moore		C. eug.	5
Coffea kapakata (A. Chev.) Bridson		C. kap.	3
Coffea racemosa Lour.		C. rac.	9
Coffea stenophylla G. Don		C. ste.	6
Coffea canephora Pierre ex A. Froehner	cv. Apoatã	C. can. ap.	3
	var. *kouilouensis* De Wild. pop. 68	C. can. ko68	10
	var. *kouilouensis* pop. 69	C. can. ko69	7
	var. *kouilouensis* pop. 70	C. can. ko70	6
	subvar. *robusta* (L. Linden) A. Chev.	C. can. rob.	6
	pop. *bukobensis*	C. can. buk.	5
	cv. Guarini	C. can. gua.	10
Coffea liberica Bull. ex Hiern	var. *liberica*	C. lib. lib.	9
	var. *dewevrei* (De Wild. et T. Durand) Lebrun pop. *abeokutae*	C. lib. dew. ab.	10
	var. *dewevrei* pop. *dewevrei*	C. lib. dew.	10
	var. *dewevrei* pop. *dibowskii*	C. lib. dew. dy.	7
	var. *dewevrei* pop. *excelsa*	C. lib. dew. ex.	6
	var. *dewevrei* pop. *uganda*	C. lib. dew. ug.	10
Piatã coffee (C387)		Piatã	1
Coffea arabica L.	Commercial cultivars	C. ara. cv (cultivar code)	8
	Mutants	Mar., Pol. and Vol.	10
Coffea heterocalyx Stoff.		C. het. R20P	3
Psilanthus ebracteolatus Hiern		P. ebr.	3
Coffea congensis A. Froehner		C. con.	14
	pop. *bangelan*	C. con. ban.	13
	pop. *uganda*	C. con. uga.	6
		Total	180

fragments were separated according to size on 1.0% agarose gel, stained with ethidium bromide. Gel documentation and fragment size determination were performed by the software Image Master Total Lab (*Pharmacia*). RAPD reactions were repeated two times for all individuals to confirm the reproducibility of this method.

Data analysis

The gels were scored for the presence or absence of only reproducible and clearly amplified fragments. Genetic diversity within species and groups was evaluated by proportion of polymorphic loci (*P*) and Shannon's genetic index (*H′*) (Bussell 1999). Groups consisted of accessions of the same variety (for *C. liberica*), cultivar, variety and population (for *C. canephora*), populations (for *C. congensis* and *C. liberica* var. *dewevrei*) and origin (for *C. arabica*). The Piatã hybrid was excluded of the genetic diversity analysis because it was represented by only one accession. *P* was calculated dividing the number of polymorphic bands by total number of amplified bands in each group. Shannon's genetic index for each RAPD locus was calculated for each group as:

$$H' = - \sum p_i^* \log_2 p_i$$

where p_i is the frequency of the presence or absence of a band in that group.

Following the method of Bussell (1999) the partitioning of genetic variation within and between species was estimated (Analysis 1). The average diversity over all species for each locus (H'_{pop}) and the total diversity in the 179 coffee accessions for each locus (H'_{sp}) were calculated (see Bussell 1999 for details). Then the component of diversity within species (H'_{pop}/H'_{sp}) and the component between species ($G'_{ST} = (H'_{sp}–H'_{pop})/H'_{sp}$) were calculated. Distribution of diversity was also studied within *C. arabica* (Analysis 2), *C. canephora* (Analysis 3), *C. congensis* (Analysis 4) and *C. liberica* (Analysis 5) species. Thus, the component of diversity within (H'_{pop}/H'_{sp}) and between groups (G'_{ST}) for each species was calculated. In addition, the components of diversity within and between the varieties of *C. liberica* (Analysis 6) and within and among groups (or populations) of *C. liberica* var. *dewevrei* (Analysis 7) were calculated. Finally, we considered *C. canephora* and *C. congensis* as a single group and analyzed the partitioning of the diversity within and between species (Analysis 8) and groups (Analysis 9). This was carried out due to the morphological and molecular similarity showed by both species (Moncada and McCouch 2004; Prakash et al. 2005).

Genetic distance among all 180 accessions was estimated as the complement of Jaccard's (1908) coefficient (Link et al. 1995). Cluster analysis was performed using the matrix distance based on the complement of Jaccard's coefficient employing the UPGMA method. Bootstrap analysis (Felsenstein 1985) was performed to evaluate the tree topology reliability for 1000 simulations. These analyses were carried out using the software Treecon (Van de Peer and Watcher 1994). The cophenetic coefficient between the matrix of genetic distance and the dendrogram were computed using the NTSYS-PC package (Rohlf 1998). The dendrogram obtained by NTSYS was the same generated by Treecon. The significance of the cophenetic correlation was tested by the Mantel correspondence test (Mantel 1967).

Since the cluster analysis of all accessions is a very large tree, it was selected three accessions at random from each group (except cultivars of *C. arabica*) in order to visualize the dendrogram. For a scrutinized study of *C. canephora* and *C. congensis* relationships,

a cluster analysis including all accessions of these species was performed.

Results

RAPD markers

Thirteen primers generated a total of 100 fragments. The number of fragments per primer varied from 2 to 16 with a mean of 7.7 fragments per primer. Fragment sizes ranged from 300 to 2240 bp. Five bands were monomorphic for all accessions studied. Polymorphism was not detected among 8 accessions of *C. liberica* var. *liberica*. Some bands were species or group specific.

Genetic diversity

H', H'_{pop}, H'_{sp} and G'_{ST} were averaged over all loci from the 179 accessions, including monomorphic loci, to yield estimates of $\underline{H}'$, $\underline{H}'_{pop}$, $\underline{H}'_{sp}$ and $\underline{G}'_{ST}$ (Tables 2 and 3). The highest value of genetic diversity was found in *C. canephora*. For this species, genetic diversity within species ($\underline{H}'_{sp}$) was of 0.414 and within groups was of $\underline{H}' = 0.343$ for cv. Guarini and $\underline{H}' = 0.248$ for var. *kouilouensis* pop. 68. *C. liberica* var. *dewevrei* also show high genetic diversity ($\underline{H}'_{sp} = 0.356$), with $\underline{H}' = 0.288$ for *abeokutae* population and $\underline{H}' = 0.199$ for *uganda* population. The lowest values were of *C. liberica* var. *liberica*, the autogamous *C. arabica* and *C. heterocalyx* species. Also, *C. kapakata*, the species with the lowest number of individuals in the collection showed low $\underline{H}'$. Values of proportion of polymorphic loci (P) in each group (Table 2) followed the same pattern of Shannon's index.

Genetic relationships

Analysis of diversity with Shannon's index (Table 3) revealed that most RAPD variation was partitioned between, rather than within species ($\underline{G}'_{ST} = 0.713$). However, *C. canephora* and *C. liberica* showed a high genetic diversity within the species ($\underline{H}'_{sp}$) mostly partitioned between the groups and varieties (high

Table 2 Diversity genetic within groups of coffee species from the Coffee Collection of IAC assessed by proportion of polymorphic loci (P) and Shannon's genetic index averaged over all loci ($\underline{H}'$)

Species	Group	P (%)	$\underline{H}'$
Coffea eugenioides		20.2	0.165
Coffea kapakata		4.2	0.040
Coffea racemosa		21.2	0.178
Coffea stenophylla		18.4	0.149
Coffea canephora	Apoatã	18.0	0.168
	kouilouensis 68	35.4	0.248
	kouilouensis 69	24.5	0.201
	kouilouensis 70	25.3	0.204
	robusta	21.2	0.171
	bukobensis	18.6	0.154
	Guarini	45.0	0.343
Coffea liberica	*liberica*	8.2	0.042
	dewevrei pop. abeokutae	38.3	0.288
	dewevrei pop. *dewevrei*	22.6	0.181
	dewevrei pop. dibowskii	21.6	0.172
	dewevrei pop. excelsa	21.6	0.175
	dewevrei pop. uganda	23.9	0.199
Piatã coffee (hybrid)		–	–
Coffea arabica	Commercial cultivars	11.5	0.066
	Mutants	2.3	0.016
Coffea heterocalyx		2.9	0.029
Psilanthus ebracteolatus		10.4	0.096
Coffea congensis		20.0	0.119
	bangelan	21.3	0.152
	uganda	13.0	0.091

Table 3 Partitioning of genetic diversity generate by 100 RAPD loci into within and between group components for nine species of genus *Coffea* and one species of genus *Psilanthus* included in the Coffee Collection of IAC

Analyzed accessions	$\underline{H}'_{pop}$	$\underline{H}'_{sp}$	$\underline{G}'_{ST}$
(1) All species	0.199	0.686	0.713
(2) *Coffea arabica*	0.043	0.126	0.391
(3) *Coffea canephora*	0.215	0.414	0.473
(4) *Coffea congensis*	0.133	0.218	0.338
(5) *Coffea liberica* (groups)	0.174	0.380	0.516
(6) *Coffea liberica* (varieties)	0.199	0.380	0.459
(7) *Coffea liberica* var. *dewevrei* (groups)	0.202	0.356	0.428
(8) *C. canephora* and *C. congensis* (species)	0.350	0.442	0.186
(9) *C. canephora* and *C. congensis* (groups)	0.205	0.442	0.529

$\underline{H}'_{pop}$, $\underline{H}'_{sp}$ and $\underline{H}'_{ST}$ are the average per locus values calculated over all loci. Numbers in brackets refer to the grouping of accessions analyzed as described in the Material and Methods

identified a level of genetic distance between species ranging from 0 (*C. canephora* var. *kouilouensis* 69-13 and *C. congensis* pop. *uganda* 01) to 0.70 (*C. racemosa* and all other species) (Fig. 1). The dendrogram showed two major groups (Fig. 1). The first included the accessions of *C. racemosa* species clustered with accessions of *P. ebractelatus*. The second group encompassed basically all other species clustered in two subgroups. Two clusters formed the first subgroup: the first included *C. heterocalyx* grouped with Piatã and all related *C. arabica* accessions. The second involved *C. eugenioides* jointed with *C. kapakata* after clustered with *C. stenophylla*. The second subgroup comprised a cluster of the *C. liberica* accessions jointed with *C. congensis* and *C. canephora*, which were closely related to each other and formed another cluster. However, there were a few groups clearly distinguished and only some clusters were supported by high bootstrap values.

Cluster analysis of all accessions of *C. canephora* and *C. congensis* confirmed $\underline{H}'$ and $\underline{G}'_{ST}$ results (Fig. 2, Tables 2 and 3). Only accessions from the *C. congensis* group, *kouilouensis* variety and the *bukobensis* population could be clearly distinguished, but these grouping associations were not supported by high bootstrap values and the *C. congensis* cluster included some accessions of *C. canephora*.

$\underline{G}'_{ST}$). *C. congensis* and *C. arabica* also showed the genetic diversity distributed between the groups, but with lower levels of genetic diversity. Special analysis of *C. canephora* and *C. congensis* accessions showed that diversity among groups of both species ($\underline{G}'_{ST} = 0.529$) is higher than diversity between the species ($\underline{G}'_{ST} = 0.186$).

Cluster analysis corroborated results of Shannon's genetic index (Fig. 1). The cophenetic correlation ($r = 0.84$) indicated that the clustering of genotypes demonstrated in the dendrogram accurately represented the estimates of genetic distance among accessions. Jaccard's coefficient

Fig. 1 Dendrogram of *Coffea* accessions (3 accessions per group) based on Jaccard genetic distance obtained from RAPD markers using the UPGMA method. Numbers (%) on the branches correspond to bootstrap values above 50% (1,000 replications). Letters indicate the geographical origin of accessions: C (Central Africa), WC (West and Central Africa) and E (East Africa)
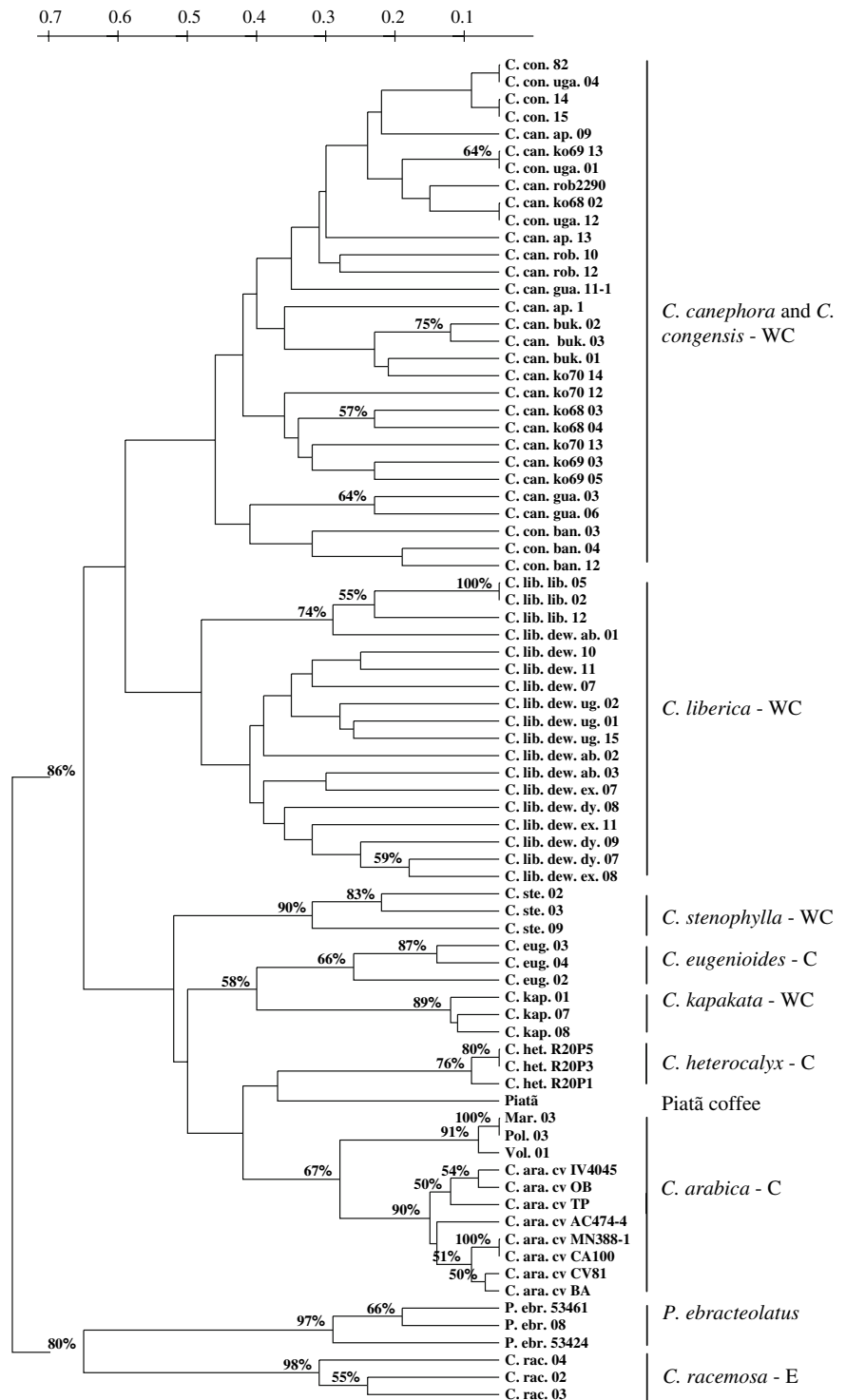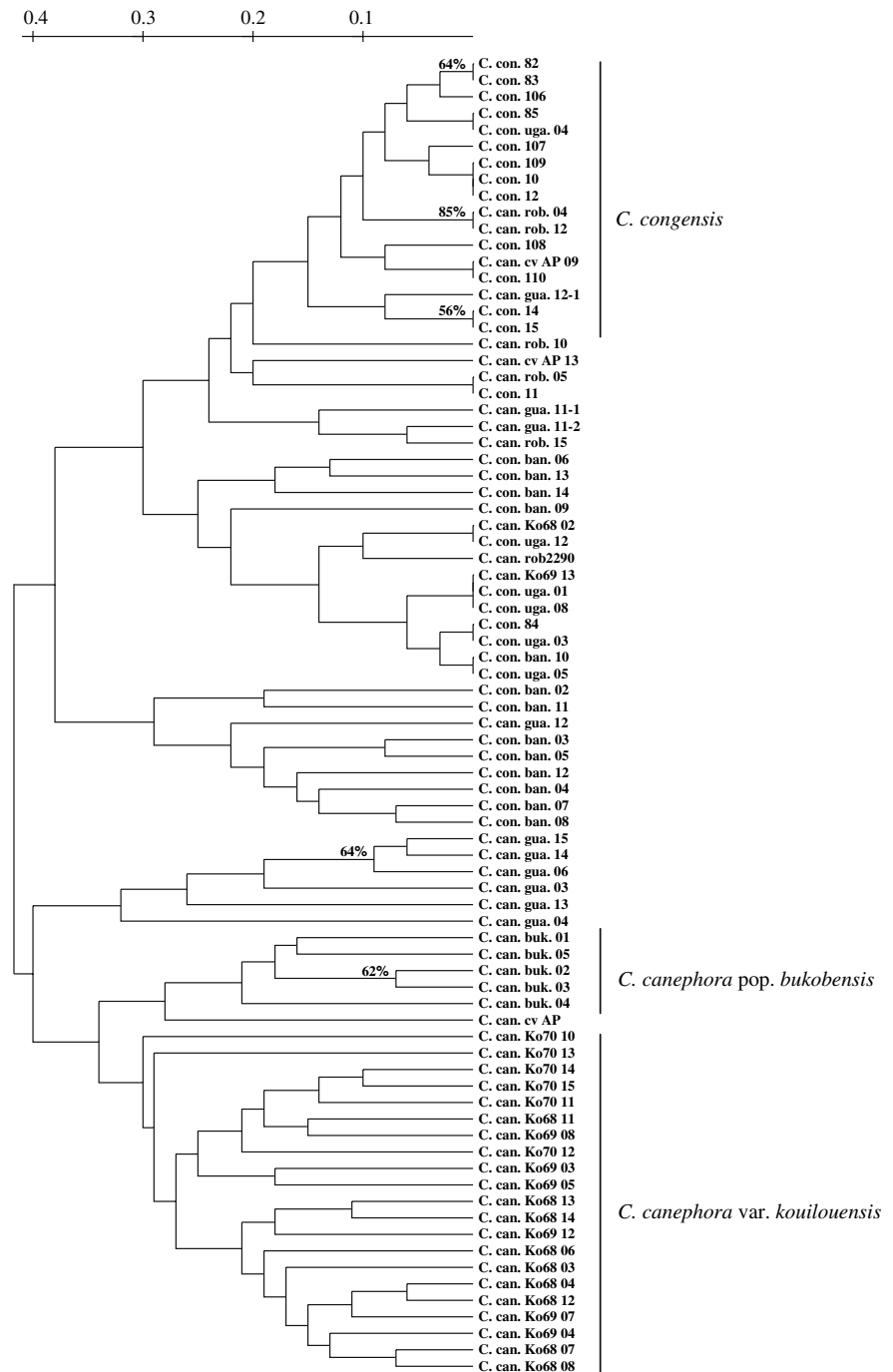
Fig. 2 Dendrogram of
Coffea canephora and
Coffea congensis accessions
based on Jaccard genetic
distance obtained from
RAPD markers using the
UPGMA method. Numbers
(%) on the branches
correspond to bootstrap
values above 50% (1,000
replications)

Discussion

Genetic diversity

The level of polymorphism, the genetic diversity evaluated through Shannon's index and the genetic distances values observed in this study were in the range of previous studies carried out in coffee species (Lashermes et al. 1993; Ruas et al. 2000; Aga et al. 2003; Moncada and McCouch 2004; Silvestrini et al. 2007). Using SSR markers, Silvestrini et al. (2007) showed slightly higher H' values than found here for

C. racemosa and *C. eugenioides* accessions and lower values for cultivars of *C. arabica*. These values were obtained even with a lower number of accessions of *C. racemosa* evaluated and with a greater number of arabica cultivars sampled from the same genebank of IAC. Despite the fact that RAPD and SSR have grouped the *C. arabica* cultivars similarly, they have showed different genetic index (Maluf et al. 2005). These results suggested that these molecular markers sample genetic diversity in different ways in coffee.

Coffea canephora and *C. liberica* var. *dewevrei* accessions showed a fairly high genetic diversity, and *C. arabica* cultivars exhibited a very low $\underline{H}'$ values. These results were not unexpected since the analyzed diploid species such as *C. canephora* are outcrossed and come from different geographical origins, while cultivated *C. arabica* have a well-reported narrow genetic basis (Lashermes et al. 1996; Moncada and McCouch 2004; Silvestrini et al. 2007).

Although *C. congensis*, *C. stenophylla*, *C. racemosa* and *C. eugenioides* showed intermediate values of $\underline{H}'_{sp}$ and $\underline{H}'$, genetic analyses indicated that there is a variation within and between species which could be useful as source of valuable genes for introgression of desired characteristics in commercial coffee cultivars. From the breeding standpoint, this is quite important if we consider the potential genes locked up in plant collections that can be identified and used in development of new plant varieties by advanced breeding methods.

Even with a reasonable number of individuals studied, *C. liberica* var. *liberica* showed very low genetic diversity. Driven by such unexpected results, a careful documentary survey was carried out in the original archives of the *C. liberica* accessions revealing that among the nine studied plants, two were half-sibs and the other seven clones of a single tree originated by cleftgrafts in the 40's. For this group, inclusion of new accessions and exclusion of duplicates are necessary to increase the effective population size and genetic diversity. Inclusion of new accessions could also be important for *C. kapakata*, *C. heterocalyx* and *P. ebracteolatus* to enlarge their genetic basis in the collection.

Genetic relationships

Most species could be distinguished in the cluster analysis (Fig. 1) corroborating the high diversity found between species. Also, there was a strong genetic structure within studied species, all of them showing high values of $\underline{G}'_{ST}$. This genetic structure could be explained in part by artificial selection and genetic drift due to dramatic reduction of population size during formation of collection. Some materials are result of intensive process of selection like the *C. canephora* cultivar Apoatã and *C. arabica* cultivars (Fazuoli et al. 2002; Fazuoli et al. 2007). It is noteworthy, however, that there was no separation between cultivated and noncultivated accessions of *C. canephora* (Fig. 2). The high genetic diversity within the groups, mainly the cultivars, could explain the lack of arrangement of the accessions in clusters. In agreement with Poncet et al. (2004), this result also suggested that domestication process in *C. canephora* might have been independently initiated from various wild populations in different regions of its natural geographical distribution.

Groups of *C. arabica* exhibited low within ($\underline{H}'$), but high between ($\underline{G}'_{ST}$) genetic variability, confirming the divergence between mutant and cultivated plants found in the cluster analysis. These results revealed a variation in mutants group, which can be explored for improvement of arabica cultivars. Yet, for Maragogipe plants (Mar.), this use has not been considered viable through conventional improvement methods due to the low productivity of the segregant progenies (Monaco 1960; Carvalho et al. 1991).

Interestingly, *C. congensis* and *C. canephora* accessions analyzed together showed lower values of $\underline{G}'_{ST}$ between species than among groups. Therefore there was more genetic variability among groups of both species than between the species. This was also observed in the cluster analysis (Figs. 1 and 2) where accessions of both species were very closely related to each other. Similar relationships between these species were found by Moncada and McCouch (2004) and Prakash et al. (2005). This high similarity plus the high production of fertile hybrids between the species (Louarn 1993) agreed with phylogenetics studies (Cros 1996; Lashermes et al. 1997; Cros et al. 1998). According to Cros (1996), *C. canephora* and *C. congensis* are clearly defined as different species by botanic, biochemistry and isoenzymatic rather than by phylogenetic characteristics.

A different result was observed in *C. liberica* species, which includes *liberica* and *dewevrei* varieties (Bridson and Verdcourt 1988). The species

showed a high $\underline{G}'_{ST}$ value when evaluated considering the varieties ($\underline{G}'_{ST} = 0.459$). Partitioning of diversity between groups demonstrated an increase of $\underline{G}'_{ST}$ values from 0.428 to 0.516 due to the inclusion of *C. liberica* var. *liberica* (Table 3). Also, cluster analysis showed a significant genetic distance between both varieties (Fig. 1). This divergence corroborated the results of N'Diaye et al. (2005) that pointed out that *C. liberica* var. *liberica* and *C. liberica* var. *dewevrei* exhibited a higher level of differentiation than that commonly observed among botanical varieties. The same is true for flower and fruit morphology and structure of chromosomes attached to the nucleolus investigated by Carvalho (1991) and Pinto-Maglio and Cruz (1987), respectively.

In general, cluster analysis of *Coffea* accessions from IAC (Fig. 1) showed similar relationships between species as those observed by phylogenetic (Lashermes et al. 1997; Cros et al. 1998) and molecular marker studies (Lashermes et al. 1993; Ruas et al. 2000; Moncada and McCouch 2004; Silvestrini et al. 2007). In other agronomic studies, accessions from the IAC *Coffea* Collection were evaluated regarding intrinsic coffee bean outturn and fruit chemical composition (Gaspari-Pezzopane et al. 2004; Aguiar et al. 2005). The resulting groups (Aguiar et al. 2005) were very similar to that observed with RAPD analysis. Three chemically distinct group were observed through PCA analysis: group 1 included *C. canephora*, *C. congensis*, *C. stenophylla* and *C. racemosa*; a second group including *C. eugenioides* and *C. kapakata*; and *C. liberica* as the only species in group 3. Overall, RAPD analyses confirmed the genetic distribution of diversity of both studies and the species relationships of genus *Coffea* evaluated. The major difference is the presence of *C. racemosa* among *C. canephora* chemical group. This result is interesting because although *C. racemosa* is the most genetically diverse species of the *Coffea* genera, it is considered closely related to *C. arabica* and *C. canephora*, from the botanical and agronomical point of view, exhibiting fair rates of artificial crossing with these species (Medina-Filho et al. 2007). The similar results observed therefore reinforce the use of RAPD method as an efficient tool for evaluating overall genetic diversity in *Coffea* collections. Also, they indicate that both agronomic and molecular

descriptors can be used for the characterization of *Coffea* diversity.

In conclusion, this study represents the first report of an *ex-situ Coffea* Germplasm Collection where parameters necessary for genetic diversity evaluation were determined, and may therefore serve as reference for comparisons with other similar collections. Also, the genetic relationship among accessions evaluated in this study not only corroborated previous classifications but also demonstrated that the IAC *Coffea* Collection is representative of the phylogenetic structure observed in the genera. Genetic diversity analysis showed a fairly high and structured genetic diversity, although some groups and species are still not very well represented, such as *C. liberica* var. *liberica* and *C. kapakata*. Variation was partitioned between, rather than within species, but some species have exhibited high $\underline{H}'_{sp}$ and $\underline{G}'_{ST}$ values and high genetic diversity was also found within most groups. This distribution of diversity between species, between groups within species and within groups indicated that there is not a single collecting strategy to represent the entire variability of the genus *Coffea* in germplasm collections. Good sampling of species as well as varieties, cultivars and individuals are necessary. Also, these results showed the importance of these germplasm resources as source of variability for coffee breeding programs.

References

Aga E, Bryngelsson T, Bekele E, Salomon B (2003) Genetic diversity of forest arabica coffee (*Coffea arabica* L.) in Ethiopia as revealed by random amplified polymorphic DNA (RAPD) analysis. Hereditas 138:36–46

Aguiar ATE, Fazuoli LC, Salva TJG, Favarin JL (2005) Chemical diversity in coffee species of genebank of Instituto Agronômico do Estado de São Paulo. Crop Breed Appl Biotechnol 5:460–466

Agwanda CO, Lashermes P, Trouslot P, Combes MC, Charrier A (1997) Identification of RAPD markers for resistance to coffee berry disease, *Colletotrichum kahawae*, in arabica coffee. Euphytica 97:241–248

Bridson DM, Verdcourt B (1988) Rubiaceae (Part 2). In: Polhill RM (ed) Flora of tropical East Africa. Balkema, Rotterdam/Brookfield, pp 703–723

Bussel JD (1999) The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). Mol Ecol 8: 775–789

Carvalho A (1991) Letter addressed to Dr. D. Bridson on January 10th, 1991. (Proc. L. Gen. 02/91 AC.efo. 10.01.91)

Carvalho A, Medina-Filho HP, Fazuoli LC, Guerreiro-Filho O, Lima MMA (1991) Aspectos genéticos do cafeeiro (Genetic Aspects of the coffee tree). Rev Bras Genét 14:135–183

Cros J (1996) Implications phylogénétiques des variations de l'ADN chloroplastique chez les caféiers (genres *Coffea* L. et *Psilanthus* Hook. f.). Thèses et documents microfichés, 147. Université Monpellier II, Orstom éditions, Monpellier

Cros J, Combes MC, Trouslot P, Anthony F, Hamon S, Charrier A, Lashermes P (1998) Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. Mol Phyl Evol 9:109–117

Fazuoli LC, Medina-Filho HP, Gonçalves W, Guerreiro-Filho O, Silvarolla MB (2002) Melhoramento do cafeeiro: Variedades do tipo arábica obtidas no Instituto Agronômico de Campinas. In: Zambolim L (ed) O estado da arte de tecnologias na produção de café. Viçosa, pp 163–251

Fazuoli LC, Guerreiro-Filho O, Silvarolla MB, Medina-Filho HP, Bordignon R, Braghini MT (2007) Pré-melhoramento do cafeeiro. In: Fávero AP, Guimarães EP, Faleiro FG, Fonseca MAJ, Lopes MA, Folle SM (eds) Pré-Melhoramento de Plantas – Estado da Arte e Experiências de Sucesso (in press)

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Gaspari-Pezzopane C, Medina-Filho HP, Bordignon R (2004) Variabilidade genética do rendimento intrínseco de grãos em germoplasma de *Coffea*. Bragantia 63(1):39–54

Jaccard P (1908) Nouvelle recherches sur la distribution florale. Bull Soc Vaud Sci Nat 44:223–270

Lashermes P, Cros J, Marmey P, Charrier A (1993) Use of random amplified DNA markers to analyze genetic variability and relationships of *Coffea* species. Genet Resour Crop Evol 40:91–93

Lashermes P, Trouslot P, Anthony F, Combes MC, Charrier A (1996) Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea arabica*. Euphytica 87:59–64

Lashermes P, Combes MC, Trouslot P, Charrier A (1997) Phylogenetic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. Theor Appl Genet 94:947–955

Link W, Dixkens C, Singh M, Schwall M, Melchinger AE (1995) Genetic diversity in European and Mediterranean faba bean germ plasm revealed by RAPD markers. Theor Appl Genet 90:27–32

Louarn J (1993) Structure génétique des caféiers africains diploies basée sur la fertilité des hybrides interspécifiques. In: 15th International Scientific Colloquium on Coffee, ASIC, Montpellier, pp 243–252

Maluf MP, Silvestrini M, Ruggiero LMC, Guerreiro-Filho O, Colombo C (2005) Genetic diversity of cultivated *Coffea arabica* inbred lines assessed by RAPD, AFLP and SSR marker systems. Sci Agric 62(4):366–373

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Medina-Filho HP, Maluf MP, Bordignon R, Guerreiro-Filho O Fazuoli LC (2007) Traditional breeding and modern genetics: a summary of tools and developments to exploit biodiversity for the benefit of the coffee agroindustrial chain. Acta Hort 745:351–368

Monaco LC (1960) Melhoramento do cafeeiro XVII – Seleção do café Maragogipe AD. Bragantia 19:459–492

Moncada P, McCouch S (2004) Simple sequence repeat diversity in diploid and tetraploid *Coffea* species. Genome 47:501–509

N'Diaye A, Poncet V, Louarn J, Hamon S, Noirot M (2005) Genetic differentiation between *Coffea liberica* var. *liberica* and *C. liberica* var. *dewevrei* and comparison with *C. canephora*. Pl Syst Evol 253:95–104

Orozco-Castillo C, Chalmers KJ, Waugh R, Powel W (1994) Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. Theor Appl Genet 87:934–940

Paillard M, Lashermes P, Pétiard V (1996) Construction of a molecular linkage map in coffee. Theor Appl Genet 93:41–47

Pinto-Maglio CAF, Cruz ND (1987) Pachytene chromosome morphology in *Coffea* L. I – Nucleolar chromosomes. Caryologia 40:7–23

Poncet V, Hamon P, Minier J, Carasco C, Hamon S, Noirot M (2004) SSR cross-amplification and variation within coffee trees (*Coffea* spp.). Genome 47:1071–1081

Prakash NS, Combes MC, Dussert S, Naveen S, Lashermes P (2005) Analysis of genetic diversity in Indian robusta coffee genepool (*Coffea canephora*) in comparison with a representative core collection using SSRs and AFLPs. Genet Resour Crop Evol 52:333–343

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding – RAPDs, microsatellites and machines. Trends Genet 9(8):275–280

Ram AS, Sreenath HL (2000) Genetic fingerprinting of coffee leaf rust differentials with RAPD markers. In: Sera T, Soccol CR, Pandey A, Roussos S (eds) Coffee biotechnology and quality. Proceedings of the 3rd international seminar on biotechnology in the coffee agroindustry, Londrina, Brazil. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 197–208

Rohlf FJ (1998) NTSYS-pc, Numerical taxonomy and multivariate analysis system, version 2.02i. Exeter Publishing, Setauket, New York

Ruas PM, Diniz LEC, Ruas CF, Sera T (2000) Genetic polymorphism in species and hybrids of *Coffea* revealed by RAPD. In: Sera T, Soccol CR, Pandey A, Roussos S (eds) Coffee biotechnology and quality. Proceedings of the 3rd international seminar on biotechnology in the coffee agroindustry, Londrina, Brazil. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 187–195

Silvarolla MB, Mazzafera P, Medina-Filho HP, Fazuoli LC (1999) Ploidy level and caffeine content in leaves of *Coffea*. Sci Agric 56(3):661–663

Silvestrini M, Junqueira MG, Favarin AC, Guerreiro-Filho O, Maluf MP, Silvarolla MB, Colombo CA (2007) Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. Genet Resour Crop Evol 54:1367–1379

Van De Peer Y, Wachter R (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Appl Biosci 10:569–570