



A Methodological Approach for Prioritization and Rationalization of Field Genebank Accessions of Coffee Genetic Resources: A Case Study of CATIE International Coffee Collection, Costa Rica

Mohammad Ehsan Dulloo^{1*}, William Solano², Dominique Dessauw^{2,3,4}, Carlos Astorga² and Luigi Guarino⁵

¹ Bioversity International, Rome, Italy, ² Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Turrialba, Costa Rica, ³ Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Amélioration Génétique et Adaptation des Plantes Méditerranéennes et Tropicales Institute, Montpellier, France, ⁴ Amélioration Génétique et Adaptation des Plantes Méditerranéennes et Tropicales Institute, University of Montpellier, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Institut National de Recherche pour l'agriculture, l'alimentation et l'environnement, Institute Agro, Montpellier, France, ⁵ Global Crop Diversity Trust, Bonn, Germany

OPEN ACCESS

Edited by:

Paulo Mazzafera,
State University of Campinas, Brazil

Reviewed by:

Douglas S. Domingues,
São Paulo State University, Brazil
Andreas Wilhelm Ebert,
World Vegetable Center, Taiwan

*Correspondence:

Mohammad Ehsan Dulloo
e.dulloo@cgiar.org

Specialty section:

This article was submitted to
Crop Biology and Sustainability,
a section of the journal
Frontiers in Sustainable Food Systems

Received: 15 September 2021

Accepted: 11 November 2021

Published: 17 December 2021

Citation:

Dulloo ME, Solano W, Dessauw D, Astorga C and Guarino L (2021) A Methodological Approach for Prioritization and Rationalization of Field Genebank Accessions of Coffee Genetic Resources: A Case Study of CATIE International Coffee Collection, Costa Rica. *Front. Sustain. Food Syst.* 5:777415. doi: 10.3389/fsufs.2021.777415

Good management of coffee collections is important because they ensure long-term availability of germplasm to guarantee the sustainability of coffee value chain. The conservation of coffee genetic resources is essential to provide the raw materials for breeding and improvement of the crop. Many genetic resources of wild arabica coffee have been collected in the second half of the 20th century by several international collecting missions, including by Food and Agriculture Organization of the United Nations, ORSTOM (now IRD), Centre de coopération internationale en recherche agronomique pour le développement (CIRAD), and IPGRI (now Bioversity International), and are conserved in several national genebanks and at the CATIE International Coffee Collection (CICC) in Turrialba, Costa Rica. Over the past decades, many of the original accessions of the CICC have become threatened due to age, pests and diseases, inadequate management, and waterlogging. There is thus an urgent need to rejuvenate and rationalize the collection to ensure the long-term maintenance of the genetic diversity of the original accessions. Here we present the methodological approach we followed to carry out an in-depth assessment of the status of the coffee collection at CATIE and to prioritize accession-specific actions for the rationalization of the collection. This can be used as a model for other collections to assess and rationalize their own field genebank, with a view to improving their management in the most cost-effective way. The study identified many discrepancies between the number of accessions in the field and genebank records and revealed that 80 accessions have been lost from the collection since 2014 and that approximately 80% of the accessions were threatened and in need of intervention. Furthermore, the in-depth study identified the most diverse and valued accessions for the rationalization of the CICC field genebank and those that are in urgent need of safety duplication.

Keywords: coffee genetic resources, *Coffea arabica*, *ex situ* conservation, field genebank, prioritization

INTRODUCTION

The long-term *ex situ* conservation of coffee genetic resources faces many challenges. It has long been known that seeds of coffee species cannot be conserved under the standard conservation conditions in seed banks for extended period because coffee seeds are only partially tolerant to desiccation and are cold sensitive (Hong and Ellis, 1995). Other techniques for conserving coffee genetic resources *ex situ* have been developed (Dulloo et al., 1998; Engelmann and Dulloo, 2007), including cryopreservation (storage at liquid nitrogen temperature, -196°C) of *Coffea arabica* L. seeds (Dussert et al., 2001) and *in vitro* slow growth and cryopreservation for medium- to long-term conservation of zygotic or somatic embryos, apices, and buds (Dussert et al., 1997). Other options of *ex situ* conservation include pollen storage under vacuum (Walyaro and Van der Vossen, 1977) and DNA storage (Adams and Adams, 1991). All of these methods have their respective advantages and disadvantages. Although *in situ* protection of *Coffea* species and varieties, both in the wild and on farms, is a potentially important conservation approach, it has not received sufficient attention or resources. Consequently, the *ex situ* conservation of coffee genetic resources is mostly done as live plants in field genebanks (Vega et al., 2008; Dulloo et al., 2009). However, this mode of conservation suffers from many drawbacks and is vulnerable to many technical, management, and economic factors, including pest and disease outbreaks, extreme weather conditions, cyclones, fire, suboptimal ecological conditions, land availability, high labor requirements, and high costs (Dulloo et al., 2001, 2009; Bramel et al., 2017). In the long term, the maintenance of field genebanks often becomes a financial burden for institutions, and this may result in poorly maintained collections and the loss of accessions and thus genetic diversity (Dulloo et al., 2009; Bramel et al., 2017).

Coffee field genebanks were established in the second half of the 20th century following major international collecting missions undertaken by the FAO (Ferne et al., 1968), ORSTOM (Guillaumet and Hallé, 1978), CIRAD, the Museum of Natural History of Paris, and IPGRI (Bramel et al., 2017). Countries that harbor important diversity of wild *Coffea* species (Ethiopia, Madagascar, Cote d'Ivoire) and where there is a breeding program have also established their own national coffee field genebanks. Many of the samples collected by the international missions have also been sent to several national field genebank around the world (e.g., India, Ethiopia, Tanzania, Colombia, and Peru) (Ferne et al., 1968). Inventory of coffee field genebanks throughout the world has been undertaken in the past by several authors (Bettencourt and Konopka, 1988; FAO-WIEW Database (cited in Bramwell et al., 2017); Eira et al., 2007; Labouisse et al., 2008; Dulloo et al., 2009; and Phiri, 2013), but the latest inventory reveals a total of more than 21,000 accessions being conserved in field genebanks globally (Bramel et al., 2017).

CATIE International Coffee Collection (CICC), established in 1949 in Turrialba, Costa Rica, is one of the world's largest collections of *C. arabica* and a few diploid coffee species, with 1,960 accessions, which includes samples of the historic collecting missions by FAO, ORSTOM, and IPGRI (Bramel et al., 2017). It is considered as the most important coffee collection in the

world in the public domain, given its unique status by virtue of the agreement signed between CATIE and the International Treaty on Plant Genetic Resources for Food and Agriculture under its Article 15 (FAO, 2009). As such, the collection is accessible to users for research, breeding, and training under its facilitated access and benefit sharing arrangements. The collection was identified by the Global Coffee Conservation Strategy, prepared by the Crop Trust and World Coffee Research (Bramel et al., 2017), as essential for the long-term preservation of coffee diversity, one of the so-called "Origin Collections" outside the African continent that would meet the eligibility criteria to receive resources from the Crop Trust's Endowment Fund. In fact, the wild genotypes in the CICC have been used extensively in regional breeding programs in collaboration with different partners to produce highly productive and disease-resistant coffee varieties (Bramel et al., 2017). For example, the hybrid Nemaya was developed from two accessions of *Coffea canephora*, with resistance to nematodes and having a strong root system, being used as a rootstock for regenerating *C. arabica*. Another accession, the Geisha variety, has been shown to have resistance to coffee leaf rust and is widely used in breeding programs (Bramel et al., 2017). This variety was distributed in the 1960s to Boquete area in Panama, where it developed excellent organoleptic characteristics, and became recognized worldwide for its high quality, allowing this variety to reach record figures in international auctions.

The CICC has also been suffering from the loss of some of its accessions, principally due to aging trees, waterlogging, and diseases, principally leaf rust and American Leaf spot (Bramel et al., 2017). Furthermore, given the size of the collection, there has been a lack of adequate operational funds to purchase farm inputs such as fertilizers and pesticides. To ensure long-term sustainable support for management of the historic collection of CICC, the Global Crop Diversity Trust commissioned a study in 2019 to carry out an in-depth assessment of each accession in the CICC with a view to rationalizing the collection and defining strategic conservation actions (Dulloo, 2020a). As part of this study, an accession-by-accession prioritization methodology was developed to help determine a set of the most important accessions in the collection that would be part of the rationalized field collection. The aim of this article is to describe the accession by accession prioritization methodology that has been applied to CICC, to serve as a model that other field genebanks could use to better manage their own collection in a most cost-effective manner.

METHODS

The prioritization process involved carrying out five main steps (Figure 1), as follows:

Step 1: Gather, Collate, and Analyze Information on Each Accession

The first step involves gathering and analyzing all management, passport, characterization, and evaluation data for each accession contained in the field genebank that is held on record and

5 Steps to Prioritizing Accessions in a Field Collection

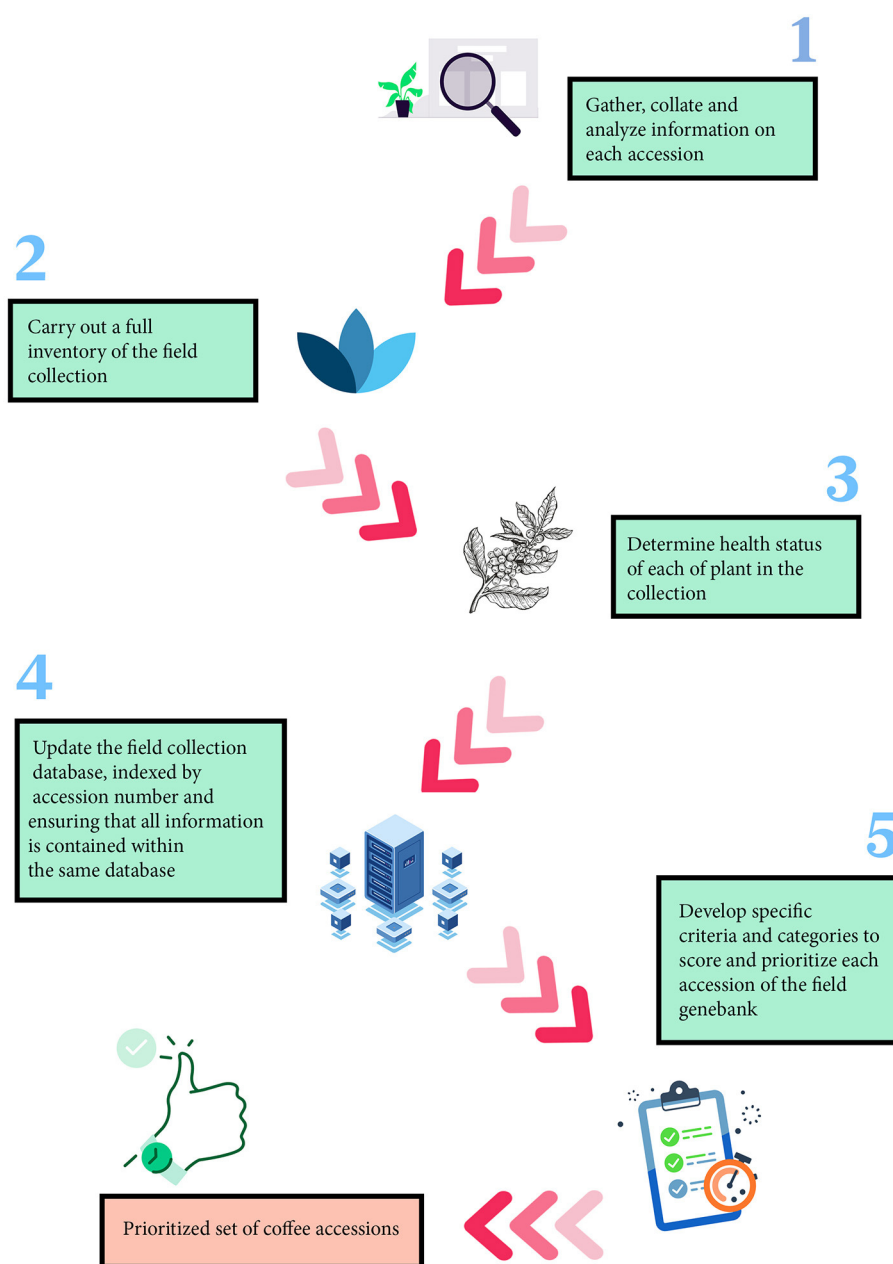


FIGURE 1 | Five steps to prioritizing accessions in a field collection.

identifying any gaps in information on each accession. Often, such information is scattered in different paper and computer files, and it is important that all the information is combined in one database to facilitate analysis. This step also involves carrying out a thorough literature review to gather information about the origins of the accessions and to identify occurrences of the same material in other field genebank, both within and outside the country, to ascertain which accessions may be duplicated elsewhere.

In the CICC, the relevant information was extracted from two principal databases: (i) CATIE's catalog database that contains the passport information (Introduction Book), to verify taxonomy information, varietal name, accession number, specific observations on the accession, its origin, and date of introduction; and (ii) the separate CAFE-BASE database, which also has passport information, but also comprises field inventories and evaluation data. In addition, the results of a 2019 inventory and individual health status assessment of plants

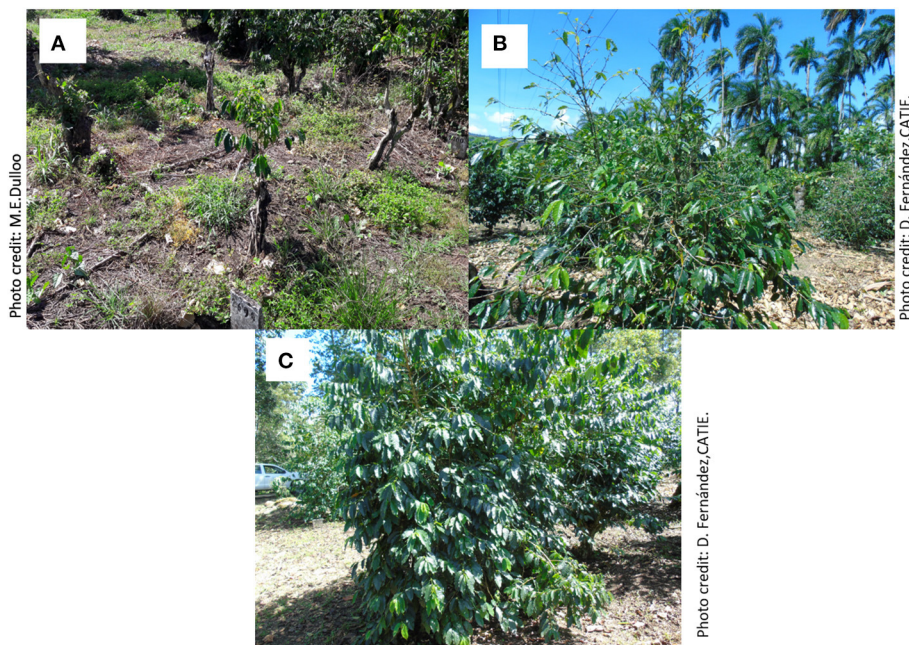


FIGURE 2 | Photographs showing health status of coffee plant in field genebank. (A), Critical; (B), poor; (C), healthy.

of each accession (see below) were held in different files. All the information was brought together in one database to enable a full analysis of the data available for all the accessions of the collection.

Step 2: Carry Out a Full Inventory of the Field Collection

The next step is to undertake a full inventory of the field collection. This involves the physical counting of the number of trees alive for each accession in the field. For each accession extant in the field, the following information should be recorded.

- Accession number.
- Description of the accession.
- The precise location of the accession, depending on how the collection is organized in the field (e.g., section, blocks, lines, etc.).
- Number of plants still alive for the accession: This should be further subdivided as the number of original plants and number of plants that have been regenerated.

A full inventory of the plants within each accession of the CICC was undertaken by the CATIE genebank staff in 2019. Prior inventories of the collection were done in 2008 and 2014.

Step 3: Determine Health Status of Each of Plant in the Collection

It is important that the health status of each plant in the accession be assessed. This involves making a physical observation of the plant in the field and classifying it into one of three categories: critical, poor, and healthy (see criteria and categories below).

Figure 2 illustrates visually the status of coffee plants that are in three health categories. This step is best done at the same time as step 2 above.

Step 4: Update the Field Collection Database, Indexed by Accession Number and Ensuring That All Information Is Contained Within the Same Database

In CICC, all information containing the accession level data on field genebank records in the CATIE catalog, CASE-BASE, inventory files, characterization, and evaluation were analyzed by CATIE information technology specialists who compiled and organized all the information into a single Excel document, using the accession number as the common denominator. This work required reviewing and debugging the information, extracting the data and generating a consolidated field genebank record database, and using unification programming algorithms. In addition, a full bibliographic search for publications containing characterization of coffee from CICC was undertaken from the CATIE's ORTON Library database, as well as from PROMECAFE Network. The extracted information, including 31 titles from ORTON Library and 48 titles from PROMECAFE Network, was compiled in Excel sheets and integrated with the new consolidated field genebank record database.

Step 5: Develop Specific Criteria and Categories to Score and Prioritize Each Accession of the Field Genebank

To prioritize a field genebank collection for appropriate actions, a set of criteria and categories and a scoring mechanism should be developed. It may be necessary to prioritize the criteria and assign

TABLE 1 | Score table for total number of plants and number of original plants in CICC collection.

| No of plants within an accession | Score for total number of plants | Score for number of original plants |
|----------------------------------|----------------------------------|-------------------------------------|
| 1 | 10 | 5 |
| 2 | 8 | 4 |
| 3-4 | 6 | 3 |
| 5-6 | 4 | 2 |
| >6 | 2 | 1 |

different weights to them according to their priority. Each field genebank may develop its own criteria and categories and scoring mechanism depending on the type of crop and the local context.

In the case of CICC, the specific criteria and categories were developed in consultation with the genebank staff and validated by recognized genetic resources experts. Five criteria (type of genetic resources, threat, uniqueness, safety duplication, and use) were identified for prioritizing the accessions of the collection. The “use” criterion could not be used because of the lack of sufficient available information. For each of the criteria, categories were defined and given a score between 1 and 20, depending on their priority. All the criteria have the same maximum score of 20.

Type of Genetic Resources

This criterion identifies the different types of genetic materials that are conserved in the CICC. This criterion is regarded as being the most important and was used to categorize accessions according to whether they were “wild materials” (mostly derived from the historic international collecting missions from Southwestern region of Ethiopia), cultivated varieties, or breeding /experimental materials. The information was derived from the passport data file held in CATIE. Each category was prioritized as follows:

1. Wild materials: 20 points.
2. Cultivated varieties: 10 points.
3. Breeding lines and experimental materials: 5 points.

Threat

The threat criterion quantifies the risk of losing an accession from the collection and is measured by a score combining the current number of plants for each accession and their observed health status. A score is given for the total number of plants for the accession and for the number of original plants (meaning the plants first introduced in the collection) according to the Score Table (**Table 1**). In CATIE, generally an accession is represented by eight plants in the collection (but sometimes more and sometimes less). An accession starts becoming threatened and at risk of loss when it declines below six plants. Thus, the lower the number of plants in the accession, the more threatened it would be, and the higher a high score it received. In addition, we gave extra points for the number of original plants, as we considered

TABLE 2 | Score range for different threat categories.

| Score range | Threat status | Color code |
|-------------|-----------------------|------------|
| 0 | Extinct | EX |
| 18–20 | Critically endangered | CE |
| 15–17 | Endangered | EN |
| 10 to <15 | Threatened | TR |
| 5 to <10 | Vulnerable | VU |
| 1 to <5 | Not threatened | NT |

these plants as being the “most original plants” and had a higher value compared with others that had been regenerated (**Table 1**).

The health status of the accession was defined by three states, namely, critical, poor, and healthy. Each plant (irrespective of whether it was an original or regenerated plant) was assessed and given a score of 5 points for critical, 3 points for poor, and 1 point for healthy. When there were more than six plants for an accession, only the best six plants were scored. For example, if an accession had nine plants with 4 “healthy,” 3 “poor,” and 2 “critical,” we scored the 4 healthy and 2 poor plants. If an accession had nine plants with 3 “healthy,” 2 “poor,” and 4 “critical,” then we would score 3 “healthy,” 2 “poor,” and 1 “critical.” The mean value for the best six plants was then taken as the health status of the accession as a whole (see **Supplementary Table S1** for more detailed explanation on how to calculate the health status score).

The Total Threat Score was then computed as the sum of the individual scores for total plants, original plants, and health scores for each accession. They were sorted with highest score being more threatened, and we categorized accessions in five groups and named the categories as critically endangered (CE), endangered (EN), threatened (TR), vulnerable (VU), and not threatened (NT), based on the score ranges in **Table 2**. In addition, if an accession was found to have been lost from the collection, it was placed in the extinct category and received a score of 0.

Uniqueness

With this criterion, we were trying to estimate how distinctive an accession was within the collection. Uniqueness was measured by counting the total number of accessions collected from a locality, as per the locality name given in the passport data. In the case of CICC, we were able to apply this criterion only to the wild-derived accession as their precise location is well documented by collecting missions of FAO 1964/65 (Fernie et al., 1968) and ORSTOM 1966 (Guillaumet and Hallé, 1978). If, at given locality, there was only one accession in the collection, it will be more important with regard to this criterion than an accession from a locality from which there are many accessions. **Table 3** gives the score table for uniqueness. To determine the score for the accessions from a locality, we looked at the range of number of accessions per locality, grouped them (**Table 3**), and assigned scores of 0 to 15 from the highest number of accessions to the lowest.

TABLE 3 | Score table for number of locality areas.

| Number of accessions from a locality | Score |
|--------------------------------------|-------|
| 1 | 15 |
| 2–4 | 12 |
| 5–10 | 10 |
| 11–17 | 8 |
| 23–30 | 5 |
| >30 | 0 |

For each accession, the exact location of collecting has been determined from the passport data file, original reports (Ferne et al., 1968; Guillaumet and Hallé, 1978), and the literature. The number of accessions known to have been collected from each specified site was then compiled. Another consideration was whether the site was from the Southwestern part of Ethiopia, considered as the center of diversity of *C. arabica* (Sylvain, 1955; Charrier and Bethaud, 1985; Bramel et al., 2017). In this study, we considered any accession coming from provinces of Kaffa and Illubabor of Ethiopia as originating from the center of diversity, and it received an additional 5 points. The maximum score for this criterion was 20.

Note that an accession can also be unique if it is only present in the CATIE collection and in no other collection. However, this could not be accounted for under this criterion because of insufficient information from other coffee collections around the world (see *Safety duplication*).

Safety Duplication

This criterion refers to whether an accession held in the collection is safely duplicated in another genebank, within or outside the country, for which ideally there should be an official signed agreement between the two institutions. However, in practice, if an accession is known based on a reputable source to be found in another collection, it should be counted as being duplicated, although its security may not be guaranteed. If an accession is not known to be safely duplicated, then it is considered as highest priority and receives a maximum score of 20. Depending on the number of collections in which an accession is found, it receives successively fewer points. Thus, the scoring for safety duplication is as follows:

- No evidence of safety duplication (20 points).
- Safety duplicated in 1 collection (15 points).
- Safety duplicated in 2 collections (10 points).
- Safety duplicated in 3 or more collections (5 points).

In the case of CICC, data on safety duplication were obtained from three other coffee collections, namely, US Department of Agriculture (USDA), Fort Collins, CO, USA; the Nica-France Foundation farm at La Cumplida, Nicaragua; and the Institut de Recherche pour le Développement, France (IRD) collection at La Reunion (personal communication with Stephanie Greene, Melanie Bordeaux, and Thierry Joet, respectively).

RESULTS

Composition of the CICC

The 2019 inventory of the CICC revealed a total of 1,975 accessions in the field, organized in 10 sections, labeled A to H, Musas and Citricos, in CATIE's campus. However, when compared with records of the passport data and introduction files, there was a disparity among them, summarized as follows:

- 172 accessions of 2019 inventory without passport data.
- 18 of 2019 inventory without passport data and not found in introduction book.
- 111 accessions on records with passport data, but not in the field.
- 130 accessions recorded in introduction book, but not in field.
- 47 accessions found in the field (2019 inventory) with no accession number (these were hybrid varieties from CATIE's breeding program).

After accounting for these disparities, the total number of accessions remaining in the CICC as of December 2019 was 1,895 accessions.

The collection was made up principally of arabica coffee, with 91.93% (including introgressed and other interspecific hybrids), followed by *C. canephora* Pierre ex Froehner (4.06%) and *Coffea liberica* Bul ex Hiern (1.27%). Other species with very low percentages (<1% each) included *Coffea sessiliflora* Bridson (14 accessions), *Coffea brevipes* Hiern (7 accessions), *Coffea pseudozanguebariae* Bridson (11 accessions), *Coffea eugenioides* S. Moore (6 accessions), *Coffea racemosa* Lour. (3 accessions), *Coffea salvatrix* Swynn. & Philipson (2 accessions), and *Coffea congensis* A. Froehner (1 accession). There were eight accessions for which the species is unidentified (**Figure 3A**).

The CICC may be broken down as follows with regard to the type of genetic resource, in order of priority (**Figure 3B**):

1. Wild materials (661 accessions)
2. Cultivated varieties (443 accessions)
3. Breeding lines and experimental materials (784 accessions).
4. Unknowns (blanks) (7 accessions).

The wild material included mostly accessions collected during the FAO collecting mission 64/65 in Ethiopia (E series) (424 accessions [64%]) and the ORSTOM collecting mission (ET series) (90 accessions). It also included materials collected by IPGRI in Yemen (nine accessions). In addition, some clonal materials (*in vitro* plants) were received from ORSTOM (97 accessions). The remaining accessions were mostly diploid species that came from other sources, including CIRAD, and countries from South America.

It must be noted that many CICC accessions arising from the ORSTOM mission (identified by their ET codes) were introduced in CICC from three different sources (**Supplementary Table S2**). Consequently, there were three sets of accessions for the same material, but with different CICC accession numbers, totaling 155 accessions. Fifty-two accessions were received from Institut de Recherches du Café et du Cacao (IRCC), Paris (with IRCC numbers) in 1985 (T.16689 to T.16741); 45 accessions were received from Institut de Recherches du Café et du Cacao,

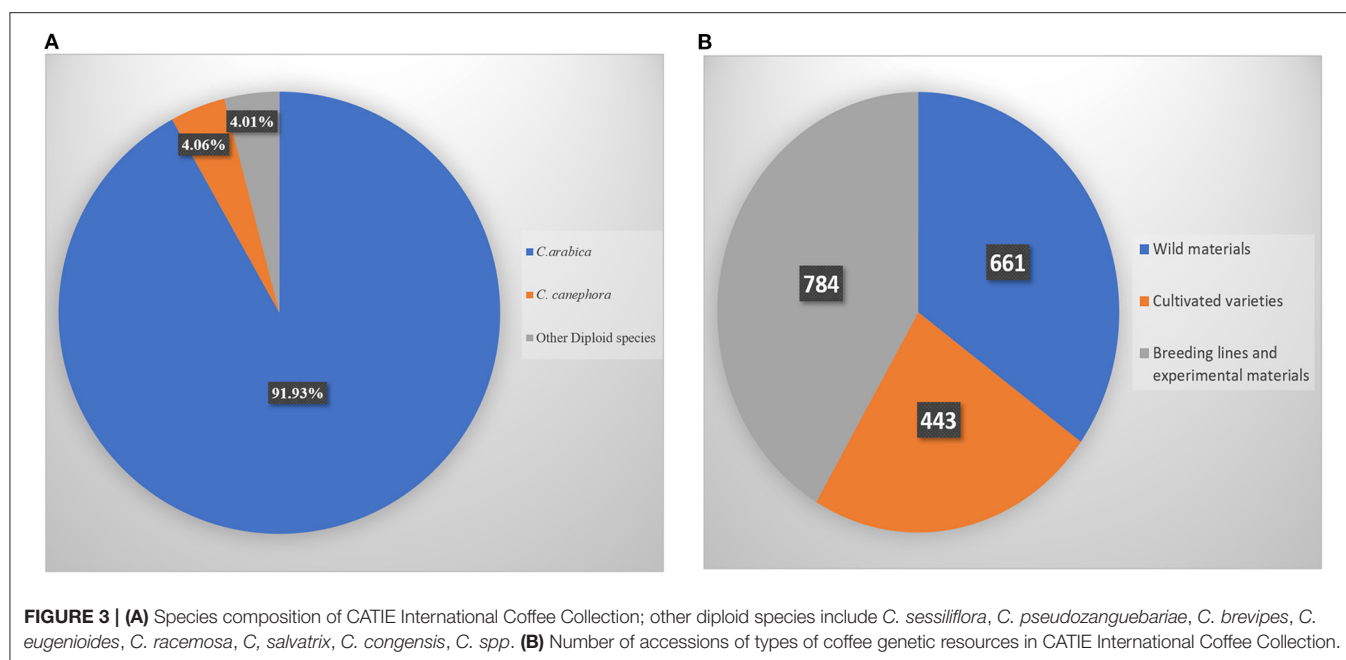


TABLE 4 | Number of accessions in threat categories [Brackets under wild total represent the effective numbers of accessions, after accounting for duplications of ORSTOM accessions (ET codes)].

| Threat categories | Code | Score range | Wild <i>C. arabica</i> | Wild diploids | Wild total | Cultivated varieties <i>C. arabica</i> | Combined |
|-------------------------|------|-------------|------------------------|---------------|--------------|--|----------------------|
| Extinct | EX | | 6 | 23 | 28 | 19 | 47 (+33 other types) |
| Critically endangered | CE | 18–20 | 13 | 9 | 22 (14) | 27 | 49 |
| Endangered | EN | 15–17 | 45 | 17 | 62 (38) | 41 | 103 |
| Threatened | TR | 10–<15 | 166 | 36 | 202 (155) | 166 | 368 |
| Vulnerable | VU | 5–<10 | 215 | 19 | 234 (212) | 145 | 379 |
| Total threatened | | | 439 | 81 | 520 | 379 | 899 |
| Not threatened | NT | 1–<5 | 140 | 1 | 141 (120) | 64 | 205 |
| TOTAL | | | 579 | 82 | 661 | 443 | 1104 |
| % threatened | | | | | 78.6% | 85.6% | 81.4% |

Nkolbisson, Cameroun with code cX (number of Bouharmont) or L series number, received in 1986 (T.17173 to T.17242) (Dulloo, 2020a). It was uncertain from where these materials were derived. It could have been introduced from the ORSTOM missions in West and East Africa in 1970s and 1980s (Charrier and Bethaud, 1985). Furthermore, 58 accessions were received as clones introduced and multiplied *in vitro* from ORSTOM, Montpellier, in 1995 (T.21259 to T.21316). Thus, the same ORSTOM mission population (ET code) had been introduced and assigned a different CATIE accession number (T series). For example, accessions T.16689, T.17173, and T.21259 all represented the same ORSTOM mission population ET-01 and were introduced from Paris, Cameroon, and as *in vitro* material from Montpellier, respectively.

In this study, we focused on the wild material and cultivated arabica varieties as being the top priority from a conservation perspective; for each of these two groups, the accessions were ranked using the criteria and categories for which data were available.

Lost Accessions

According to the 2019 inventory, 80 accessions from the collection were lost (Table 4). Fifty-six accessions were of *C. arabica*, 23 accessions were of diploid species, and one unknown species (T.04466). Among diploids, the largest losses were those of *C. racemosa* (six accessions), *C. pseudozanguebariae* (five accessions), and *C. eugenoides* (three accessions); *C. canephora*, *C. brevipes*, and *C. sessiliflora* had each lost two accessions, whereas the accessions *C. congensis* (T.04098), *C. liberica* (T.02536), and *Coffea stenophylla* (T.03416) were also lost.

Of the 80 lost accessions, 28 accessions were of wild materials, 19 were cultivated varieties, and 30 were of breeding lines or experimental material and three unknown materials (T.04466, T.05056, T.055058) (Table 4). Only two accessions (T.04538-E292 and T.04819-E419) of the FAO collecting mission 1964/65 had been lost from the historic collection. Four accessions of the ORSTOM collection were also lost (T.17224 [ET-39], T.21284 [ET24], T.21291 [ET029B], and T.21295 [ET-33B]). The last three accessions were *in vitro* materials received from France.

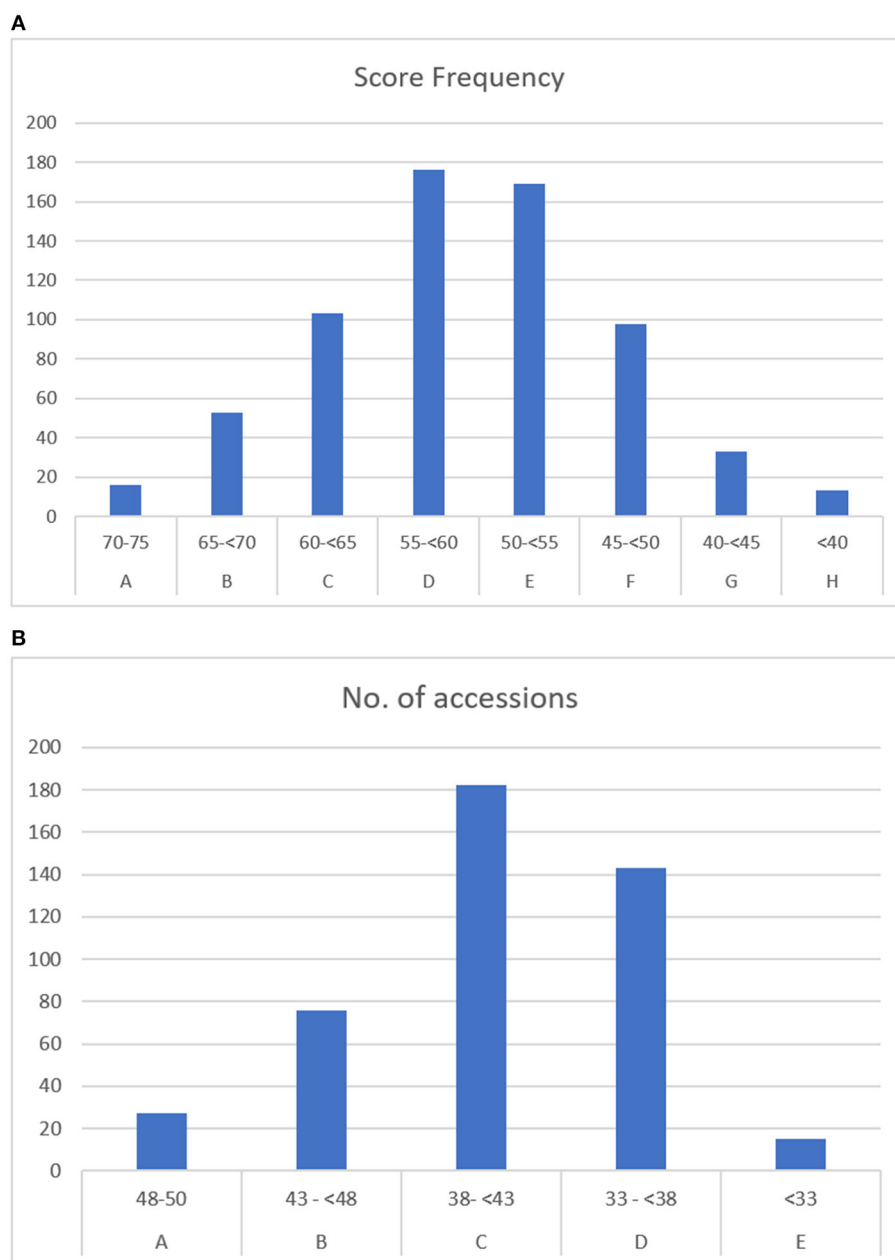


FIGURE 4 | (A) Overall score frequency of wild accessions in priority groups A to H. **(B)** Overall score frequency of cultivated accessions in priority groups A to E.

Status of Wild Material

The overall prioritized accessions of the so-called wild materials were divided into eight groups according to score range (Figure 4A). The highest-priority accessions were those belonging to groups A to C (having a score of ≥ 60), giving a total of 172 accessions of wild materials.

Threat Status of Wild Materials

Besides the 28 accessions lost from the collection, there were 22 accessions of the wild materials considered as critically endangered (Table 4), of which five were *C. arabica* accessions

collected from the FAO collection mission 1964/65 collection, eight are accessions of *C. arabica* from the ORSTOM mission, and nine were diploid accessions received as *in vitro* clones from ORSTOM Montpellier.

The most threatened accessions are the accession T.04738 (*C. arabica*, E215) and the two diploid accessions T21320 (*C. pseudozanguebariae*) and T21359 (*C. racemosa*), which received the maximum score of 20. They are all represented by a single remaining plant in the collection, and their health status is critical. Accession T.04848 is the next most threatened accession, with 19 points. In addition, there are three wild

TABLE 5 | Accessions of ET codes that are critically endangered.

| ET code | Accessions critically endangered | Comments | ET code combined treat status |
|---------|----------------------------------|---|-------------------------------|
| ET-01 | T.21259 | Represented by two other accessions of T.17173 and T.16689, that have in combination a total of 18 plants of which 11 plants are in healthy state | NT |
| ET-20 | T.21279 and T.21280 | Also represented in the collection under 3 other accessions numbers – T.17236, T.17205, T.17206. Combined, there are 16 individual plants and 9 are healthy. | NT |
| ET-32B | T.21292 | Also represented in the collection under 3 other accessions numbers – T.17236, T.17205, T.17206. Combined, there are 16 individual plants and 9 are healthy. | NT |
| ET-38 | T.21301 | Also represented in the collection by another accession T 17223, which has 9 plants, of which 6 plants are healthy | NT |
| ET-41 | T.21303 | Represented in the collection by another accession (T 16725) which has 11 plants, of which 8 plants are healthy. It had another accession, T.17201, that have been lost | NT |
| ET-44 | T.21304 | Represented in the collection by another accession, T 16728, which has 6 plants, of which 4 plants are healthy | NT |
| ET-52 | T.21308 | Represented in the collection by another accession, T 16728, which combined have 6 plants, of which 4 plants are healthy. | TR |

accessions (T.04704, T.04768, and T.04868) from the FAO mission originating from the center of diversity in Kaffa and Illubabor that are critically endangered.

The multiple introductions of the same ORSTOM material posed some challenges in the assessment of the threats of these CICC accessions. A separate analysis of the ET series codes was carried out (Dulloo, 2020a), and a combined threat score was calculated for the same ET code accessions (**Supplementary Table S3**). The results here showed that none of the ET codes were critically endangered, although eight CATIE accessions from ORSTOM mission were classified as critically endangered (**Table 5**). As there were other CICC accessions representing the same ORSTOM mission population, the combined assessment gave a “not threatened” result, as together they would have a greater number of healthy plants, thus making them not critically endangered. For example, CICC accession T21259-ET-01 was scored as critically endangered, but the same population (ET-01) was represented by two other accessions (T.17173 and T.16689), which have in combination a total of 18 plants, of which 11 plants were in healthy state. When assessed in combination, the ORSTOM mission population ET-01 was considered as not threatened. Thus, taking this into account, only 14 accessions in CICC were considered as “critically endangered” (**Table 4**).

There were 62 accessions that scored between 15 and 17 points and were considered as “endangered.” Nearly 80% of them (49 accessions) were from the ORSTOM mission, and a significant number had been received as *in vitro* clones. Seven populations collected by the ORSTOM mission were represented by a single accession, but the remaining 26 populations had more than one accession, and when these were combined, they were not endangered. Two accessions (T.17181 [ET-11C] and T.17241 [ET-35D]) were represented by several other accessions in the collection, but they were the original germplasm from Cameroon that should be considered as priority to conserve. They were represented by only one or two plants in the collection and were thus endangered. Thus, there were 38 accessions that were effectively “endangered,” of which 20 were *C. arabica*.

Regarding the “threatened” accessions, there were 202 accessions, with 50% coming from the FAO mission. As with other categories, many of the unique ORSTOM mission populations were represented by multiple CICC accessions. Consequently, of the 202 accessions, only 155 were effectively threatened, of which 119 were *C. arabica* accessions.

The great majority of “vulnerable” accessions (189 of 234) were *C. arabica* from the FAO mission. Four populations of the ORSTOM mission were represented by a single accession that is vulnerable. In addition, there were several combinations of accessions with a single ET code that together gave a vulnerable threat score. Consequently, of 234 accessions, there were 212 accessions that were effectively vulnerable (with 196 of *C. arabica* accessions). In addition, there were 10 combined sets of accessions of ET series codes that were considered as “vulnerable.”

The remaining 141 accessions were considered as “not threatened,” with a total score of less than 5 (**Table 4**). They included 115 *C. arabica* accessions from the FAO mission, 25 *C. arabica* accessions of the ORSTOM mission, and only one accession (T21329) of *C. canephora* (*in vitro* clone from ORSTOM). However, among the 25 CATIE accessions of the ORSTOM mission, there were four ORSTOM mission populations ET-03, ET-09, ET-26, and ET42 that were represented by a single CATIE accession, namely, T.16691, T.16697, T.16713, and T.16726, respectively, and were not threatened (**Supplementary Table 3** [Annex 8-ET codes]). For the remaining 21 CATIE accessions, there were other accessions that represented the same ORSTOM mission population, which belonged to the threatened category. Consequently, in total, there were 120 accessions that were effectively “not threatened.”

Uniqueness

Table 6 shows the number of accessions that belong to the different “uniqueness” categories, within and outside the center of diversity (see *Methods* section). Unfortunately, we do not have the full passport data (especially the site information) on all the accessions in the collection. This limits the kind of analysis

that we can do, and overall, the uniqueness criterion is difficult to apply to all accessions in the CICC. It was only possible to do this analysis for the accessions for which information on locality names was available, which included mostly the historic collection missions of FAO (Ferne et al., 1968), IPGRI, and ORSTOM (Guillaumet and Hallé, 1978).

The most unique accessions from the FAO and ORSTOM missions were those that came from only a single locality within the center of diversity. There were five accessions that satisfied these criteria and scored the maximum score of 20 points (Table 7). These accessions were not highly endangered and were either VU or TR. There was also a set of 20 accessions (Table 8) that came from only one locality, but not within the center of diversity.

Status of Cultivated Material

The 443 accessions of the cultivated materials were ranked using the criteria of germplasm type, threat score, and safety duplication. The accessions were grouped into five groups (A–E) (Figure 4B), with the highest priority the accessions in groups A and B (scoring ≥ 43 points), which contained 103 accessions, 27 from group A and 76 from group B. It is interesting to note that all the accessions from group A were introduced into the collection in the 1950s and 1960s. None of the accessions in groups A and B were safety duplicated. There were six accessions that were duplicated in group D, and the rest of the duplicated materials were from group E. This means that none of the higher-priority groups were duplicated, and only a few of the least important accessions were duplicated.

TABLE 6 | Number of accessions within and outside center of diversity.

| Uniqueness categories | Number of accessions in center of diversity (Provinces of Kaffa and Illubabor) | Number of accessions outside center of diversity | Total |
|-----------------------|--|--|------------|
| 1 | 5 | 20 | 25 |
| 2–4 | 24 | 17 | 41 |
| 5–10 | 60 | 23 | 83 |
| 11–17 | 88 | 37 | 125 |
| 23–30 | 99 | 32 | 131 |
| >30 | 99 | 69 | 168 |
| Subtotal | 374 | 198 | 572 |
| Blanks | | | 88 |
| TOTAL | | | 661 |

TABLE 7 | Most unique accessions in the FAO and ORSTOM collections (*C. arabica*).

| CICC accession number | Identification | Country of introduction | Locality name | Center of diversity | Treat status |
|-----------------------|----------------|-------------------------|-------------------|---------------------|--------------|
| T.04775 | E-272 | Ethiopia | Kolu | kaffa | VU |
| T.04780 | E-277 | Ethiopia | Komba Agaro | kaffa | TR |
| T.04781 | E-287 | Ethiopia | Shebe | kaffa | VU |
| T.04924 | E-555 | Ethiopia | Kaffa | kaffa | TR |
| T.17231 | ET-45 | Cameroon | Filola Hotsprings | kaffa | TR |

Threat Status of “Cultivated Materials”

Table 4 gives the summary of the number of accessions of cultivated materials belonging to the five threat groups. There were 19 accessions of cultivated varieties that had been lost from the collection as per the 2019 inventory. Twenty-seven accessions of cultivated varieties were critically endangered and were represented by a single plant in the collection. Among them, there were seven accessions for which the state of health of the single plant left is critical. These needed to be salvaged as a matter of urgency. The other 20 accessions were also in a poor state and needed to be salvaged as soon as possible. These accessions were also not duplicated elsewhere (i.e., in USDA and La Cumplida farm).

There were 41 “endangered” accessions, of which 11 accessions were represented by only a single plant, but they were all healthy, and the rest were represented by two plants only. Among the latter, there were six accessions (T.02544, T.03685, T.04295, T.04310, T.17557, T.02699, and T.17931) whose health

TABLE 8 | Accessions from one locality outside center of diversity (*C. arabica*).

| CICC accession number | Identification | Country of introduction | Locality name |
|-----------------------|----------------|-------------------------|------------------|
| T.04472 | E-007 | Ethiopia | Combulchia |
| T.04500 | E-036 | Ethiopia | Welkitte |
| T.04612 | E-089 | Ethiopia | Gera- Jimma |
| T.04622 | E-124 | Ethiopia | Geisha |
| T.04623 | E-125 | Ethiopia | Tui |
| T.04624 | E-126 | Ethiopia | Gorei -Geisha |
| T.04758 | E-237 | Ethiopia | Yirgalem |
| T.04759 | E-238 | Ethiopia | Aleta Wondo |
| T.04892 | E-524 | Ethiopia | Omonadda |
| T.04945 | E-579 | Ethiopia | Faghena Erythrea |
| T.04946 | E-001 | Ethiopia | Finote Selam |
| T.04950 | E-012 | Ethiopia | Harar city |
| T.04952 | E-020 | Ethiopia | Dilla |
| T.21231 | PDRY-01 | France | Hashsh |
| T.21234 | PDRY-05 | France | Dinakheb |
| T.21236 | PDRY-07 | France | Hanaka |
| T.21237 | PDRY-09 | France | Baynassayal |
| T.21239 | PDRY-014 | France | El-Kor |
| T.21240 | PDRY-15 | France | Diraa |
| T.21242 | PDRY-22 | France | Hewle |

status was all “critical,” except accessions T.04310 and T17931, which had one plant as “critical” and one plant are “poor.” These accessions were thus of highest priority in this category considering health status.

One hundred sixty-six accessions of cultivated varieties were threatened, with three accessions represented by a single individual (T17545, T.17548, and T.17549) and were all in poor health. Twenty-seven accessions were represented by two plants, with 10 accessions healthy and the remaining accessions with one plant either healthy or poor. Furthermore, there were 66 accessions with three plants. Only nine accessions of these were healthy. It should be noted that accession T.04308 has only three plants, all of which were in a critical condition and were among the top two most threatened accessions. There were 63 accessions that were also represented by four plants, four accessions by five plants, and three accessions by six plants.

Among the 145 vulnerable accessions, the number of plants per accession varied from 3 to 19. There were two accessions (T.17541 and T.05038) with no original plants and were represented by three regenerated plants. There were eight accessions that were having a high health status score (3 points), and the rest of the accessions had fewer than 3 points and were doing fairly well. These accessions with high numbers of plants (>8) can be rationalized to reduce the size of the collection.

There were 64 accessions that were considered as “not threatened” and were all doing very well. The outlier was the variety Geisha (T.02722), which was represented by no fewer than 276 plants in the collection. However, the majority of the plants were used for producing seeds for distribution and were not strictly part of the collection. The rest of the accessions had 7 to 19 plants. These accessions with very high numbers of plants (>8) can be rationalized to reduce the size of the collection. It was suggested that only eight plants of Geisha (T.02722) be kept in the rationalized long-term collection and the rest kept in a working collection.

Safety Duplication

The CICC has 221 accessions, representing 11.7% of the total, which are considered as being safely duplicated in three institutions outside Costa Rica, namely, USDA, Nica-France Foundation, and IRD-La Reunion. Of these, only USDA has a formal agreement with CATIE for safety duplication under black-box conditions in cryopreservation. The safety duplicates in the Nica-France Foundation and IRD-La Reunion are both held in field genebanks, but IRD also holds its collection in cryopreservation. Besides these, there is a high probability that some of historic collections are also being conserved in the genebanks in Ethiopia, Tanzania, India, Peru, and elsewhere (Brazil, Colombia), as materials from the historic collections are known to have been sent to these places (Ferne et al., 1968). However, we have no information to date about which accessions are still extant in these national collections. It is recommended that these collections should also be studied using the methodology described in this article to determine which accessions they share in common.

DISCUSSION

Prioritization Methodology

Coffee field genebanks have often been criticized as being highly vulnerable (Vega et al., 2008; Dulloo et al., 2009; Bramel et al., 2017). Most field genebanks in the world suffer from the vagaries of changing climate and weather, inappropriate field conditions, pest and disease outbreaks, fire, and aging plants (FAO, 2013; Alemayehu and Merga, 2017; Bramel et al., 2017). In addition, collections keep growing with time as new accessions are added, and this makes the collection more difficult and expensive to maintain, with the results that the collection is not properly curated, labels are lost from the field, and records are not well kept. These technical, management, and economic constraints severely impact the sustainability of field genebanks, in general, and coffee field genebanks, in particular. In this article, we describe the application of an accession-by-accession methodology to effectively monitor, prioritize, and rationalize the field coffee genebank maintained by CATIE under Article 15 of the Plant Treaty. This methodology can be used by other field genebank curators to assess the status of their collection and ensure that they are properly managed in a cost-effective manner and at high international standards, as recommended by FAO (2013).

The five criteria and categories proposed (type of genetic resources, threat, uniqueness, safety duplication, and use) are key to the prioritization process. However, their successful application is highly dependent on the availability of the relevant information. In this study, the “use” criterion was initially regarded as a very important criterion because it documents the value to the accession and promotes its use, which should be the ultimate goal for maintaining the collection. The “use” criterion should consider the different traits that are important for coffee industry and include resistance to rust and nematodes, cupping quality, biochemical profile (caffeine, chlorogenic acid, sucrose content), resilience to climatic factors, yield, vigor, male sterility, and level of heterozygosity. However, the CICC has not been evaluated for all these traits, and only partial information is available, which makes this criterion impossible to use for prioritizing the entire collection. For example, approximately 50% of the collection has been evaluated for vigor, whereas only 25 accessions have been tested for their organoleptic characteristics. A limited number of accessions have been evaluated, and the characteristics evaluated for some accessions include tolerance to rust, nematodes, vigor, physical characteristics of the grain, and organoleptic qualities. The World Coffee Research has also characterized 847 wild accessions using molecular markers (Klein et al., 2016). Currently, the characterization data of 7 traits for only 34 accessions are made available and are uploaded on the Genesys portal <https://www.genesys-pgr.org/>; however, the full passport data set of the complete CATIE coffee collection can be found on the portal.

Among the remaining four criteria, the type of genetic resources, that is, whether they were of wild origin, cultivated varieties, or resulting from experimental and breeding activities, was considered as the most important criterion given that it has been demonstrated that wild types are genetically more diverse

compared with cultivated arabica varieties (Lashermes et al., 1996; Scalabrin et al., 2020). However, it is important to clarify that among the so-called “wild materials” accessions in CICC, very few were truly from the wild (Ferne et al., 1968). This was confirmed by a molecular genetic diversity study of the historic collection in CICC, indicating that less than 1.5% of the so-called wild accessions were actually derived from the wild and small farms (Klein et al., 2016).

With regard to the uniqueness criterion, it could only be applied to the so-called wild materials, where the names of the sites from where the seeds were collected, during the FAO and ORSTOM collecting missions, were available (Ferne et al., 1968). This criterion is also considered as being important as it informs us about the origin of the accessions and whether they came from the same or different sites and provides an indication of the genetic distinctiveness of the subpopulations from which they originate. Geographic distribution is often considered as a good proxy for genetic diversity (Pelletier and Carstens, 2018). Thus, the extent of accessions from different locations can be a good indication of the breadth of genetic diversity conserved in the field collection. Often, field genebanks contain a disproportionate number of accessions from the same subpopulation, and knowledge of this can inform how to rationalize the collection to maximize cost-effectiveness in genetic resources conservation, as well as identify gaps in collections. Where feasible, the use of molecular markers would help significantly in characterizing the genetic distinctiveness of the accessions and be used to score this criterion.

In the threat criterion, the category “number of original plants” was considered important because they are the most original plants that were planted and not regenerated and thus was given extra points. It is true that genetically this should not make a difference in terms of genetic diversity, if clonally propagated. However, the fact that the original trees of an accession persisted in the collection can be an indication that they are more adapted to the field conditions than those accessions that have lost their original plants and also to be less likely for errors to be made.

The multiple introductions of the same genetic resources in a collection can make the genebank management at the accession level become complicated, as illustrated by ORSTOM missing introductions in this study. It is important that new introductions be properly screened to see if similar genetic materials do not exist anymore and that records are properly cross referenced. In the case of ORSTOM mission populations, the same materials were introduced from different sources, but in different forms, as plantlets and *in vitro* materials. It was interesting to note that among the accessions that have been introduced to the CICC as *in vitro* clonal materials from ORSTOM, a relatively high proportion are doing poorly compared with the original accessions, and eight accessions (T.21259 to T.21308) were critically endangered (Dulloo, 2020a). This may suggest that plants derived from *in vitro* propagated materials may have a shorter longevity in the field compared with those propagated by cuttings or seeds. However, more research is required to verify this.

Composition of the Accessions Conserved at CICC

The Global Strategy on Coffee Genetic Resources (Bramel et al., 2017) reported that CATIE's accessions from the historic collecting missions of 1964/1965 in Ethiopia by FAO (Ferne et al., 1968) and of 1966 by ORSTOM (now IRD) (Guillaumet and Hallé, 1978) constitute approximately 40% of the conserved accessions. The material from the IPGRI collecting expedition in Yemen (Eskes, 1989) is represented by only a few (17) accessions. However, the present study has shown that the FAO collection makes up of 22% of the collection (424 accessions), ORSTOM 8% (145 accessions), and only nine accessions (0.5%) come from the Yemen expedition. This difference may be partly due to the fact that the collection may be losing accessions; the study showed that 80 accessions have actually been lost since. But it also may be due to new materials arriving, mainly as breeding lines and experimental materials. The collection lost a disproportionate number of diploid species (23 accessions) such as *C. racemosa*, *C. pseudozanguebariae*, and *C. eugenoides*, *C. canephora*, *C. brevipes*, and *C. sessiliflora*, compared with *C. arabica* (56 accessions), which make up more than 90% of the collection. This is probably due to the low adaptive potential of these diploid wild species to grow under the environmental conditions existing at CICC. They generally exhibit narrow climatic envelopes with restricted habitat (niche) specificity and are mostly forest dwelling.

Rationalization of the CICC Collection

The application of this prioritization methodology to CICC allowed a set of recommendations for the rationalization of collection at CICC and to define cost strategic conservation actions, including re-establishment and safety duplication of the CICC (Dulloo, 2020b). It is recommended that the so-called wild collected accessions (661 accessions) and cultivated varieties (443 accessions), a total of 1,104 accessions, are included in the rationalized collection, as they represent the most diverse and high-value material in the collection and that the breeding lines and experimental materials including all hybrid materials be moved to separate working collection. This recommendation would allow for a more manageable cost-effective rationalized collection to be established. Furthermore, it was recommended that the number of individual plants in each accession be brought down to six individual plants, keeping the original plants as far as possible. In cases where the numbers are fewer than 6, they should be multiplied urgently to bring back the numbers to six per accession.

The in-depth study (Dulloo, 2020a) also showed that there are 899 accessions (81.4%) considered as being threatened. There were 22 accessions of wild collected genetic material and 27 accessions of cultivated materials that are critically endangered, 49 accessions in total. These are accessions that are down to the last individual that has a critical or poor health status. Immediate action needs to be taken to propagate them by the best available technique to ensure their survival and be replanted in a new location of the rationalized collection (Dulloo, 2020b). There are also 62 accessions from wild collected materials and 41 accessions

cultivated varieties that are endangered, 103 accessions in total. These are accessions that are down to one or two plants and have poor or healthy plants and also need to be rescued urgently.

The safety duplication of CICC was also considered as a high priority, and the study report (Dulloo, 2020b) recommended that priority for safety duplication be given to the historic collection of *C. arabica* belonging to the most threatened groups and that diploid species should be safety duplicated as live plants in another collection site suitable for these species. A total of 403 accessions were identified for urgent safety duplication in cryopreservation.

CONCLUSION

The in-depth study of the CICC allowed its management to carry out a full inventory of the field collection in December 2019, including an assessment of the health status of each tree across all surviving accessions, and to reassemble all its field genebank records into one database system, which henceforth will greatly facilitate the monitoring of the coffee field genebank. It allowed the management to reconcile its records on file with what is actually conserved in the field. Furthermore, the study also allowed CATIE to develop a full rationalization plan for its collection. A new site has been identified, and work has started for the multiplication of 1,104 priority accessions, which will be part of the new collection.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

REFERENCES

- Adams, R. P., and Adams, J. E., (1991). *Conservation of Plant Genes: DNA Banking and in Vitro Biotechnology*. Academic Press, New York
- Alemayehu, D., and Merga, W. (2017). Current status of arabica coffee (*Coffea arabica* L.) Genetic resources. Conservations, constraints, and mitigation strategies in Ethiopia. *Int. J. Res. Stud. Sci. Eng. Technol.* 4, 1–11. ISSN: 2349-476X.
- Bettencourt, E., and Konopka, J. (1988). Directory of germplasm collections. 5. II. *Industrial Crops: Beet, Coffee, Oil Palm, Cotton and Rubber*. International Board of Plant Genetic Resources, Rome.
- Bramel, P., Krishnen, S., Horna, D., Lainoff, B., and Montagnon, C. (2017). *Global Conservation Strategy for coffee genetic resources*. Global Crop Diversity Trust and World Coffee Research. Available online at: https://cdn.croptrust.org/wp/wp-content/uploads/2017/07/Coffee-Strategy_Mid_Res.pdf. (accessed September 2021).
- Charrier, A., and Bethaud, J. (1985). Botanical classification of coffee. In: Clifford and Wilson, eds, *Coffee Botany biochemistry and production of beans and beverages*. Croom Helm, London and Sydney. pp. 13–47. doi: 10.1007/978-1-4615-6657-1_2
- Dulloo, M. E. (2020a). Report on the in-depth study of CATIE International Coffee Collection. In: *Consultancy Report Submitted to Global Crop Diversity Trust*, 2020.
- Dulloo, M. E. (2020b). Rationalization plan for CATIE International Coffee Collection. In: *Consultancy Report Submitted to Global Crop Diversity Trust*, May 2020.
- Dulloo, M. E., Charrier, A., Dussert, S., Anthony, F., Tesfaye, S., Rakotomalala, J. J., et al. (2001). Conservation of coffee genetic resources: constraints and opportunities. In: *Proceedings of 19th World Colloquium on Coffee Science*, ASIC, Trieste, Italy. May 2001. Association Scientifique Internationale du Café, Paris.
- Dulloo, M. E., Ebert, A. W., Dussert, S., Gotor, E., Astorga, C., Vasquez, N., et al. (2009). Cost efficiency of cryopreservation as a long-term conservation method for coffee genetic resources. *Crop Sci.* 49:2123–2138. doi: 10.2135/cropsci2008.12.0736
- Dulloo, M. E., Guarino, L., Engelmann, F., Maxted, N., Newbury, H. J., Attore, F., et al. (1998). Complementary conservation strategies for the genus *Coffea*: a case study of Mascarene *Coffea* species. *Genet. Resour. Crop Evol.* 45, 565–579. doi: 10.1023/A:1008621028343
- Dussert, S., Chabrilange, N., Anthony, F., Engelmann, F., Recalt, C., and Hamon, S. (1997). Variability in storage response within a coffee (*Coffea* spp.) core collection under slow growth conditions. *Plant Cell Rep.* 16, 344–348. doi: 10.1007/s002990050236
- Dussert, S., Chabrilange, N., Rocquelin, G., Engelmann, F., Lopez, M., and Hamon, S. (2001). Tolerance of coffee (*Coffea* spp.) seeds to ultra-low temperature exposure in relation to calorimetric properties of tissue water, lipid composition and cooling procedure. *Physiol. Plant.* 112, 495–504. doi: 10.1034/j.1399-3054.2001.1120406.x
- Eira, M. T. S., Fazuoli, L. C., Filhó, O. G., Silvarolla, M. B., Ferrão, M. A. G., Aymbire Francisco A., et al. (2007). Bancos de Germoplasma de Café no Brasil. In: Brasília, D.F. *Embrapa Recursos Genéticos e Biotecnologia*. p. 18.

AUTHOR CONTRIBUTIONS

MD, WS, DD, and LG conceived the work. MD designed the work. CA acquired the material. MD analyzed and interpreted the data. MD drafted the manuscript which was reviewed by LG, CA, and DD. All authors read and approved the final manuscript.

FUNDING

This study has been commissioned by the Global Crop Diversity Trust and was sponsored by FELCO SA.

ACKNOWLEDGMENTS

We wish to thank the management of the Global Crop Diversity Trust and CATIE for facilitating this study. Our thanks also go to Stephane Dussert and Thierry Joet of IRD, Stephanie Greene of USDA Fort Collins and Benoit Bertrand of CIRAD and Melanie Bordeaux of Foundation Nica-France (La Cumplida, Nicaragua) for providing information on safety duplicates held by them in their genebanks. We thank Andreas Ebert, Jan Engels, Francisco Anzueto for providing comments on the prioritization methodology developed for the study. Thanks are also due to CATIE field genebank staff for undertaking the field inventory and Djuneid Dulloo for making the flowchart diagram in **Figure 1**.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fsufs.2021.777415/full#supplementary-material>

- Engelmann, F., and Dulloo, M. E. (2007). Introduction. In: *Complementary Strategies for ex situ Conservation of Coffee (Coffea arabica L.) Genetic Resources. A case study in CATIE, Costa Rica.* eds Engelmann, F., Dulloo, M.E., Astorga, C., Dussert, S., Anthony, F. *Topical reviews in Agricultural Biodiversity. Bioversity International, Rome, Italy.* p. 1–11.
- Eskes, A. B. (1989). Identification, description and collection of coffee types in P.D.R. Yemen. Technical report [from IRCC/CIRAD to IBPGR] of the IBPGR/PDR Yemen Ministry of Agriculture/IRCC-CIRAD mission to Yemen PDR, 15 April–7 May 1989. IBPGR (now Bioversity International) internal report.
- FAO (2009). *The International Treaty on Plant Genetic Resources for Food and Agriculture.* Rome, Italy. p. 56.
- FAO (2013). *Genebank Standards for Plant Genetic Resources for Food and Agriculture.* Rome, Italy. p. 166.
- Fernie, L. M., Greathead, D. J., Meyer, F. G., Monaco, L. C., and Narasimhaswamy, R. L. (1968). *FAO Coffee Mission to Ethiopia, 1964-65.* FAO, Rome, Italy. p. 204.
- Guillaumet, J. L., and Hallé, F. (1978). Echantillonnage du matériel *Coffea arabica* récolté en Ethiopie. In: Charrier, A. (ed.) *Etude de la structure et de la variabilité génétiques des caféiers: résultats des études et des expérimentations réalisées au Cameroun, en Côte d'Ivoire et à Madagascar sur l'espèce Coffea arabica L. collectée en Ethiopie par une mission ORSTOM en 1966.* ORSTOM Paris and IFCC Paris. Bulletin IFCC no 14 September 1978.
- Hong, T. S., and Ellis, R. H. (1995). Interspecific variation in seed storage behaviour within two genera, *Coffea* and *Citrus*. *Seed Sci. Technol.* 23, 165–181.
- Klein, P., Murray, S., Solano, W., Montagnon, C., Schilling, T., and Bertrand, B. (2016). An extensive study of the genetic diversity of Arabica species evidences two genetic groups in Ethiopia and a sub-group in Yemen. In: *26th International Conference on Coffee Science.* November 13–17, 2016, Kunming, China.
- Labouisse, J.-P., Bellachew, B., Kotecha, S., and Bertrand, B. (2008). Current status of coffee (*Coffea arabica* L.) genetic resources in Ethiopia: implications for conservation. *Genet. Resour. Crop Evol.* 55, 1079–1093. doi: 10.1007/s10722-008-9361-7
- Lashermes, P., Trouslot, P., Anthony, F., Combes, M. C., and Charrier, A. (1996). Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea arabica*. *Euphytica.* 87, 59–64. doi: 10.1007/BF0022965
- Pelletier, T. A., and Carstens, B. C. (2018). Geographical range size and latitude predict population genetic structure in a global survey. *Biol. Lett.* 14, 20170566. doi: 10.1098/rsbl.2017.0566
- Phiri, N.A. (2013). *Increasing the resilience of coffee production to leaf rust and other diseases in India and four African countries.* Final Technical Report- Project number CFC/ICO/4. p. 27–45.
- Scalabrini, S., Toniutti, L., Di Gaspero, G., Scaglione, D., Magrie, G., Vidotto, M., et al. (2020). A single polyploidization event at the origin of the tetraploid genome of *Coffea arabica* is responsible for the extremely low genetic variation in wild and cultivated germplasm. *Sci. Rep.* 10, 4642. (2020). doi: 10.1038/s41598-020-61216-7
- Sylvain, P. G. (1955). Some observations on *Coffea arabica* L. in Ethiopia. *Turrialba.* 5, 37–53.
- Vega, F. E., Ebert, A. W., and Ming, R. (2008). Coffee germplasm resources, genomics and breeding. *Plant Breed.* 30, 415–447. doi: 10.1002/9780470380130.ch9
- Walyaro, D. J., and Van der Vossen, H. A. M. (1977). Pollen longevity and artificial cross – pollination in *Coffea arabica* L. *Euphytica* 26, 225–231. doi: 10.1007/BF00032090

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Dulloo, Solano, Dessauw, Astorga and Guarino. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.